

# INVESTIGAÇÃO OPERACIONAL

Domingos Moreira Cardoso  
Universidade de Aveiro

Cidade da Praia, Cabo Verde  
Setembro de 2007



# Conteúdo

<b>Introdução</b>	<b>1</b>
<b>I Técnicas e Modelos de Optimização Linear</b>	<b>3</b>
<b>1 Modelos de Programação Linear</b>	<b>5</b>
1.1 Exemplos de programas lineares . . . . .	6
1.2 Formas padrão e canónica de programas lineares . . . . .	8
1.3 Resolução gráfica de programas lineares com duas variáveis . . . . .	11
1.4 Exercícios . . . . .	13
<b>2 Noções Elementares de Análise Convexa</b>	<b>17</b>
2.1 Conjuntos, cones e poliedros convexos . . . . .	17
2.2 Pontos Extremos e Direcções Extremas . . . . .	28
2.3 Relações entre pontos extremos e soluções básicas . . . . .	30
2.4 Direcções extremas e representação de poliedros convexos . . . . .	34
2.5 Exercícios . . . . .	39
<b>3 Método Simplex e Algumas Variantes</b>	<b>43</b>
3.1 O algoritmo simplex básico . . . . .	43
3.1.1 Descrição do método simplex . . . . .	49
3.1.2 Determinação de uma solução básica admissível inicial . . . . .	55
3.1.3 Determinação da inversa da base corrente . . . . .	60
3.2 Método simplex revisto . . . . .	61
3.3 Convergência do método simplex . . . . .	65
3.4 Variantes do método simplex . . . . .	70
3.4.1 Método simplex para variáveis limitadas . . . . .	70
3.4.2 Método de decomposição de Dantzig-Wolfe . . . . .	77
3.5 Exercícios . . . . .	84
<b>4 Dualidade e Análise Pós-optimal</b>	<b>89</b>
4.1 Resultados básicos sobre dualidade . . . . .	89
4.2 Método dual simplex . . . . .	95
4.3 Análise pós-optimal . . . . .	100

4.3.1	Modificação dos coeficientes da função objectivo . . . . .	101
4.3.2	Modificação dos termos independentes . . . . .	103
4.3.3	Modificação das entradas da matriz de decisão . . . . .	104
4.3.4	Adição de uma nova actividade . . . . .	106
4.3.5	Adição de uma nova restrição . . . . .	107
4.4	Cortes de Gomory em programação inteira . . . . .	109
4.5	Exercícios . . . . .	111
<b>II Técnicas e Modelos de Optimização Combinatória</b>		<b>113</b>
<b>5</b>	<b>Redes e Grafos</b>	<b>115</b>
5.1	Conceitos e resultados fundamentais . . . . .	115
5.2	Problemas métricos em grafos . . . . .	122
5.2.1	Problemas de caminho mais curto . . . . .	123
5.3	Grafos bipartidos, árvores e florestas . . . . .	127
5.3.1	Árvores abrangentes de custo mínimo . . . . .	129
5.4	Exercícios. . . . .	133
<b>6</b>	<b>Problemas Clássicos em Optimização Combinatória</b>	<b>137</b>
6.1	Poliedros combinatórios e teorema de König . . . . .	137
6.2	Conjuntos independentes e colorações . . . . .	139
6.2.1	Emparelhamentos em grafos bipartidos . . . . .	140
6.2.2	Estáveis e conjuntos de cobertura . . . . .	144
6.2.3	Colorações de vértices e arestas . . . . .	145
6.3	Circuitos de Euler e ciclos de Hamilton . . . . .	148
6.3.1	Grafos de Euler . . . . .	148
6.3.2	Grafos de Hamilton . . . . .	152
<b>Bibliografia</b>		<b>161</b>
<b>Índice</b>		<b>163</b>

# Introdução

Uma formação adequada em investigação operacional facilita a integração em equipas multi-disciplinares, de elevado desempenho, na implementação de projectos de grande complexidade e dimensão, como sejam, a determinação de rotas em sistemas de entrega de encomendas a longa distância, a definição de políticas de investimento energético, a produção de equipamentos de alta tecnologia produzidos a baixo custo, etc. Na verdade, são inúmeras as organizações e equipamentos cujos modelos e princípios de funcionamento são de natureza matemática. Vivemos numa sociedade dominada pela noção de produto industrial (i.e., objecto com uma função bem definida cuja existência se deve ao facto de se ter identificado um mercado para ele). O seu aparecimento, porém, exige um processo precedente de optimização de vários aspectos que vão desde a sua concepção até à sua comercialização, passando, naturalmente, pela respectiva produção. Em todo este percurso existem vários constrangimentos e é cada vez mais crítico reflectir sobre eles, modelar convenientemente os problemas associados, simular cenários alternativos, aplicar técnicas de análise apropriadas, etc. A aquisição de competências na área da modelação matemática de problemas faz-se, em geral, estudando modelos já existentes e analisando diferentes formulações. Este estudo e análise requer, porém, o conhecimento das propriedades que os objectos formalizados têm nos ambientes em que se inserem e, a partir delas, dos instrumentos e técnicas matemáticas de resolução ou simulação analítica e/ou combinatória. Muitos dos modelos matemáticos da Investigação Operacional são problemas de programação matemática que se designam, simplesmente, por programas matemáticos. Um programa matemático, consiste na determinação de  $x^* \in X$  tal que  $f(x^*) = \inf\{f(x) : x \in S\}$ , com  $S \subseteq X$ . Mais formalmente, os programas matemáticos exprimem-se como problemas de optimização do tipo:

$$(P) \quad \begin{array}{ll} \inf & f(x) \\ \text{s.a.} & x \in S, \end{array}$$

onde  $S$  denota o conjunto de pontos admissíveis, geralmente, representado por um conjunto de restrições,  $S = \{x \in X : g_i(x) \geq 0, i = 1, \dots, m\}$ , no qual, para cada  $i \in \{1, \dots, m\}$ ,  $g_i$  uma função real cujo domínio é  $X$  (i.e.,  $g_i : X \rightarrow \mathbb{R}$ ). Se o conjunto de pontos admissíveis é definido por um conjunto de restrições lineares e a função objectivo é linear (ou afim), então o problema designa-se por programa linear (problema de programação linear ou ainda problema de optimiza-

ção linear). No caso de qualquer das funções,  $f(x)$  ou  $g_i(x)$ , com  $i \in \{1, \dots, m\}$ , ser não linear,  $(P)$  designa-se por programa não linear (problema de programação não linear ou problema de optimização não linear). Quando  $S$  é um conjunto finito ou infinito numerável, o problema  $(P)$  designa-se por problema de optimização combinatoria ou de optimização discreta (nesta linha, no caso particular de  $S$  ser um subconjunto de  $n$ -uplos de inteiros, este problema, também se designa por problema de programação inteira ou de optimização inteira). Se  $S \neq \emptyset$ , então  $(P)$  diz-se consistente, caso contrário diz-se inconsistente.

Usualmente, os modelos matemáticos de investigação operacional dividem-se em lineares, não lineares e combinatorios. Este curso aborda, essencialmente, os modelos lineares e os combinatorios, por serem muito comuns na vida real.

Os tópicos em estudo dividem-se em duas partes, Técnicas e Modelos de Optimização Linear e Técnicas Modelos de Optimização Combinatória. A primeira parte inclui: Modelos de optimização linear (regras e princípios básicos de modelação, exemplos de problemas reais a que correspondem modelos de optimização linear, formas padrão e canónica de programas lineares e técnicas particulares de resolução de problemas simples); Método simplex e algumas variantes (pontos extremos de poliedros convexos e soluções básicas admissíveis, caracterização de direcções extremas e representação de poliedros, algoritmo simplex básico, determinação de uma solução básica admissível inicial, regras de prevenção de entrada em ciclo, método simplex para variáveis limitadas e método de decomposição de Dantzig-Wolfe); Dualidade e análise pós-optimal (método dual-simplex e método primal-dual, análise de sensibilidade a modificações nos diferentes dados do problema e cortes de Gomory e aplicações em programação inteira); Extensões das técnicas do tipo simplex (aplicação de técnicas do tipo simplex em programação paramétrica e na resolução de problemas de transporte e de afectação). A segunda parte inclui: Redes e Grafos (conceitos e resultados básicos, colorações de vértices e arestas, espaços vectoriais associados grafos, algoritmo de Disjkstra para problemas de caminho mais curto em grafos e algoritmos de Prim e Kruskal para a determinação de árvores abrangentes de custo mínimo); Circuitos de Euler e ciclos de Hamilton (circuitos de Euler, problema do carteiro chinês, ciclos de Hamilton, problema do caixeiro viajante (TSP) e método de Little (de "branch and bound") de aproximação ao TSP); Problemas de fluxo em redes (problemas de fluxo máximo, teorema do fluxo máximo corte mínimo, algoritmo de Ford e Fulkerson, problemas de fluxo de custo mínimo e algoritmo simplex para redes); Planeamento de projectos e problemas de sequenciamento (pert/cpm; sequenciamento de tarefas numa única máquina e em máquinas paralelas - modelos deterministas).

Parte I

Técnicas e Modelos de  
Optimização Linear



# Capítulo 1

## Modelos de Programação Linear

Muitos problemas de grande interesse prático podem ser formulados como programas lineares. Existem modelos de Programação Linear em áreas como a dos transportes aéreos (gestão de frotas, sequenciamento de tripulações, controlo de reservas, planeamento de tráfego em aeroportos), transportes terrestres (“routing”, localização de armazéns e/ou depósitos, sequenciamento e afectação de veículos), financeira (selecção e optimização de investimentos, gestão de “cash”, seguros), processos industriais (refinação de produtos, logística, optimização de misturas), gestão da produção (planeamento da produção, gestão de stocks, sequenciamento de tarefas), gestão e planeamento da produção agropecuária, definição de redes de transporte (de distribuição de energia eléctrica, de abastecimento de água, de gás natural), gestão de “staff” (sequenciamento de equipas médicas e/ou de apoio hospitalar), definição de “layout”, avaliação da exposição à radiação, entre muitas outras.

A arte da modelação matemática de problemas reais, requer não só uma adequada preparação matemática, como também um grande espírito de observação e elevada experiência. Questões como “O que é o problema”, “Quais os dados relevantes?”, “Como se relacionam?”, “Quais as decisões em jogo?” estão frequentemente associadas à modelação de situações complexas perante as quais é necessário tomar decisões. De acordo com os procedimentos usuais da *Investigação Operacional* (actividade científica dedicada à resolução de problemas relacionados com a tomada ou o apoio à tomada de decisões) um processo de tomada de decisão envolve, geralmente, os seguintes passos:

1. Observação da situação associada à decisão a tomar;
2. Reconhecimento e definição do problema;
3. Modelação conceptual e formulação matemática do problema;
4. Determinação de uma solução para o problema matemático obtido;

5. Tomada de decisão com base na solução encontrada;
6. Implementação da decisão e avaliação dos resultados.

## 1.1 Exemplos de programas lineares

Neste capítulo vamos procurar, recorrendo a vários exemplos, ilustrar unicamente a passagem do passo 2 (reconhecimento e definição do problema) para o passo 4 (modelação conceptual e formulação matemática do problema) centrandonos nas múltiplas situações que podem ser resolvidas recorrendo a modelos de Programação Linear. Seguem-se alguns exemplos de modelos de problemas Programação Linear.

**Exemplo 1.1** (problema da dieta óptima). *Um nutricionista pretende pôr em prática um tipo de alimentação que satisfaça certas exigências em termos dos nutrientes que pode fornecer, como sejam, proteínas, cálcio e vitaminas, com o menor custo possível. Assim, suponha que são necessárias  $b_1$  unidades de nutrientes do tipo  $N_1$ ,  $b_2$  unidades de nutrientes do tipo  $N_2, \dots$ , e  $b_m$  unidades de nutrientes do tipo  $N_m$ . Suponha ainda que estão disponíveis  $n$  tipos diferentes de alimentos e que  $a_{ij}$ , para  $i = 1, \dots, m$  e  $j = 1, \dots, n$ , corresponde ao número de unidades de nutriente do tipo  $N_i$  contidos em cada unidade de alimento do tipo  $j$ . Sendo  $x_1, x_2, \dots, x_n$ , respectivamente, a quantidade de alimento do tipo  $1, \dots, n$  e  $c_1, \dots, c_n$  os respectivos custos unitários de cada um dos tipos de alimento em causa, com a formulação que a seguir se indica*

$$\begin{array}{ll} \min & f(x) = \sum_{j=1}^n c_j x_j \\ \text{sujeito a} & \sum_{i=1}^m a_{ij} x_j \geq b_i, \quad \text{para } i = 1, \dots, m \\ & x_j \geq 0, \quad \text{para } j = 1, \dots, n, \end{array}$$

determina-se o (ou os)  $n$ -uplo(s) de números reais  $(x_1, x_2, \dots, x_n)$  que, de entre os que satisfazem as restrições indicadas, minimiza (ou minimizam) a função  $f(x)$ .

**Exemplo 1.2** (problema de transportes). *Uma dada companhia necessita abastecer  $n$  fábricas com um determinado produto, a partir de  $m$  armazéns. Suponha que, relativamente ao produto em causa, existem  $b_i$  unidades no  $i$ -ésimo armazém (com  $i \in \{1, \dots, m\}$ ) e que a  $j$ -ésima fábrica pretende adquirir  $a_j$  unidades (com  $j \in \{1, \dots, n\}$ ). Supondo ainda que a quantidade disponível, na totalidade dos armazéns, é igual à totalidade da quantidade pretendida pelas diferentes fábricas, ou seja,*

$$\sum_{i=1}^m b_i = \sum_{j=1}^n a_j$$

e que o custo do transporte de uma unidade de produto do  $i$ -ésimo armazém para a  $j$ -ésima fábrica é  $c_{ij}$  (para  $i = 1, \dots, m$  e  $j = 1, \dots, n$ ). Nestas condições, o problema em questão resume-se à determinação das quantidades de produto a transportar entre cada um dos armazéns e cada uma das fábricas de modo a que

este transporte se faça com o menor custo possível. Assim, sendo  $x_{ij}$  o número de unidades de produto a transportar entre o  $i$ -ésimo armazém e a  $j$ -ésima fábrica, o problema pode formular-se do seguinte modo:

$$\begin{aligned} \min \quad & \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \\ \text{sujeito a} \quad & \sum_{j=1}^n x_{ij} = b_i, \quad \text{para } i = 1, \dots, m \\ & \sum_{i=1}^m x_{ij} = a_j, \quad \text{para } j = 1, \dots, n \\ & x_{ij} \geq 0, \quad \text{para } i = 1, \dots, m \text{ e } j = 1, \dots, n. \end{aligned}$$

**Exemplo 1.3** (problema de otimização da produção). *Suponha que uma dada fábrica é capaz de produzir  $n$  produtos distintos utilizando  $m$  recursos limitados, os quais podem ser horas de trabalho ou tempos de operação de várias máquinas por semana, matérias primas, etc. Sendo  $c_j$  o lucro obtido por cada unidade do  $j$ -ésimo produto,  $b_i$  a quantidade de recurso  $i$  disponível e  $a_{ij}$  a quantidade de recurso  $i$  utilizada por unidade de produto  $j$ , a determinação do leque de produtos a produzir, bem como as suas diferentes quantidades com o máximo lucro, obtém-se resolvendo o problema:*

$$\begin{aligned} \max \quad & \sum_{j=1}^n c_j x_j \\ \text{sujeito a} \quad & \sum_{i=1}^m a_{ij} x_j \leq b_i, \quad \text{para } i = 1, \dots, m \\ & x_j \geq 0, \quad \text{para } j = 1, \dots, n. \end{aligned}$$

**Exemplo 1.4** (problema de gestão de “staff”). *Considere que num dado hospital existem 6 turnos diários em cada um dos quais o número mínimo de enfermeiras necessárias para assegurar o serviço é o indicado no quadro a seguir.*

Turno	1	2	3	4	5	6
Período	0-4h	4-8h	8-12h	12-16h	16-20h	20-24h
Número	8	10	12	10	8	6

*Supondo que as enfermeiras têm 8 horas diárias de trabalho (seguidas) e que iniciam sempre o seu trabalho no princípio de um turno, pretende-se determinar o menor número de enfermeiras de modo a que sejam garantidos os números mínimos indicados no quadro. Assim, sendo  $x_j$  (para  $j = 1, \dots, 6$ ) o número de enfermeiras que começam a trabalhar no  $j$ -ésimo turno, o problema em questão corresponde ao seguinte programa linear:*

$$\begin{aligned} \min \quad & x_1 + x_2 + x_3 + x_4 + x_5 + x_6 \\ \text{sujeito a} \quad & x_1 + x_6 \geq 8 \\ & x_1 + x_2 \geq 10 \\ & x_2 + x_3 \geq 12 \\ & x_3 + x_4 \geq 10 \\ & x_4 + x_5 \geq 8 \\ & x_5 + x_6 \geq 6 \\ & x_1, x_2, x_3, x_4, x_5, x_6 \geq 0. \end{aligned}$$

Embora, em rigor, se devesse também impor que as variáveis  $x_1, x_2, x_3, x_4, x_5$  e  $x_6$  fossem inteiras, neste caso, esta condição pode ser relaxada, uma vez que o problema indicado dá uma solução óptima inteira. Com efeito,  $x_1^* = 8, x_2^* = 4, x_3^* = 8, x_4^* = 2, x_5^* = 6$  e  $x_6^* = 0$  constitui uma solução óptima.

**Exemplo 1.5** (problema de inventário). Num dado período  $t$ , seja  $x_t$  a produção (número de artigos produzidos) de uma dada fábrica, seja  $I_t$  o inventário (número de artigos em “stock”) e  $p_t$  a procura (número de artigos pretendidos pelo mercado). Então, a relação geral entre estas variáveis para os diferentes períodos vem dada por:

$$x_t + I_{t-1} = p_t + I_t, \text{ para } t = 1, \dots, T.$$

Suponha que, para  $t = 1, 2$  e  $3$ , os custos unitários de produção são 15 euros, 30 euros e 40 euros, respectivamente, e que os custos unitários de armazenamento são de 10 euro por período. Supondo ainda que a procura nestes três períodos é de 10 000 artigos no 1º período, 20 000 no 2º e 25 000 no 3º, com o objectivo de minimizar os custos globais de modo a satisfazer a procura, dever-se-à resolver o seguinte programa linear:

$$\begin{array}{rcccccccc} \min & 15x_1 & + & 30x_2 & + & 40x_3 & + & 10I_1 & + & 10I_2 & + & 10I_3 \\ \text{sujeito a} & x_1 & & & & & & - & I_1 & & & = & 10\,000 \\ & & & x_2 & & & & + & I_1 & - & I_2 & = & 20\,000 \\ & & & & & x_3 & & & + & I_2 & - & I_3 & = & 25\,000 \\ & x_1, & & x_2, & & x_3, & & I_1, & & I_2, & & I_3 & \geq & 0. \end{array}$$

## 1.2 Formas padrão e canónica de programas lineares

Tipicamente, os problemas de programação linear apresentam-se na forma padrão ou na forma canónica e quando tal não acontece é possível transformá-los de modo tomarem uma destas formas.

**Definição 1.1.** Dada uma matriz  $A = [a_{ij}]_{i=1, \dots, m, j=1, \dots, n}$  (com  $m$  linhas e  $n$  colunas), um  $m$ -uplo de reais  $b^t = [b_1, \dots, b_m]$  e um  $n$ -uplo de reais  $c^T = [c_1, \dots, c_n]$ , designa-se por programa linear de minimização (maximização) na forma padrão (ou standard), e denota-se por  $PL(A, b, c)$  de minimização (maximização) na forma padrão, o problema de determinação de um  $n$ -uplo de reais  $x^T = [x_1, \dots, x_n]$  que minimiza (maximiza) a função objectivo  $f(x) = c_1x_1 + c_2x_2 + \dots + c_nx_n$ , no conjunto de pontos de  $\mathbb{R}^n$  que satisfazem as restrições:

$$\begin{array}{cccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1, \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = & b_2, \\ \vdots & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = & b_m, \end{array}$$

$$\text{e } x_j \geq 0, \text{ para } j = 1, 2, \dots, n.$$

Numa forma mais condensada, este problema corresponde a determinar o

$$\begin{array}{ll} \min(\max) & c^T x \\ \text{sujeito a} & Ax = b \\ & x \geq 0. \end{array}$$

**Definição 1.2.** Dada a matriz  $A = [a_{ij}]_{i=1,\dots,m, j=1,\dots,n}$  (com  $m$  linhas e  $n$  colunas), um  $m$ -uplo de reais  $b^T = [b_1, \dots, b_m]$  e um  $n$ -uplo de reais  $c^T = [c_1, \dots, c_n]$ , designa-se por programa linear de minimização (maximização) na forma canônica, e denota-se por  $PL(A, b, c)$  de minimização (maximização) na forma canônica, o problema de determinação de um  $n$ -uplo de reais  $x^T = [x_1, \dots, x_n]$  que minimiza (maximiza) a função objectivo  $f(x) = c_1x_1 + c_2x_2 + \dots + c_nx_n$ , no conjunto de pontos de  $\mathbb{R}^n$  que satisfazem as restrições:

$$\begin{array}{l|l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \geq b_1 & (a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \leq b_1), \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \geq b_2 & (a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \leq b_2), \\ \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \geq b_m & (a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \leq b_m), \end{array}$$

$$\text{e } x_j \geq 0, \text{ para } j = 1, 2, \dots, n.$$

Numa forma mais condensada, o programa linear  $PL(A, b, c)$  na forma canônica corresponde, respectivamente, a determinar o

$$\begin{array}{l|l} \min & c^T x \\ \text{sujeito a} & Ax \geq b \\ & x \geq 0 \end{array} \quad \left| \quad \begin{array}{l} \max & c^T x \\ \text{sujeito a} & Ax \leq b \\ & x \geq 0. \end{array} \right.$$

Em ambas as definições as entradas da matriz  $A$  (coeficientes  $a_{ij}$ ) designam-se por coeficientes tecnológicos.

De acordo com a definição anterior, num programa linear (ou problema de programação linear) há que destacar as **variáveis de decisão**, denotadas por  $x_1, \dots, x_n$ , a **função objectivo** denotada por  $c^T x$  (ou  $f(x) = c_1x_1 + c_2x_2 + \dots + c_nx_n$ ), as restrições que, embora se possam (conforme definição 1.1) apresentar a forma de equações, na definição 1.2 (que se refere unicamente à forma canônica) são constituídas por inequações do tipo “ $\leq$ ” ou “ $\geq$ ” e as condições de não negatividade das variáveis.

A especificação de um problema de programação linear (optimização linear),  $PL(A, b, c)$ , na forma canônica ou padrão, faz-se, sinteticamente, adoptando uma das formulações a seguir indicadas.

	(maximização)	(minimização)
(forma padrão)	$\begin{array}{ll} \max & c^T x \\ \text{s. a.} & Ax = b \\ & x \geq 0 \end{array}$	$\begin{array}{ll} \min & c^T x \\ \text{s. a.} & Ax = b \\ & x \geq 0 \end{array}$
(forma canónica)	$\begin{array}{ll} \max & c^T x \\ \text{s. a.} & Ax \leq b \\ & x \geq 0 \end{array}$	$\begin{array}{ll} \max & c^T x \\ \text{s. a.} & Ax \geq b \\ & x \geq 0 \end{array}$

Qualquer  $n$ -uplo de reais que satisfaça o conjunto de restrições do  $PL(A, b, c)$  diz-se uma solução admissível e o conjunto de todas as soluções admissíveis designa-se por conjunto ou região admissível. Por outro lado,  $x^*$  diz-se uma solução óptima para o  $PL(A, b, c)$  de minimização (maximização) se

$$c^T x^* = \min (\max) \{c^T x : x \text{ é solução admissível para } PL(A, b, c)\}.$$

Note-se ainda que o facto de estarmos a considerar um problema de maximização ou minimização é irrelevante uma vez que, relativamente à solução ou conjunto de soluções óptimas a determinar, o problema de minimização  $\min\{c^T x : x \in \mathbb{R}^n, Ax = b \text{ e } x \geq 0\}$  é equivalente ao problema de maximização  $PL(A, b, -c)$ . Com efeito, se  $x^*$  é admissível para  $PL(A, b, c)$  e é tal que  $-c^T x^* \geq -c^T x$  para todo o  $x$  admissível para  $PL(A, b, -c)$ , então  $c^T x^* \leq c^T x$  para todo o  $x$  admissível para o  $PL(A, b, c)$ . As duas formas (canónica ou padrão) em que um programa linear é apresentado são completamente equivalentes, uma vez que é possível passar de uma para outra sem que o conjunto de soluções se altere. Com efeito, a partir de uma restrição do tipo

$$\sum_{j=1}^n a_{ij} x_j \leq b_i \Leftrightarrow b_i - \sum_{j=1}^n a_{ij} x_j \geq 0 \Leftrightarrow x_{n+i} = b_i - \sum_{j=1}^n a_{ij} x_j, \text{ com } x_{n+i} \geq 0,$$

acrescentando-se a variável de desvio (ou folga)  $x_{n+i}$ , obtém-se a restrição de igualdade

$$\sum_{j=1}^n a_{ij} x_j + x_{n+i} = b_i, \text{ com } x_{n+i} \geq 0.$$

De modo a que se faça a distinção entre as variáveis originais  $(x_1, \dots, x_n)$  e as variáveis de desvio (ou folga) as primeiras designam-se por variáveis de decisão.

Deve observar-se ainda que à variável de desvio pode, em geral, dar-se uma certa interpretação económica ou física, consoante o problema. Com efeito, se no Exemplo 1.3 passarmos o problema para a forma padrão, o valor que a variável de desvio  $x_{n+i}$  toma, para uma dada solução admissível corrente, corresponde à quantidade de  $i$ -ésimo recurso que não é utilizada na referida solução.

No caso de se ter uma restrição de igualdade, esta é facilmente convertida em duas restrições de desigualdade, uma vez que

$$\sum_{j=1}^n a_{ij}x_j = b_i \Leftrightarrow \begin{cases} a_{i1}x_1 + \dots + a_{in}x_n \leq b_i \\ a_{i1}x_1 + \dots + a_{in}x_n \geq b_i \end{cases} \Leftrightarrow \begin{cases} a_{i1}x_1 + \dots + a_{in}x_n \leq b_i \\ -a_{i1}x_1 - \dots - a_{in}x_n \leq -b_i \end{cases}$$

Por outro lado, qualquer variável livre, i.e., qualquer variável não restringida (aos reais não negativos),  $x_j$ , pode converter-se num par de variáveis não negativas,  $x'_j \geq 0$  e  $x''_j \geq 0$ , escrevendo-se  $x_j = x'_j - x''_j$ . Deste modo, dado um sistema de equações  $Ax = b$ , em que as variáveis são não restringidas, fazendo  $x = x' - x''$  com  $x' \geq 0, x'' \geq 0$ , obtém-se o sistema com variáveis restringidas:

$$A(x' - x'') = b \Leftrightarrow Ax' - Ax'' = b \Leftrightarrow [A, -A] \begin{bmatrix} x' \\ x'' \end{bmatrix} = b,$$

ou seja,  $\bar{A}y = b$  com  $\bar{A} = [A, -A]$  e  $y = \begin{bmatrix} x' \\ x'' \end{bmatrix} \geq 0$ .

### 1.3 Resolução gráfica de programas lineares com duas variáveis

Os problemas de programação linear com duas variáveis podem resolver-se graficamente, uma vez que, neste caso, a função objectivo define uma recta que pode ser deslocada no sentido do seu gradiente, até se tornar tangente à região admissível. Os pontos de tangência obtidos (um, nenhum, ou uma infinidade) correspondem aos pontos da região admissível que optimizam a função objectivo.

Os dois exemplos que a seguir se indicam, onde, por intermédio das respectivas representações gráficas, no primeiro se determina a solução óptima e no segundo se conclui a ausência de óptimo finito, ilustram o modo de actuar nestes casos.

**Exemplo 1.6.**

$$\begin{aligned} \max \quad & x_1 + x_2 \\ \text{s. a.} \quad & -x_1 + x_2 \leq 1 \\ & x_1 - 2x_2 \leq 2 \\ & x_2 \leq 2 \\ & x_1, x_2 \geq 0 \end{aligned}$$

Deve observar-se que, após a introdução das variáveis de desvio, a região admissível fica definida pelo conjunto

$$\left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} \in \mathbb{R}^5 : \begin{bmatrix} -1 & 1 & 1 & 0 & 0 \\ 1 & -2 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, x_1, x_2, x_3, x_4, x_5 \geq 0 \right\}$$

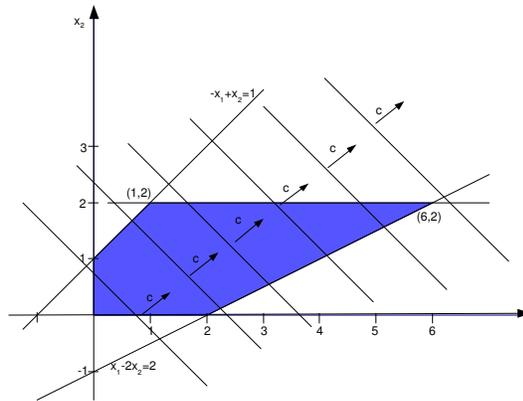


Figura 1.1: Aplicação do método gráfico a um problema com ótimo finito.

e, por exemplo, os pontos  $(1, 2)$  e  $(6, 2)$  correspondem, respectivamente, aos pontos  $(1, 2, 0, 5, 0)^T$  e  $(6, 2, 5, 0, 0)^T$ .

**Exemplo 1.7.**

$$\begin{aligned} \max \quad & 2x_1 + x_2 \\ \text{s. a.} \quad & -x_1 + x_2 \leq 1 \\ & x_1 - 2x_2 \leq 2 \\ & x_1, x_2 \geq 0 \end{aligned}$$

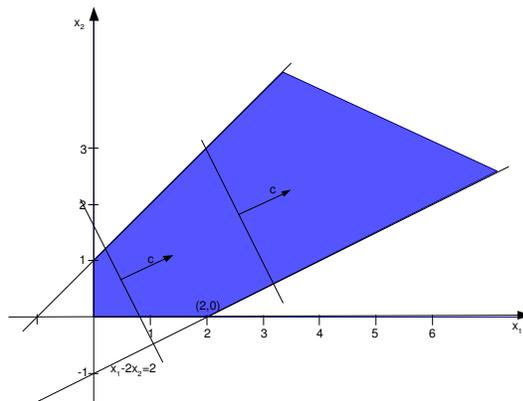


Figura 1.2: Aplicação do método gráfico a um problema sem ótimo finito.

Neste caso, transformando o problema na forma padrão, obtém-se

$$\max\{2x_1 + x_2 : \begin{bmatrix} -1 & 1 & 1 & 0 \\ 1 & -2 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, x_1, x_2, x_3, x_4 \geq 0\}$$

e, a partir da base associada ao ponto  $(2, 0)$  que, após reordenação das variáveis,

corresponde ao ponto  $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 0 \\ 0 \end{bmatrix}$  e tendo em conta que a base correspon-

dente é  $B = \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix}$ , obtêm-se as direcções linearmente independentes do subespaço nula da matriz de decisão

$$d' = \begin{bmatrix} d'_B \\ d'_3 \\ d'_4 \end{bmatrix} = \begin{bmatrix} -B^{-1}a_2 \\ 1 \\ 0 \end{bmatrix} \text{ e } d'' = \begin{bmatrix} d''_B \\ d''_3 \\ d''_4 \end{bmatrix} = \begin{bmatrix} -B^{-1}a_4 \\ 0 \\ 1 \end{bmatrix}.$$

Tendo em conta que  $B^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ ,  $a_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$  e  $a_4 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ , vem que  $d' =$

$$\begin{bmatrix} 2 \\ 1 \\ 1 \\ 0 \end{bmatrix} \text{ e } d'' = \begin{bmatrix} -1 \\ -1 \\ 0 \\ 1 \end{bmatrix}. \text{ Logo, podemos concluir que } d' \text{ constitui uma direcção}$$

extrema para a região admissível do problema em causa para a qual, tendo em conta que o gradiente da função objectivo é  $c^T = [c_B, c_N] = [2, 0, 1, 0]$ , se verifica que  $c^T d' = 5 > 0$ . Consequentemente,  $d'$  é uma direcção extrema responsável pela inexistência de óptimo finito (tal como se observa graficamente).

## 1.4 Exercícios

**Exercício 1.1.** *Considere uma fábrica em que para a produção de camisas, calças e casacos são executadas operações de corte, costura e empacotamento. Considere ainda que os tempos semanais disponíveis (em minutos) para as diferentes operações, os tempos requeridos por cada um dos produtos e os respectivos lucros unitários correspondem aos valores apresentados no quadro a seguir.*

	Tempos (min.)/unidade produzida			lucros unitários
	corte	costura	empacotar.	
camisas	1	4	1	3 euros
calças	2	5	1	7 euros
casacos	1	10	2	8 euros
tempos disp.	9 000	38 000	6 000	

Apresente uma formulação de Programação Linear, sabendo que se pretende determinar a produção semanal que maximiza o lucro.

**Exercício 1.2.** *Numa unidade agro-pecuária criam-se galinhas com rações obtidas à custa da mistura de milho, lima e alfalfa de modo a satisfazer os valores*

mínimos diários em proteínas, hidratos de carbono e vitaminas que se indicam no quadro a seguir. Nesse quadro, são também indicados os valores dos nutrientes fornecidos por cada unidade das substâncias que compõem a ração bem como os respectivos custos.

Nutrientes	Ingredientes			Valores
	Milho	Lima	Alfalfa	Mínimos Diários
Proteínas	2	4	4	10
Hidratos de carbono	4	2	4	6
Vitaminas	2	3	4	5
Custos/Unitários	10	6	7	

Sabendo que se pretende determinar a mistura de custo mínimo, apresente uma formulação de Programação Linear para este problema.

**Exercício 1.3.** Considere que uma dada fábrica de brinquedos pretende lançar no mercado uma boneca, um carro e um “puzzle”, com produções que envolvem a utilização de 3 tipos de máquinas (A, B e C) de tal forma que, no caso da boneca, existem duas linhas de fabrico alternativas (L1 e L2) as quais não utilizam os mesmos tipos de máquinas. Considere ainda que os tempos consumidos por cada artigo, nas diferentes máquinas e (para a boneca) segundo as diferentes linhas de produção, o número diário de horas disponíveis por tipo de máquina e os lucros unitários obtidos são os que se indicam.

	Artigos				n° de horas
	boneca(L1)	boneca(L2)	carro	“puzzle”	disponíveis
Maq. A	2 h.	0 h.	2 h.	4 h.	100 h.
Maq. B	0 h.	4 h.	5 h.	2 h.	200 h.
Maq. C	6 h.	4 h.	3 h.	6 h.	250 h.
Lucros Unitários	4	4	3	8	

Apresente uma formulação de Programação Linear, sabendo que se pretende determinar a produção diária (n° de bonecas segundo cada uma das linhas de produção, n° de carros e n° de “puzzles”) de maior lucro.

**Exercício 1.4.** Suponha que uma dada empresa produz dois tipos de artigos A e B, os quais, na sua fabricação, ocupam necessariamente as máquinas I e II. Suponha ainda que os tempos disponíveis para a produção nestas máquinas, bem como os tempos necessários (em cada máquina) para a produção de uma unidade dos diferentes artigos são os indicados na tabela a seguir.

	Artigo A	Artigo B	Nº de Horas Disponíveis
Máquina I	2 h.	4 h.	720 h.
Máquina II	4 h.	4 h.	880 h.

Sabendo ainda que os lucros unitários para os artigos A e B são, respectivamente, 4 e 2 euros, determine a produção (nº de artigos do tipo A e nº de artigos do tipo B) que maximiza o lucro total.

**Exercício 1.5.** Considere cada um dos problemas de Programação Linear a seguir indicados

$$\begin{array}{ll} \max & 2x_1 + 5x_2 \\ \text{s. a.} & 2x_1 + 3x_2 \geq 6 \\ & x_1 + x_2 \leq 4 \\ & x_1, x_2 \geq 0 \end{array}$$

$$\begin{array}{ll} \min & 2x_1 - x_2 \\ \text{s. a.} & -2x_1 + x_2 \leq 1 \\ & x_1 + 2x_2 \geq 4 \\ & x_1, x_2 \geq 0 \end{array}$$

e resolva-os por aplicação do método gráfico.



## Capítulo 2

# Noções Elementares de Análise Convexa

Em termos gerais, podemos afirmar que os problemas de optimização linear se caracterizam pela determinação do melhor (ou um dos melhores) ponto(s) de um poliedro convexo para uma certa funcional linear (função objectivo). Este facto justifica que se estudem, com algum detalhe, os conjuntos convexos e se analisem algumas das suas propriedades.

### 2.1 Conjuntos, cones e poliedros convexos

**Definição 2.1.** *Dado um conjunto de pontos de  $\mathbb{R}^n$ ,  $\{x^1, \dots, x^p\}$ , e um conjunto de escalares,  $\{\lambda_1, \dots, \lambda_p\}$ , se  $\sum_{i=1}^p \lambda_i = 1$ , então  $\sum_{i=1}^p \lambda_i x^i$  designa-se por combinação afim dos pontos  $x^1, \dots, x^p$ . Adicionalmente, se  $\lambda_i \geq 0 \forall i \in \{1, \dots, p\}$ , então  $\sum_{i=1}^p \lambda_i x^i$  diz-se uma combinação linear convexa dos pontos  $x^1, \dots, x^p$ . Um conjunto  $S \subseteq \mathbb{R}^n$  diz-se convexo se*

$$\forall x, y \in S \forall \lambda \in [0, 1], \lambda x + (1 - \lambda)y \in S$$

(i.e., se toda a combinação linear convexa de quaisquer dois pontos de  $S$  está em  $S$ ).

Dados dois pontos de um conjunto convexo  $S$ ,  $x$  e  $y$ , o conjunto de combinações lineares convexas de  $x$  e  $y$ ,  $\{z = \lambda x + (1 - \lambda)y : 0 \leq \lambda \leq 1\}$ , designa-se por segmento que une  $x$  e  $y$  e denota-se por  $\overline{xy}$ .

Da Definição 2.1 decorre a seguinte definição alternativa de conjunto convexo: um conjunto  $S$  é convexo sse, para  $p \geq 2$ ,

$$\forall x^1, \dots, x^p \in S \forall \lambda_1, \dots, \lambda_p \geq 0 : \sum_{i=1}^p \lambda_i = 1, \sum_{i=1}^p \lambda_i x^i \in S.$$

Uma vez que é imediato verificar que a condição é suficiente, apenas vamos provar que se trata de uma condição necessária. Assim, sendo  $S$  um conjunto

convexo, para  $p = 2$ , a condição é imediatamente satisfeita (por aplicação da Definição 2.1). Suponha-se que a condição é verdadeira para  $p \leq n$  e sejam  $x^1, \dots, x^{n+1}$  pontos de  $S$  e  $\lambda_1, \dots, \lambda_{n+1}$  escalares não negativos tais que  $\sum_{i=1}^{n+1} \lambda_i = 1$ . Se  $\lambda_{n+1} \in \{0, 1\}$ , é claro que  $\sum_{i=1}^{n+1} \lambda_i x^i \in S$ . Assim, vamos supor que  $0 < \lambda_{n+1} < 1$ . Em tais condições,  $\sum_{i=1}^n \lambda_i > 0$  e, tendo em conta que  $\frac{\lambda_i}{\sum_{i=1}^n \lambda_i} \geq 0 \quad \forall i \in \{1, \dots, n\}$  e  $\sum_{i=1}^n \frac{\lambda_i}{\sum_{i=1}^n \lambda_i} = 1$ , fazendo

$$\sum_{i=1}^{n+1} \lambda_i x^i = \sum_{i=1}^n \lambda_i x^i + \lambda_{n+1} x^{n+1} = \left( \sum_{i=1}^n \lambda_i \right) \sum_{i=1}^n \frac{\lambda_i}{\left( \sum_{i=1}^n \lambda_i \right)} x^i + \lambda_{n+1} x^{n+1},$$

da hipótese decorre que  $x = \sum_{i=1}^n \frac{\lambda_i}{\sum_{i=1}^n \lambda_i} x^i \in S$ . Logo, tendo em conta que  $\sum_{i=1}^n \lambda_i \geq 0$ ,  $\lambda_{n+1} \geq 0$  e  $\sum_{i=1}^n \lambda_i + \lambda_{n+1} = 1$ , obtém-se

$$\sum_{i=1}^{n+1} \lambda_i x^i = \sum_{i=1}^n \lambda_i x^i + \lambda_{n+1} x^{n+1} \in S,$$

o que completa a prova pretendida.

Igualmente se prova que a intersecção de uma família de conjuntos convexos é um conjunto convexo (ver Exercício 2.1).

**Definição 2.2.** *Um conjunto de pontos  $\{x^1, \dots, x^p\}$  diz-se dependente afim se existem escalares  $\lambda_1, \dots, \lambda_p$ , não todos nulos, tais que  $\sum_{i=1}^p \lambda_i = 0$  e  $\sum_{i=1}^p \lambda_i x^i = 0$ . Caso contrário o conjunto  $\{x^1, \dots, x^p\}$  diz-se independente afim.*

De acordo com esta definição, se um dado ponto  $x$  é combinação linear convexa de um conjunto de pontos  $\{x^1, \dots, x^p\}$ , então o conjunto  $\{x\} \cup \{x^1, \dots, x^p\}$  é dependente afim. Com efeito, se  $\exists \lambda_1, \dots, \lambda_p \geq 0$  tais que  $\sum_{i=1}^p \lambda_i = 1$  e  $x = \sum_{i=1}^p \lambda_i x^i$ , então  $x - \lambda_1 x^1 - \dots - \lambda_p x^p = 0$  e, além disso,  $1 - \lambda_1 - \lambda_2 - \dots - \lambda_p = 0$ . Podemos ainda afirmar que um conjunto de  $p + 1$  pontos  $\{x^0, \dots, x^p\}$  é independente afim sse o conjunto de  $p$  vectores  $\{x^1 - x^0, \dots, x^p - x^0\}$  é linearmente independente (como exercício, faça a prova desta afirmação).

**Teorema 2.1.** *Seja  $S = \{x^0, \dots, x^k\}$  um conjunto de pontos independentes afim. Então cada ponto  $x$  do conjunto das combinações afim de pontos de  $S$  tem uma representação única como combinação afim dos pontos  $x^0, \dots, x^k$ .*

**Demonstração.** Seja  $x = \sum_{i=1}^k \lambda_i x^i$ , com  $\sum_{i=1}^k \lambda_i = 1$ , e suponha-se que também se verifica a equação  $x = \sum_{i=1}^k \alpha_i x^i$ . Então  $x - x = \sum_{i=1}^k (\lambda_i - \alpha_i) x^i = 0$  e  $\sum_{i=1}^k \lambda_i - \alpha_i = 0$ . Logo, uma vez que os pontos  $x^0, \dots, x^k$  são independentes afim, conclui-se que  $\lambda_i = \alpha_i$ , para  $i = 0, \dots, k$ . □

De um modo análogo se prova que sendo  $S = \{x^0, \dots, x^k\}$  um conjunto de pontos independentes afim, então cada ponto  $x$  do conjunto das combinações lineares convexas de pontos de  $S$  tem uma representação única como combinação linear convexa dos pontos  $x^0, \dots, x^k$ .

**Definição 2.3.** Designa-se por *variedade linear* (ou *subespaço afim*), toda a translação de um subespaço vectorial, i.e., dado um subespaço vectorial de  $\mathbb{R}^n$ ,  $E$ , e um ponto  $x \in \mathbb{R}^n$ ,  $x + E$  designa-se por *variedade linear* (ou *subespaço afim*). A dimensão da variedade linear  $U = x + E$  denota-se por  $\dim(U)$  e corresponde à dimensão de  $E$ . Dado um conjunto de pontos  $S \subseteq \mathbb{R}^n$ , designa-se por *dimensão de  $S$*  e denota-se por  $\dim(S)$ , a dimensão da variedade linear de menor dimensão que contém  $S$ .

Como consequência desta definição conclui-se que um ponto é uma variedade linear de dimensão 0.

**Definição 2.4.** Dado um conjunto de pontos  $S$ , designa-se por *invólucro convexo de  $S$*  e denota-se por  $H(S)$  a intersecção de todos os conjuntos convexos que contêm  $S$  e designa-se por *invólucro afim de  $S$*  e denota-se por  $Aff(S)$  a intersecção de todas as variedades lineares (subespaços afim) que contêm  $S$ .

Como consequência desta definição vem que um conjunto  $S$  é convexo sse  $S = H(S)$ .

1. Seja  $S$  um conjunto convexo. Pela definição de  $H(S)$ , (i)  $H(S) \supseteq S$ . Por outro lado, qualquer que seja o conjunto convexo  $C$  que contenha  $S$ ,  $C \supseteq H(S)$ . Logo, uma vez que  $S$  é convexo e  $S \supseteq S$ , conclui-se que (ii)  $S \supseteq H(S)$ . Tendo em conta (i) e (ii) vem que  $S = H(S)$ .
2. Suponha-se que  $S$  não é convexo. Uma vez que  $H(S)$  é a intersecção de conjuntos convexos, pode concluir-se que  $H(S)$  é convexo. Logo  $S \neq H(S)$ .

De um modo idêntico ao anterior se conclui que  $S$  é uma variedade linear sse  $S = Aff(S)$ .

Se  $S$  é um conjunto arbitrário de pontos no qual existe um conjunto maximal de  $k + 1$  pontos independentes afim  $\{x^0, x^1, \dots, x^k\}$  (maximal no sentido em que  $\forall x \in S \{x^0, x^1, \dots, x^k\} \cup \{x\}$  é um conjunto dependente afim), sendo  $E$  o subespaço vectorial gerado pelo conjunto de vectores  $\{x^1 - x^0, \dots, x^k - x^0\}$ , então pode concluir-se que  $x^0 + E$  é a variedade linear de menor dimensão que contém  $S$ , ou seja,  $x^0 + E = Aff(S)$ . Note-se que a variedade linear de menor dimensão que contém  $S$  corresponde à intersecção de todas as variedades lineares que contêm  $S$ .

**Teorema 2.2.** Qualquer que seja o conjunto  $S$ ,  $H(S)$  coincide com o conjunto de todas as combinações lineares convexas de pontos de  $S$  e  $Aff(S)$  coincide com o conjunto de todas as combinações afim de pontos de  $S$ .

**Demonstração.** Seja  $T$  o conjunto de todas as combinações lineares convexas de pontos de  $S$ . Dado que  $H(S)$  é convexo<sup>1</sup> e  $S \subseteq H(S)$ , então podemos concluir que as combinações lineares convexas de pontos de  $S$  estão em  $H(S)$ , ou seja, que

$$H(S) \supseteq T. \tag{2.1}$$

---

<sup>1</sup>Note-se que  $H(S)$  é intersecção de conjuntos convexos.

Por outro lado,  $\forall x, y \in T$ ,  $\exists x^1, \dots, x^p \in S$ ,  $\exists y^1, \dots, y^q \in S$ ,  $\exists \lambda_1, \dots, \lambda_p \geq 0$ , com  $\sum_{j=1}^p \lambda_j = 1$ , e  $\exists \beta_1, \dots, \beta_q \geq 0$ , com  $\sum_{i=1}^q \beta_i = 1$ , tais que  $x = \sum_{j=1}^p \lambda_j x^j$  e  $y = \sum_{i=1}^q \beta_i y^i$ . Logo,  $\forall \alpha \in [0, 1]$   $\alpha x + (1 - \alpha)y = \sum_{j=1}^p \alpha \lambda_j x^j + \sum_{i=1}^q (1 - \alpha) \beta_i y^i$ , com  $\alpha \lambda_j \geq 0 \forall j \in \{1, \dots, p\}$ ,  $(1 - \alpha) \beta_i \geq 0 \forall i \in \{1, \dots, q\}$  e

$$\sum_{j=1}^p \alpha \lambda_j + \sum_{i=1}^q (1 - \alpha) \beta_i = \alpha \sum_{j=1}^p \lambda_j + (1 - \alpha) \sum_{i=1}^q \beta_i = 1.$$

Consequentemente, conclui-se que  $T$  é um conjunto convexo que contém  $S$ , donde

$$H(S) \subseteq T. \quad (2.2)$$

De (2.1) e (2.2) vem finalmente que  $H(S) = T$ . A prova de que  $Aff(S)$  coincide com o conjunto de todas as combinações afim de pontos de  $S$  é idêntica à anterior.  $\square$

O teorema que se segue, conhecido por Teorema de Caratheodory, constitui um dos resultados mais relevantes em *Análise Convexa*.

**Teorema 2.3** (de Caratheodory). *Se  $S$  é um subconjunto não vazio de  $\mathbb{R}^n$ , então  $\forall x \in H(S)$ ,  $\exists x^1, \dots, x^p \in S$ , com  $p \leq n + 1$ , tais que  $x \in H(\{x^1, \dots, x^p\})$  (ou seja,  $x$  pode exprimir-se como combinação linear convexa de não mais do que  $n + 1$  pontos de  $S$ ).*

**Demonstração.** Seja  $x \in H(S)$ , então o Teorema 2.2 garante que  $x$  pertence ao conjunto das combinações lineares convexas de pontos de  $S$ , i.e.,  $\exists x^1, \dots, x^p \in S$  e  $\exists \lambda_1, \dots, \lambda_p \geq 0$ , com  $\sum_{i=1}^p \lambda_i = 1$ , tais que  $x = \sum_{i=1}^p \lambda_i x^i$ . Resta provar que é possível escolher os pontos de modo que  $p$  seja não superior a  $n + 1$ . Se  $p > n + 1$ , então os pontos  $x^1, \dots, x^p$  são dependentes afim. Logo, existem escalares  $\alpha_1, \dots, \alpha_p$ , não todos nulos, tais que  $\sum_{i=1}^p \alpha_i = 0$  e  $\sum_{i=1}^p \alpha_i x^i = 0$ . Consequentemente, tendo em conta que  $\{\alpha_i > 0 : 1 \leq i \leq p\} \neq \emptyset$ <sup>2</sup>, fazendo  $\theta = \min\{\frac{\lambda_i}{\alpha_i} : \alpha_i > 0, 1 \leq i \leq p\}$ ,  $\beta_i = \lambda_i - \theta \alpha_i$ , para  $i = 1, \dots, p$ , e supondo  $\theta = \frac{\lambda_s}{\alpha_s}$ , obtém-se

$$\sum_{i=1}^p \beta_i x^i = \sum_{i=1}^p (\lambda_i - \theta \alpha_i) x^i = \sum_{i=1}^p \lambda_i x^i - \theta \sum_{i=1}^p \alpha_i x^i = x,$$

com  $\beta_i \geq 0 \forall i \in \{1, \dots, p\}$ <sup>3</sup>,  $\beta_s = \lambda_s - \frac{\lambda_s}{\alpha_s} \alpha_s = 0$  e

$$\sum_{i=1}^p \beta_i = \sum_{i=1}^p (\lambda_i - \theta \alpha_i) = \sum_{i=1}^p \lambda_i - \theta \sum_{i=1}^p \alpha_i = 1.$$

<sup>2</sup>Observe-se que uma vez que os escalares  $\alpha_1, \dots, \alpha_p$ , não são todos nulos, da equação  $\sum_{i=1}^p \alpha_i = 0$  retira-se que existem pelo menos dois escalares de sinal contrário.

<sup>3</sup>Com efeito, dado que  $\theta > 0$ ,  $\forall i \in \{1, \dots, p\}$  se  $\alpha_i \leq 0$  então  $\lambda_i - \theta \alpha_i \geq 0$  e se  $\alpha_i > 0$  então  $\theta \leq \frac{\lambda_i}{\alpha_i} \Leftrightarrow \lambda_i - \theta \alpha_i \geq 0$ .

Deste modo, exprime-se  $x$  como combinação linear convexa de  $p - 1$  pontos de  $S$ . Se  $p - 1 \leq n + 1$ , então obtém-se o que se pretendia, caso contrário este processo repete-se até que tal aconteça.  $\square$

**Definição 2.5.** *O invólucro convexo de um conjunto finito de pontos designa-se por polítopo. Se  $S = \{x^0, \dots, x^k\}$  é um conjunto independente afim, então  $H(S)$  designa-se por simplex de dimensão  $k$  (ou  $k$ -dimensional) e os pontos  $x^0, \dots, x^k$ , designam-se por vértices.*

Na figura a seguir representam-se dois conjuntos simpliciais de  $\mathbb{R}^2$ .

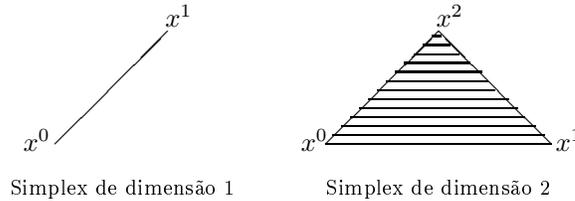


Figura 2.1: Exemplos de conjuntos simpliciais de  $\mathbb{R}^2$ .

Os cones e os poliedros convexos desempenham um papel relevante em otimização linear.

**Definição 2.6.** *Uma aplicação,  $f : \mathbb{R}^n \mapsto \mathbb{R}^m$ , diz-se linear se é aditiva e homogénea, ou seja, se*

$$\forall u, v \in \mathbb{R}^n \forall \lambda \in \mathbb{R} \quad f(u + v) = f(u) + f(v) \quad \wedge \quad f(\lambda u) = \lambda f(u).$$

*Uma aplicação linear de  $\mathbb{R}^n$  em  $\mathbb{R}$  diz-se uma funcional linear de  $\mathbb{R}^n$ .*

Esta definição dá origem ao conceito de hiperplano e de meio espaço (ou semi-espaço). Sendo  $f$  uma funcional linear de  $\mathbb{R}^n$  não identicamente nula, i.e., para qual  $\exists u \in \mathbb{R}^n$  tal que  $f(u) \neq 0$ ,  $H_\alpha = \{x \in \mathbb{R}^n : f(x) = \alpha\}$  determina um hiperplano (variedade linear de dimensão  $n - 1$ ) que divide  $\mathbb{R}^n$  nos dois meios espaços:

$$H_\alpha^+ = \{x \in \mathbb{R}^n : f(x) \geq \alpha\} \quad \text{e} \quad H_\alpha^- = \{x \in \mathbb{R}^n : f(x) \leq \alpha\},$$

no sentido em que  $\mathbb{R}^n = H_\alpha^+ \cup H_\alpha^-$  e  $H_\alpha = H_\alpha^+ \cap H_\alpha^-$ .

Qualquer equação do tipo  $a^T x = \alpha$ , com  $a \in \mathbb{R}^n \setminus \{0\}$  e  $\alpha \in \mathbb{R}$  (equação linear), pode ser representada pelo sistema equivalente de inequações lineares:

$$\begin{cases} a^T x \leq \alpha, \\ a^T x \geq \alpha, \end{cases}$$

o qual define a intersecção de dois meios espaços.

**Definição 2.7.** Designa-se por *poliedro convexo* a intersecção de um número finito de meios espaços.

De acordo com esta definição, um conjunto vazio é um poliedro convexo. Por outro lado, podemos ainda concluir que um poliedro fica definido por um número finito de equações e ou inequações lineares. Assim, sendo  $A$  uma matriz com  $m$  linhas e  $n$  colunas ( $A \in \mathbb{R}^{m \times n}$ ) e  $b$  um vector de  $\mathbb{R}^m$  ( $b \in \mathbb{R}^m$ ), tem-se que o sistema de equações

$$Ax = b \Leftrightarrow \begin{bmatrix} A \\ -A \end{bmatrix} x \leq \begin{bmatrix} b \\ -b \end{bmatrix}.$$

define um poliedro.

**Definição 2.8.** Um subconjunto não vazio de  $\mathbb{R}^n$ ,  $K$ , diz-se um *cone* se

$$\forall \lambda > 0 \quad \forall x \in K, \quad \lambda x \in K.$$

Se, adicionalmente,  $K$  é convexo, o cone diz-se *convexo*.

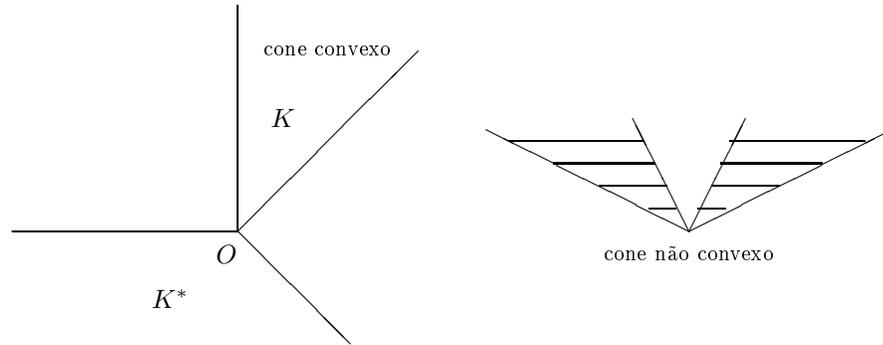


Figura 2.2: Exemplos de cones (convexo e não convexo).

Dado um conjunto,  $K$ , é fácil concluir que o conjunto  $K^* = \{y \in \mathbb{R}^n : y^T x \leq 0 \quad \forall x \in K\}$  é um cone convexo fechado. Com efeito, se  $y \in K^*$  então  $\lambda y \in K^* \quad \forall \lambda > 0$  (logo é um cone) e dados dois pontos  $u, v \in K^*$  (ou seja, tais que  $u^T x \leq 0$  e  $v^T x \leq 0 \quad \forall x \in K$ ) então  $\forall \alpha \in [0, 1]$  vem que  $(\alpha u + (1 - \alpha)v)^T x = \alpha(u^T x) + (1 - \alpha)(v^T x) \leq 0 \quad \forall x \in K$ , donde  $\alpha u + (1 - \alpha)v \in K^*$  (logo é convexo). Para se obter a prova que  $K^*$  é fechado, considerando uma sucessão de  $K^*$ ,  $(u^k)_{k \in \mathbb{N}}$ , tal que  $u^k \rightarrow u$ , basta provar que  $u^T x \leq 0 \quad \forall x \in K$  (ou seja, que  $u \in K^*$ ). Com efeito, suponha-se, por absurdo, que existe  $x' \in K$  tal que  $u^T x' = \epsilon > 0$  e considere-se a aplicação contínua  $x'$  de  $\mathbb{R}^n$  em  $\mathbb{R}$ , definida por

$y \rightarrow x'(y) = y^T x'$ . Pela continuidade de  $x'$  em  $u$ ,  $\exists n \in \mathbb{N}$  tal que  $\forall k \geq n$

$$\begin{aligned} |x'(u^k) - x'(u)| < \frac{\epsilon}{2} &\Leftrightarrow -\frac{\epsilon}{2} < x'(u^k) - \epsilon < \frac{\epsilon}{2} \\ &\Leftrightarrow \frac{\epsilon}{2} < x'(u^k) < \frac{3\epsilon}{2} \\ &\Leftrightarrow \frac{\epsilon}{2} < (u^k)^T x' < \frac{3\epsilon}{2}, \end{aligned}$$

o que é absurdo (tendo em conta que  $\forall k \in \mathbb{N}$ ,  $(u^k)^T x \leq 0 \forall x \in K$ ). Usualmente,  $K^*$  designa-se por *cone polar*<sup>4</sup> *negativo* de  $K$ .

**Teorema 2.4.** *Se  $A = [a_1, \dots, a_n] \in \mathbb{R}^{m \times n}$ , então  $K = \{Ax : x \geq 0\}$  é um cone convexo fechado.*

**Demonstração.** Uma vez que a prova que  $K$  é um cone convexo é imediata, apenas provaremos que  $K$  é fechado.

Se  $A$  é a matriz nula, então  $K = \{0\}$  e o resultado é trivialmente verdadeiro. Supondo  $A \neq 0$  e considerando uma sucessão,  $\{y^k\}_{k \in \mathbb{N}}$ , de  $K$  (logo  $y^k = Ax^k$ , com  $x^k \geq 0$ ) convergente para  $y'$ , vamos provar que  $y' \in K$ .

1. Se as colunas de  $A$ ,  $a_1, \dots, a_n$  são linearmente independentes, então para cada  $y^k$  existe um único  $x^k$  tal que  $Ax^k = y^k$  e, pela continuidade da aplicação  $y \rightarrow x = (A^T A)^{-1} A^T y$  a sucessão  $(x^k)_{k \in \mathbb{N}}$  converge para  $x'$  tal que  $Ax' = y'$ .

Como consequência, dado que  $\mathbb{R}_+^n = \{x \in \mathbb{R}^n : x \geq 0\}$  é fechado e  $\{x^k\}_{k \in \mathbb{N}} \subseteq \mathbb{R}_+^n$ , podemos concluir que  $x' \in \mathbb{R}_+^n$ , i.e,  $x' \geq 0$ .

2. Seja  $v \in K$  e suponha que  $v = Au$  (onde  $u$  é tal que as suas componentes são os escalares não negativos  $\beta_1, \beta_2, \dots, \beta_n$ ) e que, para  $I_u = \{j : \beta_j > 0\}$ , as colunas de  $A$  associadas a este conjunto de índices são linearmente dependentes. Então existem  $|I_u|$  escalares, não todos nulos ( $\alpha_j$  com  $j \in I_u$ ), tais que  $\sum_{j \in I_u} \alpha_j a_j = 0$ , onde, pelo menos um dos escalares  $\alpha_j$  é positivo (caso contrário multiplica-se a equação anterior por  $-1$ ). Determinando-se  $\frac{\beta_p}{\alpha_p} = \min\{\frac{\beta_j}{\alpha_j} : j \in I_u, \alpha_j > 0\}$  vem que

$$\begin{aligned} v &= \sum_{j \in I_u} \beta_j a_j - \frac{\beta_p}{\alpha_p} \sum_{j \in I_u} \alpha_j a_j \\ &= \sum_{j \in I_u} (\beta_j - \frac{\beta_p}{\alpha_p} \alpha_j) a_j \\ &= \sum_{j \in I_u \setminus \{p\}} (\beta_j - \frac{\beta_p}{\alpha_p} \alpha_j) a_j, \text{ com } \beta_j - \frac{\beta_p}{\alpha_p} \alpha_j \geq 0 \quad \forall j \in I_u, \end{aligned}$$

<sup>4</sup>Associado a um subconjunto não vazio de  $\mathbb{R}^n$ ,  $M$ , existe também a noção de conjunto polar de  $M$ , o qual também se denota por  $M^*$  e corresponde ao conjunto  $M^* = \{y \in \mathbb{R}^n : y^T x \leq 1 \forall x \in M\}$ .

ou seja,  $v \in K(I_u \setminus \{p\})$  (onde  $K(J)$  define o cone gerado pelas colunas de  $A$  cujos índices estão  $J$ ). Se estas colunas são linearmente dependentes, então o processo repete-se até se obter um conjunto de índices  $I_w$  tal que  $\{a_j : j \in I_w\}$  é linearmente independente e  $v \in K(I_w)$ , que é um conjunto fechado, de acordo com a prova apresentada no ponto 1. Fazendo variar  $v$  em  $K$ , podemos obter  $K$  como a união dos cones convexos fechados  $K_1, K_2, \dots, K_r$ , gerados por conjuntos maximais de colunas de  $A$  linearmente independentes (cujo número é finito). Logo vem que  $K = \bigcup_{1 \leq i \leq r} K_i$ , pelo que  $K$  é fechado.

□

O teorema que se segue, conhecido por teorema da separação estrita, constitui um resultado com inúmeras aplicações em *Programação Matemática*.

**Teorema 2.5** (da separação estrita). *Seja  $S$  um subconjunto fechado, convexo, não vazio de  $\mathbb{R}^n$  e  $x \in \mathbb{R}^n \setminus S$ . Então existe  $c \in \mathbb{R}^n \setminus \{0\}$  e existe  $\epsilon > 0$  tais que  $c^T x \geq c^T y + \epsilon \forall y \in S$ .*

**Demonstração.** Sendo  $y'$  o ponto de  $S$  mais próximo de  $x$  (note-se que, uma vez que  $S$  é fechado e não vazio, um tal ponto existe (porquê?)<sup>5</sup>, tem-se que  $\|x - y'\| > 0$  (porquê?).

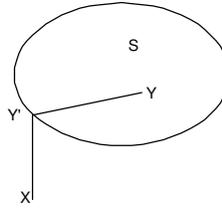


Figura 2.3: Conjunto convexo fechado.

Primeiramente, vamos mostrar que  $(x - y')^T(y - y') \leq 0 \forall y \in S$ . Sendo  $y$  um ponto arbitrário de  $S$  tem-se que  $\forall \lambda \in ]0, 1[$ ,  $\lambda y + (1 - \lambda)y' \in S$  e, uma vez que  $y'$  é o ponto de  $S$  mais próximo de  $x$ , tem-se ainda que

$$\begin{aligned} \|x - y'\|^2 &\leq \|x - \lambda y - (1 - \lambda)y'\|^2 \\ &= \|x - y' + \lambda(y' - y)\|^2 \\ &= \|x - y'\|^2 + 2\lambda(x - y')^T(y' - y) + \lambda^2\|y' - y\|^2, \end{aligned}$$

donde vem que

$$0 \leq 2(x - y')^T(y' - y) + \lambda\|y' - y\|^2 \Rightarrow \lim_{\lambda \rightarrow 0} 2(x - y')^T(y' - y) + \lambda\|y' - y\|^2 = 2(x - y')^T(y' - y) \geq 0.$$

<sup>5</sup>Com efeito, uma vez que  $S$  é não vazio existe  $z \in S$  e considerando a bola fechada de centro em  $x$  e raio  $\|x - z\|$  obtém-se um conjunto compacto cuja intersecção com o conjunto fechado  $S$  é também compacta. Como a determinação do ponto de  $S$  que está à distância mínima de  $x$  corresponde à determinação do mínimo da função contínua  $f(y) = \|x - y\|$  na referida intersecção, vem que o ponto de mínimo,  $y' \in S$ , existe.

Conclui-se assim que

$$(x-y')^T(y'-y) \geq 0 \Leftrightarrow (x-y')^T(x-x+y'-y) = (x-y')^T(y'-x) + (x-y')^T(x-y) \geq 0,$$

pelo que, fazendo  $c = x - y'$  e  $\epsilon = \|x - y'\|^2$ , se obtém

$$-\epsilon + c^T x - c^T y \geq 0 \Leftrightarrow c^T x \geq c^T y + \epsilon.$$

□

Supondo que  $M$  é um subconjunto compacto e  $N$  um subconjunto fechado, pelo que  $M - N = \{z = x - y : x \in M, y \in N\}$  é um conjunto fechado e supondo que  $M$  e  $N$  são convexos, donde  $M - N$  também é convexo, com base no Teorema 2.5 e tendo em conta que  $M$  e  $N$  são disjuntos sse  $0 \notin M - N$ , podemos obter a separação estrita entre o subconjunto compacto  $M$  e o subconjunto fechado  $N$ , disjunto de  $M$ , obtendo a separação estrita entre  $0 \notin M - N$  e  $M - N$ . Com efeito,

$$\begin{aligned} c^T 0 > c^T z \quad \forall z \in M - N &\Leftrightarrow 0 > c^T(x - y) \quad \forall x \in M \quad \forall y \in N \\ &\Leftrightarrow c^T x > c^T y \quad \forall x \in M \quad \forall y \in N. \end{aligned}$$

Deve sublinhar-se que para a obtenção da separação estrita entre dois subconjuntos disjuntos convexos, não basta que eles sejam fechados, é necessário que pelo menos um deles seja compacto. Por exemplo, embora os subconjuntos  $A = \{(x, y) : x > 0 \text{ e } y \geq \frac{1}{x}\} \subset \mathbb{R}^2$  e  $B = \{(x, y) : x \geq 0 \text{ e } y = 0\} \subset \mathbb{R}^2$  sejam disjuntos, convexos e fechados, não existe qualquer separação estrita entre eles.

**Teorema 2.6.** *Seja  $S \subseteq \mathbb{R}^n$  um subconjunto convexo com interior não vazio e seja  $x$  um ponto da fronteira de  $S$  ( $x \in \text{fr}(S)$ )<sup>6</sup>. Então existe um vector  $v \in \mathbb{R}^n \setminus \{0\}$  tal que  $v^T x \geq v^T y \quad \forall y \in S$ .*

**Demonstração.** Seja  $x'$  um ponto interior de  $S$  e, para  $\tau > 1$ , seja  $x_\tau = x' + \tau(x - x')$ . É claro que  $x_\tau \notin \text{Ad}(S) \quad \forall \tau > 1$ , onde  $\text{Ad}(S)$  denota a aderência de  $S$ <sup>7</sup>. Uma vez que  $x_\tau \notin \text{Ad}(S) \quad \forall \tau > 1$ , pelo Teorema 2.5 (da separação estrita) para cada  $\tau > 1$ , existe  $u_\tau \neq 0$  tal que  $u_\tau^T x_\tau > u_\tau^T y \quad \forall y \in S$ . Supondo, sem perda de generalidade, que  $\|u_\tau\| = 1 \quad \forall \tau > 1$ <sup>8</sup> e considerando uma sucessão  $(\tau_k)_{k \in \mathbb{N}}$  convergente para 1 ( $\lim_{\tau \rightarrow \infty} \tau_k = 1$ ), tal que  $\tau_k > 1 \quad \forall k \in \mathbb{N}$ , pelo Teorema de Bolzano-Weierstrass pode concluir-se que a sucessão limitada  $(v_k)$ ,

<sup>6</sup>Dado um conjunto  $C \subseteq \mathbb{R}^n$ , designa-se por fronteira de  $C$  e denota-se por  $\text{fr}(C)$  o conjunto de todos os pontos  $x$  relativamente aos quais, qualquer bolas aberta centrada em  $x$  intersecta  $C$  e o seu complementar  $\bar{C}$ , ou seja,  $\forall \epsilon > 0$

$$\{y \in \mathbb{R}^n : \|x - y\| < \epsilon\} \cap C \neq \emptyset \wedge \{y \in \mathbb{R}^n : \|x - y\| < \epsilon\} \cap \bar{C} \neq \emptyset.$$

<sup>7</sup>Dado um conjunto  $C \subseteq \mathbb{R}^n$ , designa-se por aderência de  $C$  e denota-se por  $\text{Ad}(C)$ , o conjunto de todos os pontos  $x$  relativamente aos quais, qualquer bola aberta centrada em  $x$  intersecta  $C$ , ou seja  $\forall \epsilon > 0 \{y \in \mathbb{R}^n : \|x - y\| < \epsilon\} \cap C \neq \emptyset$ .

<sup>8</sup>Para tal basta multiplicar  $u_\tau$  pelo inverso da sua norma.

com  $v_k = u_{\tau_k}$ , admite uma subsucessão,  $(v_{k_p})$ , convergente para  $v \in \mathbb{R}^n$ . Adicionalmente, dado que  $1 = \lim_{p \rightarrow \infty} \|v_{k_p}\| = \|v\|$  conclui-se ainda que  $v \neq 0$ . Destas considerações, decorre finalmente que  $\forall y \in S$

$$v^T y = \lim_{p \rightarrow \infty} (v_{k_p}^T y) \leq \lim_{p \rightarrow \infty} (v_{k_p}^T x_{k_p}) = v^T x,$$

onde  $x_{\tau_k}$  se denota por  $x_k$ . □

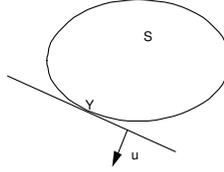


Figura 2.4: Hiperplano de suporte de um conjunto convexo.

**Lema 2.7** (de Farkas, 1ª versão). *Dada uma matriz  $A = [a_1, a_2, \dots, a_n] \in \mathbb{R}^{m \times n}$  e  $b \in \mathbb{R}^m$ , um e não mais do que um dos sistemas a seguir indicados tem solução.*

$$Ax = b, \quad e \quad x \geq 0; \tag{2.3}$$

$$y^T A \leq 0 \quad e \quad y^T b > 0. \tag{2.4}$$

**Demonstração.** Suponha que (2.3) tem solução, ou seja,  $\exists x' \in \mathbb{R}^n$  tal que  $Ax' = b$  e  $x' \geq 0$ , então  $\forall y \in \mathbb{R}^m$   $y^T Ax' = y^T b$ . Logo se  $\exists y' \in \mathbb{R}^m$  tal que  $y'^T A \leq 0$ , vem que  $y'^T Ax' = y'^T b \leq 0$ , o que torna impossível a verificação da última inequação de (2.4).

Suponha agora que (2.3) não tem solução. Então  $b \notin K = \{Ax : x \geq 0\}$  e, uma vez que (pelo Teorema 2.4)  $K$  é um convexo fechado, por aplicação do Teorema 2.5 conclui-se que  $\exists y \in \mathbb{R}^m \setminus \{0\}$  e  $\exists \epsilon > 0$  tais que  $y^T b \geq y^T z + \epsilon \quad \forall z \in K$  (ou seja,  $\forall z : z = Ax, \text{ com } x \geq 0$ ). Logo, tendo em conta que  $0 \in K$  e  $\epsilon > 0$ , podemos concluir que  $y^T b > 0$ . Por outro lado, dado que  $K$  é um cone,  $\forall z \in K \quad \forall \rho > 0, \rho z \in K$ . Consequentemente, para  $z = Ae_j$ , com  $j \in \{1, \dots, n\}$  (onde  $e_j$  denota o  $j$ -ésimo vector da base canónica de  $\mathbb{R}^n$ ), vem que  $\forall \rho > 0, \rho y^T z = \rho y^T Ae_j = \rho y^T a_j \leq y^T b - \epsilon$ , o que só é possível se  $y^T a_j \leq 0$  (caso contrário, bastaria tomar  $\rho = \frac{y^T b - \epsilon + 1}{y^T a_j}$  para que a inequação em causa fosse violada). Assim, prova-se que  $y^T a_j \leq 0 \quad \forall j \in \{1, \dots, n\}$  e, consequentemente, que  $y$  é uma solução para o sistema (2.4), ou seja,  $y^T A \leq 0$  e  $y^T b > 0$ . □

A interpretação geométrica do lema de Farkas permite concluir que, se  $b$  pertence ao cone  $K$ , gerado pelas colunas de  $A$ , então não existe  $y$  pertencente

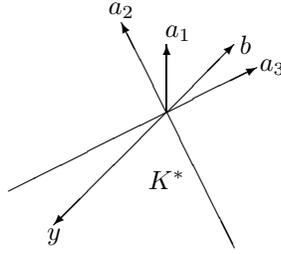


Figura 2.5: Interpretação geométrica do Lema de Farkas.

ao cone polar negativo de  $K$ ,  $K^* = \{y \in \mathbb{R}^m : y^t z \leq 0 \forall z \in K\}$ , tal que  $y^t b > 0$ . A figura a seguir ilustra esta afirmação.

Note-se que

$$K^* = \{y \in \mathbb{R}^m : y^T z \leq 0 \forall z \in K\} = \{y \in \mathbb{R}^m : y^T a_j \leq 0, j = 1, \dots, n\}.$$

Com efeito, se  $y^T a_j \leq 0 \forall j \in \{1, \dots, n\}$ , então

$$y^T Ax \leq 0 \forall x \geq 0 \Leftrightarrow y^T z \leq 0 \forall z \in K.$$

Reciprocamente, se  $y^T z \leq 0 \forall z \in K$ , então, em particular,  $y^T A e_j \leq 0 \forall j \in \{1, \dots, n\} \Leftrightarrow y^T a_j \leq 0 \forall j \in \{1, \dots, n\}$ .

Sendo  $A$  uma matriz com  $m$  linhas e  $n$  colunas ( $A \in \mathbb{R}^{m \times n}$ ) e  $b$  um vector de  $\mathbb{R}^m$  ( $b \in \mathbb{R}^m$ ), o sistema de inequações  $Ax \leq b$  que determina um poliedro convexo, pode transformar-se no sistema equivalente de equações,  $Ax + y = b$ , desde que se acrescentem as condições de não negatividade ao  $m$ -uplo de variáveis  $y$ , ou seja,

$$\{x \in \mathbb{R}^n : Ax \leq b\} = \{x \in \mathbb{R}^n : Ax + y = b, y \geq 0\}.$$

Com efeito, sendo  $S = \{x \in \mathbb{R}^n : Ax \leq b\}$  e  $T = \{x \in \mathbb{R}^n : Ax + y = b, y \geq 0\}$ , se  $x \in S$ , fazendo  $y = b - Ax \geq 0$ , conclui-se que  $\exists y \geq 0$  tal que  $Ax + y = b$ , pelo que  $x \in T$ . Logo,  $T \supseteq S$  (i). Reciprocamente, se  $x \in T$ , então  $\exists y \geq 0$  tal que  $Ax + y = b$ , pelo que  $x$  é tal que  $b - Ax \geq 0 \Leftrightarrow Ax \leq b$ , ou seja,  $x \in S$ . Consequentemente,  $S \supseteq T$  (ii). De (i) e (ii) conclui-se que  $S = T$ . Assim, supondo que  $A \in \mathbb{R}^{m \times n}$ , a partir de um conjunto

$$T = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^{m \times n} : [A, I_m] \begin{bmatrix} x \\ y \end{bmatrix} = b \right\}$$

por aplicação da projecção  $\pi : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^n$ , tal que  $\pi\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = x$ , vem que  $\pi(T) = S$ . A restrição a  $T$  desta projecção, é uma aplicação bijectiva entre  $T$  e

$S$ , a qual admite, como inversa, a aplicação  $\pi^{-1}(x) = (x, b - Ax)$ .

As variáveis  $y_1, y_2, \dots, y_m$ , introduzidas na passagem de  $S$  para  $T$ , designam-se por variáveis de desvio ou folga. Recorrendo a estas variáveis de desvio, estamos em condições de apresentar uma versão alternativa à versão do Lema de Farkas anteriormente referida.

**Lema 2.8** (de Farkas, 2ª versão). *Dada uma matriz  $A \in \mathbb{R}^{m \times n}$  e  $b \in \mathbb{R}^m$ , um e não mais do que um dos sistemas a seguir indicados tem solução.*

$$Ax \geq b \quad e \quad x \geq 0. \quad (2.5)$$

$$y^T A \leq 0, y^T b > 0 \quad e \quad y \geq 0. \quad (2.6)$$

**Demonstração.** Introduzindo variáveis de desvio no sistema (2.5), i.e., fazendo  $Ax - s = b$ , com  $s \geq 0$ , os sistemas (2.5) e (2.6) podem escrever-se (de um modo equivalente) da seguinte forma:

$$[A, -I] \begin{bmatrix} x \\ s \end{bmatrix} = b \quad e \quad \begin{bmatrix} x \\ s \end{bmatrix} \geq 0. \quad (2.7)$$

$$y^T [A, -I] \leq 0 \quad e \quad y^T b > 0. \quad (2.8)$$

Nestas condições, por aplicação directa do Lema 2.7, conclui-se o que se pretendia.  $\square$

## 2.2 Pontos Extremos e Direcções Extremas

Os conjuntos convexos têm certos pontos (pontos extremos) e direcções (direcções extremas) com especial interesse, no contexto da Optimização, muito particularmente, em Optimização Linear. Com efeito, como veremos, se o conjunto de soluções óptimas de uma funcional linear num poliedro convexo é não vazio, então esse conjunto contém, necessariamente, um ponto extremo e se essa funcional linear não tem mínimo (máximo) finito, então existe uma direcção extrema ao longo da qual a funcional linear é decrescente (crescente).

**Definição 2.9.** *Seja  $S \subseteq \mathbb{R}^n$  um conjunto convexo não vazio, diz-se que  $x \in S$  é um ponto extremo de  $S$  se não existem dois pontos  $y, z \in S$  distintos ( $y \neq z$ ) e um escalar  $\lambda \in ]0, 1[$  tais que  $x = \lambda y + (1 - \lambda)z$ .*

O teorema a seguir dá-nos uma condição necessária e suficiente para que um ponto de um conjunto convexo seja um ponto extremo.

**Teorema 2.9.** *Dado um conjunto convexo  $S \subseteq \mathbb{R}^n$ , o ponto  $x \in S$  é um ponto extremo sse  $\forall y \in S \setminus \{x\} \{z \in \mathbb{R}^n : z = x + \lambda(x - y), \lambda > 0\} \cap S = \emptyset$ .*

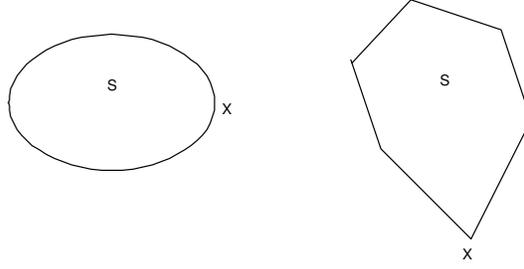


Figura 2.6: Exemplos de pontos extremos.

**Demonstração.** Suponha-se que  $x$  é um ponto extremo de  $S$ , seja  $y \in S \setminus \{x\}$  e  $z = x + \lambda(x - y)$  com  $\lambda > 0$ . Logo, vem que  $x = \frac{1}{1+\lambda}z + \frac{\lambda}{1+\lambda}y$ , donde se conclui que  $z \notin S$  (uma vez que  $x$  é um ponto extremo de  $S$ ). Como consequência, podemos afirmar que

$$\{z \in \mathbb{R}^n : z = x + \lambda(x - y), \lambda > 0\} \cap S = \emptyset.$$

Reciprocamente, suponha-se que  $x$  não é um ponto extremo de  $S$ , ou seja, que existem dois pontos de  $S$ ,  $y$  e  $z$ , distintos ( $y \neq z$ ) e um escalar  $\alpha \in ]0, 1[$  tais que  $x = (1 - \alpha)y + \alpha z$ . Assim, vem que

$$z = \frac{1}{\alpha}x - \frac{1 - \alpha}{\alpha}y = x + \frac{1 - \alpha}{\alpha}(x - y) = x + \lambda(x - y),$$

com  $\lambda = \frac{1 - \alpha}{\alpha} > 0$ , pelo que  $\{z \in \mathbb{R}^n : z = x + \lambda(x - y), \lambda > 0\} \cap S \neq \emptyset$ . □

**Definição 2.10.** Sendo  $S$  um conjunto convexo de  $\mathbb{R}^n$ ,  $d \in \mathbb{R}^n \setminus \{0\}$  diz-se uma direcção de  $S$  se

$$\forall x \in S, S \supseteq \{y : y = x + \lambda d, \lambda > 0\}.$$

Dois direcções  $d^1$  e  $d^2$  dizem-se distintas se não existe um escalar  $\alpha > 0$ , tal que  $d^1 = \alpha d^2$ .

Como é óbvio, só os conjuntos ilimitados é que admitem direcções.

**Definição 2.11.** Uma direcção  $d$  de um conjunto convexo  $S \subseteq \mathbb{R}^n$  diz-se uma direcção extrema, se não admitir uma representação do tipo

$$d = d^1 + d^2,$$

onde  $d^1$  e  $d^2$  são duas direcções distintas de  $S$ .

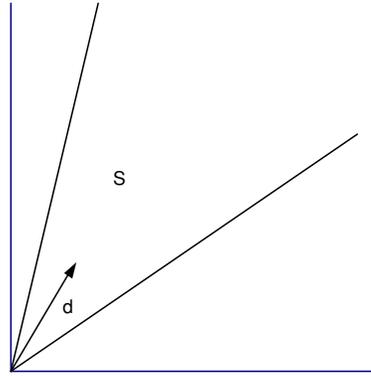


Figura 2.7: Exemplo de um conjunto que contém direcções.

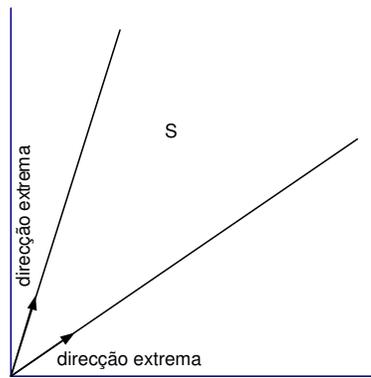


Figura 2.8: Exemplos de direcções extremas.

### 2.3 Relações entre pontos extremos e soluções básicas

No que se segue,  $S$  identificará o conjunto de pontos de  $\mathbb{R}^n$ ,

$$S = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\},$$

com  $A \in \mathbb{R}^{m \times n}$  tal que  $r(A) = m$  (onde  $r(A)$  denota a característica de  $A$ ) e  $b \in \mathbb{R}^m$ .

**Definição 2.12.** *Uma solução  $\bar{x}$  do sistema de equações  $Ax = b$  (anteriormente referido) diz-se uma solução básica se, após eventual reordenação das suas componentes e, conseqüentemente, das colunas de  $A$ , a matriz  $A$  admitir uma decomposição do tipo  $A = [B, N]$ , em que  $B$  é uma matriz invertível e se verifica que*

$$A\bar{x} = B\bar{x}_B + N\bar{x}_N = B\bar{x}_B = b$$

(com  $\bar{x}_N = 0$ ). Adicionalmente, verificando-se que as componentes de  $\bar{x}_B$  são não negativas, a solução diz-se uma solução básica admissível para  $S$ . As componentes associadas a  $\bar{x}_B$  ( $\bar{x}_N$ ) designam-se por componentes básicas (não básicas) da solução básica  $\bar{x}$ .

**Teorema 2.10.**  $x$  é um ponto extremo de  $S$  sse é uma solução básica admissível para  $S$ , i.e., sse (após uma eventual reordenação das componentes de  $x$  e, conseqüentemente, das colunas de  $A$ ) a matriz  $A$  admitir uma decomposição do tipo  $A = [B, N]$ , em que  $B$  é uma matriz invertível e

$$x = \begin{bmatrix} x_B \\ x_N \end{bmatrix} = \begin{bmatrix} B^{-1}b \\ 0 \end{bmatrix} \geq 0.$$

**Demonstração.** Suponha que  $A$  se pode decompor na forma  $[B, N]$ , onde  $B$  é uma matriz invertível tal que

$$x = \begin{bmatrix} B^{-1}b \\ 0 \end{bmatrix},$$

e se verifica que  $B^{-1}b \geq 0$ . É claro que, nestas condições,  $x$  é uma solução básica admissível para  $S$ .

Suponha que existem dois pontos de  $S$ ,  $x'$  e  $x''$ , e um escalar  $\lambda \in ]0, 1[$  para os quais se verifica que  $x = \lambda x' + (1 - \lambda)x''$ . Seja

$$x' = \begin{bmatrix} x'_B \\ x'_N \end{bmatrix} \text{ e } x'' = \begin{bmatrix} x''_B \\ x''_N \end{bmatrix},$$

então vem que

$$\begin{bmatrix} B^{-1}b \\ 0 \end{bmatrix} = \lambda \begin{bmatrix} x'_B \\ x'_N \end{bmatrix} + (1 - \lambda) \begin{bmatrix} x''_B \\ x''_N \end{bmatrix},$$

donde se conclui que  $x'_N = x''_N = 0$ . Logo  $x'_B = x''_B = B^{-1}b$ <sup>9</sup>, pelo que  $x' = x'' = x$ , o que mostra que  $x$  é um ponto extremo de  $S$ .

Reciprocamente, suponha que  $x$  é um ponto extremo de  $S$  e, sem perda de generalidade, que

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

<sup>9</sup>Observe-se que, uma vez que  $x'_N = x''_N = 0$ , então

$$\begin{cases} Bx'_B = b \\ Bx''_B = b \end{cases} \Leftrightarrow \begin{cases} x'_B = B^{-1}b \\ x''_B = B^{-1}b \end{cases}$$

é tal que  $x_i > 0$ , para  $i = 1, \dots, k$ . Suponha ainda que, sendo  $A = [a_1, a_2, \dots, a_k, \dots, a_n]$ , onde  $a_1, a_2, \dots, a_k, \dots, a_n$  denotam as colunas de  $A$ , as colunas  $a_1, a_2, \dots, a_k$  são linearmente dependentes, ou seja, que existem  $k$  escalares,  $\alpha_1, \alpha_2, \dots, \alpha_k$ , não todos nulos, tais que  $\sum_{i=1}^k \alpha_i a_i = 0$ .

Fazendo

$$u = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_k \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$x' = x + \beta u$  e  $x'' = x - \beta u$ , com  $\beta > 0$  e tal que  $x' \geq 0$  e  $x'' \geq 0$ , tem-se que  $x', x'' \in S$  (note-se que  $Ax' = Ax + \beta Au = b + \beta \sum_{i=1}^k \alpha_i a_i = b$  e  $Ax'' = Ax - \beta Au = b - \beta \sum_{i=1}^k \alpha_i a_i = b$ ). Uma vez que  $x' \neq x''$ <sup>10</sup> e  $x = \frac{1}{2}x' + \frac{1}{2}x''$  conclui-se que  $x$  não é um ponto extremo de  $S$ , o que contraria a hipótese. Esta contradição resultou de se haver suposto que  $a_1, a_2, \dots, a_k$  seriam linearmente dependentes. Assim, provou-se que se  $x$  é um ponto extremo de  $S$ , então  $a_1, a_2, \dots, a_k$  (as colunas de  $A$  associadas a componentes estritamente positivas de  $x$ ) são linearmente independentes, pelo que  $k \leq m$ . No caso de se ter  $k < m$  bastará completar as  $k$  colunas de  $A$  com  $m - k$  colunas, de entre as colunas  $a_{k+1}, \dots, a_n$ , de modo a obter-se uma matriz invertível  $B = [a_1 \dots a_k a_{j_1} \dots a_{j_{m-k}}]$ , para a qual vem que

$$b = x_1 a_1 + \dots + x_k a_k + 0 a_{j_1} + \dots + 0 a_{j_{m-k}} \Leftrightarrow B \begin{bmatrix} x_1 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{bmatrix} = b \Leftrightarrow \begin{bmatrix} x_1 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{bmatrix} = B^{-1}b.$$

□

Deve observar-se, no entanto, que não existe uma correspondência biunívoca entre pontos extremos (vértices) e soluções básicas admissíveis. Com efeito, embora a uma dada solução básica admissível corresponda um único vértice, o recíproco nem sempre é verdadeiro, conforme a seguir se exemplifica.

No poliedro convexo

$$S = \{(x, y) \in \mathbb{R}^2 : 3x + 5y \leq 40, 5x + 3y \leq 40, x + y \leq 10, x, y \geq 0\},$$

a seguir representado, facilmente se reconhece que ao ponto  $(5, 5)$  correspondem mais do que uma solução básica admissível.

<sup>10</sup>Caso contrário, ter-se-ia  $x + \beta u = x - \beta u \Leftrightarrow 2\beta u = 0$ , o que contraria o facto de se saber que  $\beta > 0$  e  $u \neq 0$ .

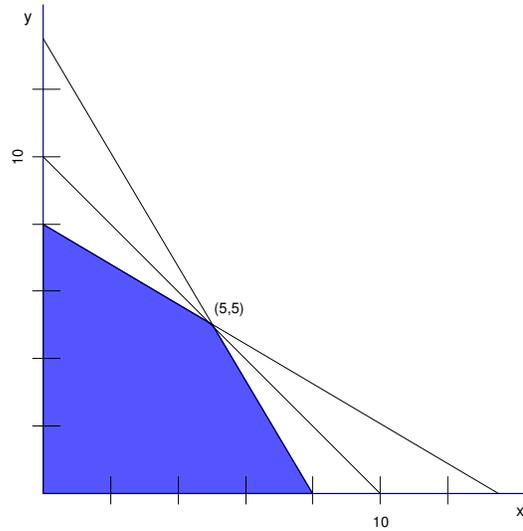


Figura 2.9: Ponto extremo correspondendo a soluções básicas degeneradas.

Estas situações ocorrem em soluções básicas admissíveis com uma ou mais componentes básicas nulas.

**Definição 2.13.** *Uma solução básica,  $\bar{x}$ , do sistema de equações  $Ax = b$ , que apresenta pelo menos uma das suas componentes básicas com valor nulo designa-se por solução básica degenerada (e, no caso de ser admissível, diz-se básica admissível degenerada).*

No exemplo da Figura 2.9, transformando o sistema de inequações no sistema de equações de variáveis não negativas:

$$\begin{bmatrix} 3 & 5 & 1 & 0 & 0 \\ 5 & 3 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} 40 \\ 40 \\ 10 \end{bmatrix},$$

$$x, y, z_1, z_2, z_3 \geq 0,$$

verifica-se que ao ponto extremo (5, 5) estão associadas as soluções básicas degeneradas  $(x, y, z_1) = (5, 5, 0)$ ,  $(x, y, z_2) = (5, 5, 0)$  e  $(x, y, z_3) = (5, 5, 0)$  às quais correspondem as bases

$$B' = \begin{bmatrix} 3 & 5 & 1 \\ 5 & 3 & 0 \\ 1 & 1 & 0 \end{bmatrix}, B'' = \begin{bmatrix} 3 & 5 & 0 \\ 5 & 3 & 1 \\ 1 & 1 & 0 \end{bmatrix}, B''' = \begin{bmatrix} 3 & 5 & 0 \\ 5 & 3 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

Uma vez que os candidatos a pontos extremos de  $S$  são as soluções básicas do sistema de equações  $Ax = b$ , podemos concluir que qualquer ponto extremo de  $S$  tem pelo menos  $n - m$  componentes nulas e que o número de pontos extremos de  $S$  é não superior a  $\binom{n}{m} = \frac{n!}{m!(n-m)!}$ .

**Teorema 2.11.** *Se  $S \neq \emptyset$  então  $S$  tem pelo menos um ponto extremo.*

**Demonstração.** Seja  $x \in S$  e, sem perda de generalidade, suponha que  $x = (x_1, \dots, x_k, x_{k+1}, \dots, x_n)^T$ , onde  $x_i > 0$ , para  $i = 1, \dots, k$  e  $x_j = 0$ , para  $j = k + 1, \dots, n$ . Seja  $A = [a_1 \dots a_k a_{k+1} \dots a_n]$ , onde  $a_1, \dots, a_k, a_{k+1}, \dots, a_n$  denotam as colunas de  $A$ .

Se  $a_1, \dots, a_k$  são vectores linearmente independentes, então  $k \leq m$  e  $x$  é uma solução básica admissível, pelo que (de acordo com o Teorema 2.10) é um ponto extremo.

Suponha que existem  $k$  escalares,  $\lambda_1, \dots, \lambda_k$ , com pelo menos um positivo, tais que  $\sum_{j=1}^k \lambda_j a_j = 0$ , ou seja,  $a_1, \dots, a_k$  são vectores linearmente dependentes (note-se que se existem  $k$  escalares não positivos que satisfazem o sistema de equações  $\sum_{j=1}^k \lambda_j a_j = 0$ , multiplicando-os por  $-1$  obtêm-se  $k$  escalares não negativos que satisfazem as mesmas equações). Fazendo-se  $\alpha = \min_{1 \leq j \leq k} \{\frac{x_j}{\lambda_j} : \lambda_j > 0\} = \frac{x_i}{\lambda_i}$  e determinando-se  $x'$  de tal forma que as suas componentes venham determinadas por

$$x'_j = \begin{cases} x_j - \alpha \lambda_j, & \text{para } 1 \leq j \leq k, \\ 0, & \text{para } j > k, \end{cases}$$

vem que  $x'_j \geq 0$  para,  $j = 1, \dots, k$ ,  $x'_j = 0$  para  $j = k + 1, \dots, n$  e  $x'_i = 0$ . Nestas condições,

$$Ax' = \sum_{j=1}^n x'_j a_j = \sum_{j=1}^k (x_j - \alpha \lambda_j) a_j = \sum_{j=1}^k x_j a_j - \alpha \sum_{j=1}^k \lambda_j a_j = b - 0 = b,$$

pelo que se obtém um novo ponto  $x'$  com, no máximo  $k - 1$  componentes estritamente positivas. Uma vez que este processo se pode repetir até que às componentes estritamente positivas correspondam colunas de  $A$  linearmente independentes, conclui-se o pretendido. □

## 2.4 Direcções extremas e representação de poliedros convexos

**Teorema 2.12.** *Pode afirmar-se que  $d$  é uma direcção extrema de  $S$  sse (após reordenação das componentes de  $d$  e, consequentemente, das colunas de  $A$ )*

1. a matriz  $A$  admite uma decomposição do tipo  $A = [B, N]$ , onde  $B$  é uma matriz invertível,

2. existe pelo menos uma coluna  $a_j$  tal que  $B^{-1}a_j \leq 0$ ,

3.  $d$  é um múltiplo escalar positivo de

$$\begin{bmatrix} -B^{-1}a_j \\ e_{j-m} \end{bmatrix},$$

com  $e_{j-m}$  denotando o  $(j-m)$ -ésimo vector da base canónica de  $\mathbb{R}^{n-m}$ .

**Demonstração.** ( $\Leftarrow$ ) Se  $B^{-1}a_j \leq 0$ , então  $d \geq 0$  e, além disso,  $Ad = 0$ , donde  $d$  é uma direcção de  $S$ . Vamos mostrar que se trata de uma direcção extrema.

Seja  $d = d' + d''$ , com  $d'$  e  $d''$  duas direcções de  $S$ , tem-se que  $n-m-1$  componentes, tanto de  $d'$  como de  $d''$ , são iguais a zero (uma vez que as correspondentes componentes de  $d$  são nulas). Nestas condições, tanto  $d'$  como  $d''$  são da forma

$$d' = \begin{bmatrix} d'_B \\ \alpha e_{j-m} \end{bmatrix} \text{ e } d'' = \begin{bmatrix} d''_B \\ \beta e_{j-m} \end{bmatrix},$$

com  $\alpha + \beta = 1$ . Dado que  $Ad' = Ad'' = 0$ , i.e.,  $Ad' = Bd'_B + \alpha a_j = 0$  e  $Ad'' = Bd''_B + \beta a_j = 0$ ,  $d'_B = \alpha(-B^{-1}a_j)$  e  $d''_B = \beta(-B^{-1}a_j)$ , ou seja,  $d'$  e  $d''$  não são duas direcções distintas. Logo  $d$  é uma direcção extrema.

( $\Rightarrow$ ) Reciprocamente, suponha que  $d$  é uma direcção extrema de  $S$  e, sem perda de generalidade, que

$$d = \begin{bmatrix} d_1 \\ \vdots \\ d_k \\ 0 \\ \vdots \\ 0 \\ d_j \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

com  $d_i > 0$ , para  $i = 1, \dots, k$  e  $i = j$ .

Suponha ainda que as  $k$  primeiras colunas de  $A$ ,  $a_1, \dots, a_k$ , são linearmente dependentes. Logo, existem  $k$  escalares,  $\lambda_1, \dots, \lambda_k$ , não todos nulos, tais

que  $\sum_{i=1}^k \lambda_i a_i = 0$ . Seja

$$u = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_k \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

e escolha-se  $\alpha > 0$ , suficientemente pequeno para que os vectores  $d' = d + \alpha u$  e  $d'' = d - \alpha u$  tenham componentes não negativas. Note-se que

$$Ad' = Ad + \alpha Au = 0 + \sum_{i=1}^k \lambda_i a_i = 0 = 0 - \sum_{i=1}^k \lambda_i a_i = Ad - \alpha Au = Ad''.$$

Assim,  $d'$  e  $d''$  são duas direcções de  $S$  distintas de  $d$  e distintas entre si, uma vez que  $\alpha > 0$ ,  $u \neq 0$  e a  $j$ -ésima componente, tanto de  $d'$  como de  $d''$ , são iguais à  $j$ -ésima componente de  $d$ <sup>11</sup>. Adicionalmente, pode concluir-se que  $d = \frac{1}{2}d' + \frac{1}{2}d''$ , o que contradiz a suposição de  $d$  ser uma direcção extrema. Consequentemente,  $a_1, \dots, a_k$  são colunas de  $A$  linearmente independentes e, uma vez que a característica de  $A$  é  $m$ , tem-se que  $k \leq m$ . Nestas condições, acrescentando as  $m - k$  colunas de  $A$  que, conjuntamente com  $a_1, \dots, a_k$ , formam uma base para  $\mathbb{R}^m$  e, consequentemente, constituem uma matriz invertível  $B$ , fazendo  $A = [B, N]$ , vem

$$Ad = Bd_B + d_j a_j = 0 \Leftrightarrow d_B = d_j B^{-1} a_j,$$

donde

$$d = d_j \begin{bmatrix} -B^{-1} a_j \\ e_{j-m} \end{bmatrix}.$$

□

Como corolário, imediato, deste teorema pode afirmar-se que o número de direcções extremas de  $S$  é não superior a  $(n - m) \binom{n}{m}$ , onde  $n - m$  corresponde ao número de possibilidades de escolha de  $a_j$  e  $\binom{n}{m}$  corresponde ao número de partições possíveis de  $A$  na forma  $[B, N]$ .

Segue-se um teorema que desempenha um papel crucial em Optimização Linear.

**Teorema 2.13** (da representação de poliedros convexos). *Dado um ponto  $x \in \mathbb{R}^n$ ,  $x \in S$  sse  $x$  admite uma representação do tipo*

$$x = \sum_{j=1}^p \alpha_j x^j + \sum_{i=1}^q \beta_i d^i,$$

<sup>11</sup>Com efeito,  $d' = \tau d''$  com  $\tau > 0$ , é equivalente a  $d + \alpha u = \tau(d - \alpha u) \Leftrightarrow (1 + \tau)\alpha u = (\tau - 1)d \Rightarrow (\tau - 1)d_j = 0$  (uma vez que a  $j$ -ésima componente de  $u$  é nula). Por outro lado, dado que (por hipótese)  $d_j > 0$ , conclui-se que  $\tau = 1$ . Logo,  $2\alpha u = 0$ , o que é absurdo, tendo em conta que  $\alpha > 0$  e  $u \neq 0$ .

com  $\sum_{j=1}^p \alpha_j = 1$ ,  $\alpha_j \geq 0$ , para  $j = 1, \dots, p$ , e  $\beta_i \geq 0$ , para  $i = 1, \dots, q$ , onde os  $x^j$ s são pontos extremos de  $S$  e os  $d^i$ s são direções extremas de  $S$  (se tais direções existem).

**Demonstração.** Uma vez que qualquer ponto  $x$  que admita uma representação do tipo da indicada no teorema é um ponto de  $S$  (porquê?), resta-nos provar que qualquer ponto de  $S$  admite uma representação deste tipo.

Suponha-se que  $\exists x \in S$  tal que  $x$  não se pode representar pela adição de uma combinação linear convexa de pontos extremos de  $S$  com uma combinação linear não negativa das suas direções extremas, i.e., tal que  $x \notin X$ , onde

$$X = \left\{ \sum_{j=1}^p \alpha_j x^j + \sum_{i=1}^q \beta_i d^i : \sum_{j=1}^p \alpha_j = 1, \alpha_j \geq 0, j = 1, \dots, p, \beta_i \geq 0, i = 1, \dots, q \right\}.$$

Dado que  $S \neq \emptyset$ , podemos concluir que  $S$  tem pelo menos um ponto extremo e, consequentemente, que  $X \neq \emptyset$ . Adicionalmente, fazendo

$$C = \begin{bmatrix} x^1 & \dots & x^p & d^1 & \dots & d^q \\ 1 & \dots & 1 & 0 & \dots & 0 \end{bmatrix} \text{ e } u = (\alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q)^T$$

dizer que  $x \notin X$ , equivale a afirmar que não existe  $u \geq 0$  tal que  $Cu = \begin{bmatrix} x \\ 1 \end{bmatrix}$ .

Logo, aplicando o Lema de Farkas

$$\exists \begin{bmatrix} c \\ \delta \end{bmatrix} \in \mathbb{R}^{n+1} \setminus \{0\} \text{ tal que } [c^T, \delta]C \leq 0 \text{ e } [c^T, \delta] \begin{bmatrix} x \\ 1 \end{bmatrix} > 0,$$

o que equivale a afirmar que

1.  $c^T x^j + \delta \leq 0 \forall j \in \{1, \dots, p\}$ ,
2.  $c^T d^i \leq 0 \forall i \in \{1, \dots, q\}$
3. e  $c^T x + \delta > 0$ .

Sendo  $x^k$  um ponto extremo de  $X$ , tal que  $c^T x^k = \max\{c^T x^j, j = 1, \dots, p\}$ , e tendo em conta 1. e 3., conclui-se que

4.  $c^T x > c^T x^k$ .

Uma vez que  $x^k$  é um ponto extremo de  $S$ , pelo Teorema 2.10,  $x^k$  é uma solução básica para  $S$ , ou seja, pode representar-se na forma

$$x^k = \begin{bmatrix} B^{-1}b \\ 0 \end{bmatrix},$$

onde  $B$  é uma submatriz de  $A$  invertível que, sem perda de generalidade, se supõe tal que  $B^{-1}b > 0$ . Dado que  $x \in S$ , tem-se que  $Ax = b$  e  $x \geq 0$ , pelo que,

decompondo  $x$  em  $\begin{bmatrix} x_B \\ x_N \end{bmatrix}$ , se obtém

$$b = Ax = Bx_B + Nx_N \Rightarrow x_B = B^{-1}b - B^{-1}Nx_N.$$

Por sua vez, decompondo  $c^T$  em  $[c_B, c_N]$ , vem

$$c^T x = c_B x_B + c_N x_N = c_B B^{-1} b + (c_N - c_B B^{-1} N) x_N = c^T x^k + (c_N - c_B B^{-1} N) x_N$$

(uma vez que  $c^T x^k = c_B B^{-1} b$ ). Logo, de 4. decorre que

$$c^T x^k + (c_N - c_B B^{-1} N) x_N > c^T x^k \Leftrightarrow (c_N - c_B B^{-1} N) x_N > 0.$$

Desta última inequação, e uma vez que cada componente de  $x_N$  é não negativa, decorre que existe pelo menos uma componente de  $x_N$ ,  $x_r$ , tal que  $x_r > 0$  e  $c_r - c_B B^{-1} a_r > 0$ .

Nestas condições,  $B^{-1} a_r$  tem pelo menos uma componente positiva, caso contrário  $B^{-1} a_r \leq 0$  e, conseqüentemente,

$$d = \begin{bmatrix} -B^{-1} a_r \\ e_{r-m} \end{bmatrix},$$

com  $e_{r-m}$  identificando o  $(r-m)$ -ésimo vector da base canónica de  $\mathbb{R}^{n-m}$ , seria uma direcção extrema de  $S$  tal que  $c^T d = -c_B B^{-1} a_r + c_r > 0$ , o que contradiz 2. Fazendo  $x' = x^k + \tau d$ , com

$$\begin{aligned} \tau &= \min_{1 \leq i \leq m} \left\{ \frac{(B^1 b)_i}{(B^{-1} a_r)_i} : (B^1 a_r)_i > 0 \right\} \\ &= \frac{(B^{-1} b)_s}{(B^{-1} a_r)_s}, \end{aligned}$$

obtém-se  $x' \geq 0$  tal que

$$Ax' = Ax^k + \tau Ad = b + 0 = b$$

e ainda que  $x'$  tem no máximo  $m$  componentes positivas (dado que  $x'_s = 0$ ). Considerando a submatriz de  $A$ ,  $B'$ , obtida de  $B$  pela substituição de  $a_s$  por  $a_r$  (note-se que, dado que  $(B^{-1} a_r)_s \neq 0$ ,  $B'$  continua invertível<sup>12</sup>) conclui-se que  $x'$  é uma solução básica para  $S$  e, conseqüentemente, um ponto extremo, tal que

$$c^T x' = c^T x^k + \tau c^T d,$$

com  $\tau > 0$  e  $c^T d > 0$ . Logo,

$$c^T x' > c^T x^k,$$

o que é absurdo. □

Como consequência, do Teorema 2.13, vamos terminar este capítulo, sobre pontos extremos e direcções extremas de conjuntos convexos, com um resultado (cuja prova é proposta no Exercício 2.17) que nos garante a existência de um ponto extremo óptimo em problemas de optimização de uma funcional linear num poliedro convexo, com valor óptimo finito.

<sup>12</sup>Com efeito, suponha-se que  $a_r$  é combinação linear das colunas  $a_1, \dots, a_{s-1}, a_{s+1}, \dots, a_m$ . Então existe um  $m$ -uplo de escalares não todos nulos  $u^T = (\alpha_1, \dots, \alpha_m)$ , tais que  $a_r = \sum_{i=1}^m \alpha_i a_i$  com  $\alpha_s = 0$ . Logo, de  $a_r = Bu$  vem que  $u = B^{-1} a_r$ , com  $(B^{-1} a_r)_s = 0$ , o que contraria a hipótese.

**Teorema 2.14.** *Seja  $f$  uma funcional linear,  $f : \mathbb{R}^n \mapsto \mathbb{R}$ . Se  $\exists z^* \in \mathbb{R}$  tal que*

$$z^* = \max\{f(x) : x \in S\},$$

*então o valor  $z^*$  é atingido num ponto extremo de  $S$ .*

## 2.5 Exercícios

**Exercício 2.1.** *Dada a família de subconjuntos convexos de  $\mathbb{R}^n$ ,  $S_j$  com  $j \in J$ , prove que  $S = \bigcap_{j \in J} S_j$  é um conjunto convexo.*

**Exercício 2.2.** *Sendo  $T_i$ , com  $i \in I$ , uma colecção de subespaços afim, prove que  $\bigcap_{i \in I} T_i$  é um subespaço afim.*

**Exercício 2.3.** *Dado um subconjunto finito de pontos de  $\mathbb{R}^n$ ,  $S = \{x^1, \dots, x^p\}$ , prove que se  $p \geq n + 2$ , então  $S$  pode partir-se em dois subconjuntos disjuntos  $S_1$  e  $S_2$  tais que  $H(S_1) \cap H(S_2) \neq \emptyset$  <sup>13</sup>.*

**Sugestão:** *Sabendo que para  $p \geq n + 2$ , os pontos de  $S$  são dependentes afim, a partir da igualdade  $\sum_{i=1}^p \lambda_i x^i = 0$ , onde os  $\lambda_i$ s não são todos nulos e são tais que  $\sum_{i=1}^p \lambda_i = 0$ , podemos concluir a existência de um ponto  $x = \frac{1}{\lambda} \sum_{i=1}^{p_1} \alpha_i x^i = \frac{1}{\lambda} \sum_{j=1}^{p_2} \beta_j x^j$ , com  $\alpha_i \geq 0$  para  $i = 1, \dots, p_1$ ,  $\beta_j \geq 0$ , para  $j = 1, \dots, p_2$ ,  $p_1 + p_2 = p$  e  $\lambda = \sum_{i=1}^{p_1} \alpha_i = \sum_{j=1}^{p_2} \beta_j > 0$ .*

**Exercício 2.4.** *Sendo  $F = \{S_1, \dots, S_p\}$  uma família de  $p$  subconjuntos convexos de  $\mathbb{R}^n$ , com  $p \geq n + 1$ , prove que se toda a subfamília de  $n + 1$  conjuntos de  $F$  tem intersecção não vazia, então  $\bigcap_{1 \leq i \leq p} S_i \neq \emptyset$ . <sup>14</sup>*

**Sugestão:** *fazer a prova por indução sobre  $p$  (tendo em conta que o resultado é trivialmente válido para  $p = n + 1$ ), considerando o resultado válido para  $p = k \geq n + 1$  e aplicando o resultado do exercício anterior (teorema de Radon) ao conjunto de pontos  $S = \{x^i \in S_1 \cap \dots \cap S_{i-1} \cap S_{i+1} \dots \cap S_{k+1} : 1 \leq i \leq k + 1\}$ . A partir deste conjunto, por aplicação do teorema de Radon, pode supôr, sem perda de generalidade, que existem dois subconjuntos,  $T_1$  e  $T_2$ , tais que  $T = T_1 \cup T_2$ , com  $T_1 = \{x^1, \dots, x^r\}$ ,  $T_2 = \{x^{r+1}, \dots, x^{k+1}\}$  e que  $\exists x \in H(T_1) \cap H(T_2)$ . Finalmente, dada a convexidade dos  $S_i$ s, e dado que para  $i \leq r$ ,  $x^i \in S_{r+1} \cap \dots \cap S_{k+1}$  e, para  $j \geq r + 1$ ,  $x^j \in S_1 \cap \dots \cap S_r$ , tenha em consideração que  $S_{r+1} \cap \dots \cap S_{k+1} \supseteq H(T_1)$  e  $S_1 \cap \dots \cap S_r \supseteq H(T_2)$ .*

<sup>13</sup>Este resultado é conhecido por Teorema de Radon .

<sup>14</sup>Este resultado é conhecido por Teorema de Helly .

**Exercício 2.5.** Considerando os pontos de  $\mathbb{R}^2$ :  $x^1 = (1, 0)$ ,  $x^2 = (1, 3)$ ,  $x^3 = (4, 3)$ ,  $x^4 = (4, 0)$  e  $x = (\frac{7}{4}, \frac{5}{4})$ , pelo que  $x = \frac{1}{2}x^1 + \frac{1}{4}x^2 + \frac{1}{6}x^3 + \frac{1}{12}x^4$ , com uma técnica idêntica à utilizada na prova do teorema de Caratheodory, exprima  $x$  como combinação linear convexa de  $x^1$ ,  $x^2$  e  $x^3$ .

**Exercício 2.6.** Supondo que  $C_1$  é um subconjunto convexo de  $\mathbb{R}^n$  e que  $C_2 = \mathbb{R}_+^n = \{x \in \mathbb{R}^n : x \geq 0\}$ , verifique que  $C_1 + C_2 = \{x \in \mathbb{R}^n : \exists \bar{x} \in C_1, \bar{x} \leq x\}$ .

**Exercício 2.7.** Supondo que  $M$  é compacto e  $N$  é fechado prove que  $M + N$  é fechado.

**Sugestão:** Considere uma sucessão  $\{x_k\}$  de  $M + N$  convergente para  $x$  (pelo que, para cada  $k \in \mathbb{N} \exists y_k \in M$  e  $\exists z_k \in N$  tal que  $x^k = y^k + z^k$ ) e tendo em conta a compacidade de  $M$  considere a subsucessão de  $\{y^k\}$ ,  $\{y^{k'}\}$ , convergente para  $y \in M$ , donde vem que  $x^{k'} = y^{k'} + z^{k'} \Leftrightarrow z^{k'} = x^{k'} - y^{k'}$  e  $\{x^{k'} - y^{k'}\} \rightarrow x - y \in N$  (uma vez que  $N$  é fechado).

**Exercício 2.8.** Sendo  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  uma aplicação linear,  $A$  um subconjunto convexo de  $\mathbb{R}^n$  e  $B$  um subconjunto convexo de  $\mathbb{R}^m$  prove que tanto  $f(A) = \{f(x) : x \in A\}$  como  $f^{-1}(B) = \{x \in \mathbb{R}^n : f(x) \in B\}$  são conjuntos convexos.

**Exercício 2.9.** Seja  $u \in \mathbb{R}^n \setminus \{0\}$  e considere a aplicação  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  definida por  $f(x) = u^T x$ .

1. Mostre que  $f$  é uma funcional linear.
2. A partir de  $f$ , determine o hiperplano que contém o ponto  $u$ .
3. Considerando que o hiperplano determinado em b) divide  $\mathbb{R}^n$  em dois meios espaços, indique qual deles contém a origem.

**Exercício 2.10.** Seja  $S$  um subconjunto de  $\mathbb{R}^n$  fechado e não vazio e  $x \in \mathbb{R}^n$  tal que  $x \notin S$ .

1. Prove que  $\exists x^* \in S$  tal que  $x^*$  é um ponto de  $S$  mais próximo de  $x$  (no sentido da métrica induzida pela norma euclidiana).
2. Prove que se  $x^*$  é um ponto nas condições da alínea anterior, então

$$\|x - x^*\| > 0$$

(onde  $\|\cdot\|$  denota a norma euclidiana).

**Exercício 2.11.** Sendo  $A$  e  $B$  dois subconjuntos de  $\mathbb{R}^n$ , prove que  $A \cap B \neq \emptyset$  sse  $0 \in A - B$ .

**Exercício 2.12.** Supondo que a matriz  $A$  é tal que  $A \in \mathbb{R}^{m \times n}$  e que a sua característica é  $n$ , prove que a matriz  $A^T A$  é não singular e que a matriz  $AA^T$  pode ser singular.

**Exercício 2.13.** Dado um cone  $K$ , gerado pelas combinações lineares não negativas dos vectores de  $\mathbb{R}^m: \{u_1, \dots, u_n\}$  e um vector  $b \in \mathbb{R}^m$ , prove que ou (i)  $b \in K$  ou (ii)  $\forall y \in K^*$  (cone polar negativo de  $K$ )  $y^T b > 0$ , mas não as duas coisas.

**Sugestão:** Recorra ao Lema de Farkas.

**Exercício 2.14.** Sendo  $S = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$ , com  $n > m$ ,  $A \in \mathbb{R}^{m \times n} \setminus \{0\}$  e  $b \in \mathbb{R}^m$  prove cada uma das seguintes afirmações:

1.  $S$  é um conjunto convexo.
2. Se  $r(A) = m$ , então  $r(A) = r([A, b])$ <sup>15</sup>.
3. Se  $r(A) \neq r([A, b])$ , então  $S = \emptyset$ .
4. Se  $b = 0$  e  $S \neq \{0\}$ , então qualquer ponto de  $S$  é uma direcção de  $S$  (ou seja,  $S$  é um cone).
5. O número de soluções básicas admissíveis de  $S$  é não superior a  $\binom{n}{m} = \frac{n!}{m!(n-m)!}$ .
6. O número de direcções extremas de  $S$  é não superior a  $(n-m)\binom{n}{m} = (n-m)\frac{n!}{m!(n-m)!}$ .

**Exercício 2.15.** Considere a região poliédrica  $S = \{x \in \mathbb{R}^5 : Ax = b, x \geq 0\}$ , onde

$$A = \begin{bmatrix} -5 & 1 & 1 & 0 & 0 \\ -3 & 2 & 0 & 1 & 0 \\ 1 & -4 & 0 & 0 & 1 \end{bmatrix} \quad e \quad b = \begin{bmatrix} 5 \\ 12 \\ 2 \end{bmatrix},$$

e responda às seguintes questões:

1. Faça o esboço da região de  $\mathbb{R}^2$  que corresponde à região  $S$ .
2. Determine duas soluções básicas para  $S$  e indique, na figura da alínea anterior, os vértices a que correspondem.
3. Determine as duas direcções extremas de  $S$ .

**Exercício 2.16.** Supondo que  $X$  é um politopo cujos pontos extremos são

$$x^1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad x^2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \quad x^3 = \begin{bmatrix} 4 \\ 3 \end{bmatrix} \quad e \quad x^4 = \begin{bmatrix} 4 \\ 0 \end{bmatrix},$$

com recurso ao teorema da representação, verifique se

$$x = \begin{bmatrix} \frac{7}{4} \\ \frac{5}{4} \end{bmatrix} \in X.$$

<sup>15</sup> $r(B)$  denota a característica da matriz  $B$ .

**Exercício 2.17.** *Faça a prova do Teorema 2.14.*

**Exercício 2.18.** *Dada uma matriz invertível  $B = [b_1, \dots, b_m] \in \mathbb{R}^{m \times m}$  e um vector  $u \in \mathbb{R}^m$  tal que  $(B^{-1}u)_i \neq 0$ , prove que a matriz  $\tilde{B}$  obtida de  $B$  substituindo a  $i$ -ésima coluna por  $u$ , continua invertível.*

**Exercício 2.19.** *Considerando o polítopo  $X$  do Exercício 2.16, determine:*

$$1. \max\{2x_1 + x_2 : \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in X\}.$$

$$2. \min\{2x_1 + x_2 : \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in X\}.$$

*(Sugestão: utilize o Teorema 2.14).*

**Exercício 2.20.** *Considere a região poliédrica  $T = \{x \in \mathbb{R} : -1 \leq x \leq 1\}$ , a qual (introduzindo variáveis de desvio) se pode definir pela região:*

$$U = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, x_2, x_3 \geq 0\}.$$

1. *Prove que ao ponto  $0 \in T$  corresponde uma solução básica para  $U$ .*
2. *Verifique se a solução básica  $(0, 1, 1) \in U$  é um ponto extremo e analise a existência ou não de uma contradição entre a conclusão obtida e o enunciado do Teorema 2.10.*

**Exercício 2.21.** *Seja  $S$  um subconjunto convexo fechado de  $\mathbb{R}^n$ , prove que*

$$\forall x^* \in fr(S) \exists c \in \mathbb{R}^n \setminus \{0\}$$

*tal que  $c^T x^* = \max\{c^T x : x \in S\}$ .*

**Exercício 2.22.** *Sejam  $U$  e  $V$  dois subconjuntos convexos de  $\mathbb{R}^n$  tais que  $W = U - V$  e seja  $c \in \mathbb{R}^n \setminus \{0\}$ . Suponha ainda que  $z^* = \min\{c^T x : x \in W\} > 0$ .*

1. *Prove que  $W$  é um conjunto convexo.*
2. *Prove que existe um hiperplano que separa estritamente  $U$  e  $V$ .*

## Capítulo 3

# Método Simplex e Algumas Variantes

Apesar de, a partir da década de 80 se terem desenvolvido diferentes métodos para *Programação Linear*, baseados em técnicas de pontos interiores, o método simplex é, ainda hoje, o método mais popular para este tipo de problemas. O método simplex foi desenvolvido no princípio da década de 50 por George Dantzig e consiste, basicamente, numa pesquisa sistemática das soluções básicas admissíveis (conceito a esclarecer mais adiante). Com ele, a partir de uma solução básica admissível inicial, gera-se uma sequência de soluções básicas admissíveis, escolhendo-se, em cada passo, uma de entre as soluções básicas admissíveis que correspondem aos vértices adjacentes (ligados por uma aresta) ao vértice corrente, de modo que o correspondente valor da função objectivo seja "melhorado" (maior se o problema é de maximização ou menor no caso contrário). Quando, a partir da solução básica admissível corrente, não existe uma solução básica adjacente para a qual a função objectivo apresente um valor "melhorado", conforme se demonstrará, pode concluir-se que a solução corrente é uma solução óptima para o problema em causa.

Uma vez que o número de soluções básicas de qualquer programa linear é finito, por aplicação do método simplex, ao fim de um número finito de passos, ou se determina uma solução básica óptima, ou se conclui que o problema não tem óptimo finito. Por outro lado, embora o método simplex apenas pesquise soluções básicas admissíveis com vista à determinação da "melhor" delas, os resultados a estudar neste capítulo garantem-nos a suficiência desta pesquisa, com vista à determinação de uma solução óptima para um programa linear.

### 3.1 O algoritmo simplex básico

Apesar de, a partir da década de 80 se terem desenvolvido diferentes métodos para *Programação Linear*, baseados em técnicas de pontos interiores, o método simplex é, ainda hoje, o método mais popular para este tipo de problemas. O mé-

todo simplex foi desenvolvido no princípio da década de 50 por George Dantzig e consiste, basicamente, numa pesquisa sistemática das soluções básicas admissíveis. Com ele, a partir de uma solução básica admissível inicial, gera-se uma sequência de soluções básicas admissíveis, escolhendo-se, em cada passo, uma de entre as soluções básicas admissíveis que correspondem aos vértices adjacentes (ligados por uma aresta) ao vértice corrente, de modo que o correspondente valor da função objectivo seja "melhorado" (maior se o problema é de maximização ou menor no caso contrário). Quando, a partir da solução básica admissível corrente, não existe uma solução básica adjacente para a qual a função objectivo apresente um valor "melhorado", pode concluir-se que a solução corrente é uma solução óptima para o problema em causa.

Uma vez que o número de soluções básicas de qualquer programa linear é finito, por aplicação do método simplex, ao fim de um número finito de passos, ou se determina uma solução básica óptima, ou se conclui que o problema não tem óptimo finito. Por outro lado, embora o método simplex apenas pesquise soluções básicas admissíveis (vértices) com vista à determinação da "melhor" delas, o Teorema 2.14 garante-nos a suficiência desta pesquisa, com vista à determinação de uma solução óptima para um programa linear.

A partir de agora vamos considerar o programa linear  $PL(A, b, c)$  na forma padrão, que consiste no problema de optimização

$$\min\{c^T x : x \in S\} \quad (3.1)$$

com  $S = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$ , onde se supõe que a matriz  $A \in \mathbb{R}^{m \times n}$  tem característica  $m$  (i.e., igual ao número de linhas). Vamos modificar (3.1), exprimindo um ponto genérico de  $S$  como sendo a adição de um ponto particular com uma combinação linear não negativa de certas direcções do espaço nulo da matriz  $A$ .

Se  $x^0$  é um ponto arbitrário de  $S$ , i.e., uma solução arbitrária admissível para o  $PL(A, b, c)$  de minimização na forma padrão, e se  $d^1, d^2, \dots, d^p$ , são  $p$  vectores linearmente independentes que geram o subespaço nulo da matriz  $A$ ,  $Ker(A)$ , então  $x \in S$  sse  $\exists \lambda_1, \lambda_2, \dots, \lambda_p \in \mathbb{R}$  tais que

$$x = x^0 + \sum_{j=1}^p \lambda_j d^j \geq 0.$$

Com efeito, se  $x \in S$ , então  $x \geq 0$  e, dado que  $x = x^0 + (x - x^0)$ , vem que  $x = x^0 + d$ , com  $d = x - x^0$ . Logo  $Ad = A(x - x^0) = 0$ , ou seja,  $d$  pertence ao subespaço nulo de  $A$  e, conseqüentemente,  $\exists \lambda_1, \lambda_2, \dots, \lambda_p \in \mathbb{R}$  tais que  $d = \sum_{j=1}^p \lambda_j d^j$ . Reciprocamente, supondo  $x = x^0 + \sum_{j=1}^p \lambda_j d^j \geq 0$ , então  $Ax = Ax^0 + \sum_{j=1}^p \lambda_j Ad^j = b + 0 = b$  e, conseqüentemente,  $x \in S$ . Nestas condições, o programa  $PL(A, b, c)$  pode escrever-se na forma:

$$\min\{c^T(x^0 + \sum_{j=1}^p \lambda_j d^j) : \sum_{j=1}^p \lambda_j d^j \geq 0\},$$

ou seja,

$$\begin{aligned} \min \quad & c^T x^0 + \sum_{j=1}^p \lambda_j c^T d^j, \\ \text{s. a.} \quad & x^0 + \sum_{j=1}^p \lambda_j d^j \geq 0. \end{aligned}$$

Supondo, sem perda de generalidade, que  $x^0$  corresponde à solução básica admissível,  $x^0 = \begin{bmatrix} x_B^0 \\ x_N^0 \end{bmatrix}$ , com  $x_B^0 > 0$  e  $x_N^0 = 0$ , vem que

$$Ax^0 = b \Leftrightarrow [B, N] \begin{bmatrix} x_B^0 \\ x_N^0 \end{bmatrix} = b \Leftrightarrow Bx_B^0 + Nx_N^0 = b \Leftrightarrow Bx_B^0 + N0 = b,$$

e, conseqüentemente, que  $x_B^0 = B^{-1}b$ .

Adicionalmente, identificando por  $a_j$  a  $j$ -ésima coluna de  $A$ , e fazendo

$$d^{j-m} = \begin{bmatrix} -B^{-1}a_j \\ e_{j-m} \end{bmatrix}, \text{ para } j = m+1, \dots, n,$$

onde  $e_{j-m}$  corresponde ao  $(j-m)$ -ésimo vector da base canônica de  $\mathbb{R}^{n-m}$ , obtêm-se  $n-m$  vectores  $d^{j-m}$ , para  $j = m+1, \dots, n$ , que constituem uma base para o espaço nulo da matriz  $A$ . Com efeito, a característica de  $A$  é  $m$  (logo  $\dim(Ker(A)) = n-m$ ) e  $\forall j \in \{m+1, \dots, n\}$ ,  $Ad^{j-m} = 0$ , uma vez que

$$Ad^{j-m} = [B, N] \begin{bmatrix} -B^{-1}a_j \\ e_{j-m} \end{bmatrix} = -BB^{-1}a_j - a_j = 0.$$

Decompondo  $c^T$  nos correspondentes subvectores linha  $c_B$  e  $c_N$ , i.e., fazendo  $c^T = [c_B, c_N]$ , o programa  $PL(A, b, c)$  anterior fica

$$\begin{aligned} \min \quad & c_B B^{-1}b - \sum_{j=m+1}^n \lambda_j [c_B, c_N] \begin{bmatrix} B^{-1}a_j \\ -e_{j-m} \end{bmatrix} \\ \text{s. a.} \quad & \begin{bmatrix} x_B^0 \\ x_N^0 \end{bmatrix} - \sum_{j=m+1}^n \lambda_j \begin{bmatrix} B^{-1}a_j \\ -e_{j-m} \end{bmatrix} \geq 0. \end{aligned}$$

Nestas condições, pode concluir-se que os escalares  $\lambda_j$  (para  $j = m+1, \dots, n$ ) vão corresponder aos valores das variáveis não básicas,  $x_{m+1}, x_{m+2}, \dots, x_n$ , pelo que o  $PL(A, b, c)$  se pode escrever na forma

$$\begin{aligned} \min \quad & c_B B^{-1}b - \sum_{j=m+1}^n x_j (c_B B^{-1}a_j - c_j) \\ \text{s. a.} \quad & \begin{bmatrix} B^{-1}b \\ 0 \end{bmatrix} - \sum_{j=m+1}^n x_j \begin{bmatrix} B^{-1}a_j \\ -e_{j-m} \end{bmatrix} \geq 0. \end{aligned}$$

Organizando-se os diferentes coeficientes de acordo com o quadro

	$x_{m+1}$	$\dots$	$x_{m+k}$	$\dots$	$x_n$	
$x_1$						
$\vdots$						
$x_q$	$B^{-1}a_{m+1}$	$\dots$	$B^{-1}a_{m+k}$	$\dots$	$B^{-1}a_n$	$B^{-1}b$
$\vdots$						
$x_m$						
$x_{m+1}$	-1	$\dots$	0	$\dots$	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$x_{m+k}$	0	$\dots$	-1	$\dots$	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$x_n$	0	$\dots$	0	$\dots$	-1	0
	$c_B B^{-1}a_{m+1} - c_{m+1}$	$\dots$	$c_B B^{-1}a_{m+k} - c_{m+k}$	$\dots$	$c_B B^{-1}a_n - c_n$	$c_B B^{-1}b$

obtêm-se todas as componentes em jogo.

A partir do quadro, vamos seleccionando um vector, de entre os que verificam a inequação  $[c_b, c_N] \begin{bmatrix} B^{-1}a_j \\ -e_{j-m} \end{bmatrix} = [c_B B^{-1}a_j - c_j] > 0$ . Assim, supondo que se seleccionou o  $m+k$ -ésimo, ou seja,  $\begin{bmatrix} B^{-1}a_{m+k} \\ e_k \end{bmatrix}$ , incrementando (tanto quanto possível) a correspondente variável  $x_{m+k}$ , mantendo-se as demais variáveis não básicas com valor nulo, vem

$$x^1 = \begin{bmatrix} B_0^{-1}b \\ \mathbf{0} \end{bmatrix} - x_{m+k} \begin{bmatrix} B^{-1}a_{m+k} \\ -e_k \end{bmatrix} = \begin{bmatrix} B^{-1}b - x_{m+k}B^{-1}a_{m+k} \\ 0 \\ x_{m+k} \\ 0 \end{bmatrix},$$

(onde 0 denota um subvector de componentes nulas de dimensão adequada). Por outro lado,  $c^T x^1 = c^T x^0 - x_j c^T \begin{bmatrix} B^{-1}a_{m+k} \\ -e_k \end{bmatrix} = c_B B^{-1}b - x_{m+k}(c_B B^{-1}a_{m+k} - c_{m+k})$ .

Como consequência, se  $x_{m+k} > 0$  e  $c_B B^{-1}a_{m+k} - c_{m+k} > 0$ , então podemos concluir a desigualdade  $c^T x^1 < c^T x^0$ , havendo todo o interesse em incrementar, tanto quanto possível,  $x_{m+k}$ . Porém, as condições de não negatividade obrigam a que se tenha  $x^1 \geq 0$ , o que é equivalente a  $B^{-1}b - x_{m+k}B^{-1}a_{m+k} \geq 0$ , ou seja,

$$x_{m+k}(B^{-1}a_{m+k})_i \leq (B^{-1}b)_i \quad \forall i \in \{1, \dots, m\}.$$

Assim, para que se cumprir a condição  $x^1 \geq 0$ , basta que se verifiquem as desigualdades

$$x_{m+k} \leq \frac{(B^{-1}b)_i}{(B^{-1}a_{m+k})_i} \quad \forall i \in \{1, \dots, m\} \text{ tal que } (B^{-1}a_{m+k})_i > 0.$$

Caso não exista  $i \in \{1, \dots, m\}$  tal que  $(B^{-1}a_{m+k})_i > 0$ , podemos concluir que o problema não tem óptimo finito (porquê?). Caso contrário, com o objectivo de incrementar o mais possível a variável  $x_{m+k}$ , faz-se

$$x_{m+k} = \min\left\{ \frac{(B^{-1}b)_i}{(B^{-1}a_{m+k})_i} : (B^{-1}a_{m+k})_i > 0, i = 1, \dots, m \right\}.$$

Supondo que  $\min\left\{\frac{(B^{-1}b)_i}{(B^{-1}a_k)_i} : (B^{-1}a_k)_i > 0, i = 1, \dots, m\right\} = \frac{(B^{-1}b)_q}{(B^{-1}a_k)_q}$ , podemos concluir que a  $k$ -ésima componente de  $x^1$  (que para  $x^0$  era nula) vem agora igual a  $\frac{(B^{-1}b)_q}{(B^{-1}a_k)_q}$  e as restantes componentes não básicas (e por conseguinte nulas) para  $x^0$  permanecem não básicas. Finalmente, as demais componentes (que eram básicas para  $x^0$ ) tomam os valores:

$$x_i = (B^{-1}b)_i - x_{m+k}(B^{-1}a_{m+k})_i, \text{ para } i = 1, \dots, m.$$

Deve observe-se ainda que, nestas condições, a  $q$ -ésima componente (que é positiva para  $x^0$ ) de  $x^1$  é igual a zero, dado que  $x_q = (B^{-1}b)_q - \frac{(B^{-1}b)_q}{(B^{-1}a_{m+k})_q}(B^{-1}a_{m+k})_q$ .

Voltando ao quadro anterior, identificando o vector coluna  $B^{-1}a_{m+j}$  por  $y_j$ , e fazendo  $z_{m+j} - c_{m+j} = c_B B^{-1}a_{m+j} - c_{m+j}$ , para  $j = 1, \dots, p$ , com  $p = n - m$ ,

	$x_{m+1}$	$\dots$	$x_{m+k}$	$\dots$	$x_n$	
$x_1$						$B^{-1}b$
$\vdots$						
$x_q$	$y_1$	$\dots$	$y_k$	$\dots$	$y_p$	
$\vdots$						
$x_m$						
$x_{m+1}$	-1	$\dots$	0	$\dots$	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$x_k$	0	$\dots$	-1	$\dots$	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$x_n$	0	$\dots$	0	$\dots$	-1	0
	$z_1 - c_{m+1}$	$\dots$	$z_k - c_{m+k}$	$\dots$	$z_p - c_n$	$c_B B^{-1}b$

estamos em condições de proceder à respectiva actualização, i.e., à determinação do quadro associado à nova solução básica admissível  $x^1$ , onde vão constar as componentes dos vectores definidores dos deslocamentos admissíveis (ao longo das arestas do poliedro que partem do novo vértice). Assim, no caso da variável escolhida para entrar na base ser  $x_{m+k}$  (pelo que  $z_k - c_{m+k} > 0$ ), dizemos que  $y_k$  é a coluna pivotal. Consequentemente, no pressuposto de que (tal como se descreveu) se obtém

$$\frac{(B^{-1}b)_q}{y_{qk}} = \min_{1 \leq i \leq m} \left\{ \frac{(B^{-1}b)_i}{y_{ik}} : y_{ik} > 0 \right\},$$

ou seja, a variável a sair da base é  $x_q$  e a  $q$ -ésima linha do quadro designa-se por linha pivotal.

A linha pivotal determina a variável a sair da base e a coluna pivotal determina a variável a entrar para a base.

O coeficiente determinado pelo cruzamento da coluna pivotal com a linha pivotal designa-se por "pivot" e a última linha do quadro designa-se por linha de custos reduzidos.

Consequentemente, a solução básica corrente passa a ser

$$x^1 = \begin{bmatrix} x_1^1 \\ \vdots \\ x_{q-1}^1 \\ 0 \\ x_{q+1}^1 \\ \vdots \\ x_m^1 \\ 0 \\ \vdots \\ x_{m+k}^1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} x_1^0 \\ \vdots \\ x_{q-1}^0 \\ x_q^0 \\ x_{q+1}^0 \\ \vdots \\ x_m^0 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix} - \frac{(B^{-1}b)_q}{y_{qk}} \begin{bmatrix} y_{1k} \\ \vdots \\ y_{(q-1)k} \\ y_{qk} \\ y_{(q+1)k} \\ \vdots \\ y_{mk} \\ 0 \\ \vdots \\ -1 \\ \vdots \\ 0 \end{bmatrix},$$

e os vectores de deslocamento associados à variável  $x_q$  (que agora é não básica relativamente a  $x^1$ ) e às variáveis  $x_j$ , com  $j \in \{m+1, \dots, n\} \setminus \{k\}$  deverão ser os vectores do subespaço nulo<sup>2</sup> de  $A$  determinados por

$$\begin{bmatrix} y_{1j}^1 \\ \vdots \\ y_{qj}^1 \\ \vdots \\ y_{mj}^1 \\ 0 \\ \vdots \\ y_{(m+k)j}^1 \\ \vdots \\ -1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} y_{1j} \\ \vdots \\ y_{qj} \\ \vdots \\ y_{mj} \\ 0 \\ \vdots \\ 0 \\ \vdots \\ -1 \\ \vdots \\ 0 \end{bmatrix} - \frac{y_{qj}}{y_{qk}} \begin{bmatrix} y_{1k} \\ \vdots \\ y_{qk} \\ \vdots \\ y_{mk} \\ 0 \\ \vdots \\ -1 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \text{ para } j \in \{1, \dots, p\} \setminus \{k\}.$$

---

<sup>2</sup>Observe-se que um vector pertence a um dado subespaço sse é combinação linear de vectores desse subespaço.

Por sua vez  $y_k^1$  vem dado por

$$\begin{bmatrix} y_{1k}^1 \\ \vdots \\ y_{qk}^1 \\ \vdots \\ y_{mk}^1 \\ 0 \\ \vdots \\ y_{(m+k)k}^1 \\ \vdots \\ 0 \end{bmatrix} = \frac{-1}{y_{qk}} \begin{bmatrix} y_{1k} \\ \vdots \\ y_{qk} \\ \vdots \\ y_{mk} \\ 0 \\ \vdots \\ -1 \\ \vdots \\ 0 \end{bmatrix}.$$

No que diz respeito à linha de custos reduzidos, a atualização das suas componentes vem dada por  $c^T y_j^1 = z_j^1 - c_{m+j} = c^T y_j - \frac{y_{qj}}{y_{qk}} c^T y_k = (z_j - c_{m+j}) - \frac{y_{qj}}{y_{qk}} (z_k - c_{m+k})$ , para  $j \in \{1, \dots, p\} \setminus \{k\}$  e  $c^T y_k^1 = z_k^1 - c_{m+k} = \frac{-1}{y_{qk}} (z_k - c_{m+k})$ . Com todas estas atualizações que se designam por operações de pivotação (nas quais o “pivot” é a  $q$ -ésima componente de  $y_k$ , ou seja,  $y_{qk}$ ), trocando a  $k$ -ésima linha do quadro com a  $q$ -ésima, bem como as posições relativas das correspondentes variáveis, obtém-se o novo quadro (quadro associado à nova solução básica admissível  $x^1$ ), numa forma idêntica à inicial.

Denotando por  $B_1$  a matriz básica associada a este quadro, ou seja, a que se obtém de  $B$  trocando a coluna  $a_q$  por  $a_{m+k}$  ( $B_1 = [a_1 \dots a_{q-1} \ a_{m+k} \ a_{q+1} \dots a_m]$ ), após reordenação das variáveis, obtém-se os vectores do subespaço nulo da matriz  $A_1$  (obtida de  $A$ , após a correspondente troca de colunas),

$$y_j^1 = \begin{bmatrix} B_1^{-1} a_j \\ -e_{j-m} \end{bmatrix}, \text{ para } j \in \{m+1, \dots, n\} \setminus \{k\} \text{ e } y_q^1 = \begin{bmatrix} B_1^{-1} a_q \\ -e_k \end{bmatrix}.$$

### 3.1.1 Descrição do método simplex

Para descrever o método simplex, vamos considerar apenas o que se designa por quadro simplex reduzido <sup>3</sup> que, relativamente ao quadro anterior, é constituído pelas  $m$  primeiras linhas acrescidas da linha de custos reduzidos. Nestas condições, o quadro simplex reduzido associado a  $x^0$  corresponde ao quadro que a seguir se indica.

	$x_{m+1}$	$\dots$	$x_{m+k}$	$\dots$	$x_n$	
$x_1$	$y_{11}$	$\dots$	$y_{1k}$	$\dots$	$y_{1p}$	$(B^{-1}b)_1$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$x_q$	$y_{q1}$	$\dots$	$y_{qk}$	$\dots$	$y_{qp}$	$(B^{-1}b)_q$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$x_m$	$y_{m1}$	$\dots$	$y_{mk}$	$\dots$	$y_{mp}$	$(B^{-1}b)_m$
	$z_1 - c_{m+1}$	$\dots$	$z_k - c_k$	$\dots$	$z_n - c_n$	$c_B B^{-1} b$

<sup>3</sup>Esta designação aparece em oposição à de quadro simplex completo no qual, além das colunas que constam no reduzido, aparecem, também, as relativas à matriz identidade (que correspondem às componentes básicas correntes).

Deve observar-se que, neste último quadro, existe toda a informação necessária para se produzir o quadro simplex reduzido associado a  $x^1$ . A determinação (actualização) dos quadros subsequentes faz-se tal como se descreveu anteriormente. Com efeito, suponha que se actualiza o quadro simplex reduzido associado a  $x^s$  (de modo a obter-se o quadro simplex reduzido associado a  $x^{s+1}$ ), escolhendo para coluna pivotar a coluna  $y_k^s$  associada à variável não básica  $x_{j_k}$ . Então, podemos concluir que, em problemas de maximização,  $z_k^s - c_{j_k} < 0$  e, em problemas de minimização,  $z_k^s - c_{j_k} > 0$ . Nestas condições, sendo  $J_{N_s} = \{j_1, \dots, j_k, \dots, j_p\}$  o conjunto de índices não básicos e  $I_{B_s} = \{i_1, \dots, i_q, \dots, i_m\}$  o conjunto de índices básicos (note-se que  $I_{B_s} = \{1, \dots, n\} \setminus J_{N_s}$ ), considerando o quadro simplex corrente

	$x_{j_1}$	...	$x_{j_k}$	...	$x_{j_p}$	
$x_{i_1}$	$y_{11}^s$	...	$y_{1k}^s$	...	$y_{1p}^s$	$(B_s^{-1}b)_1$
$\vdots$	$\vdots$	...	$\vdots$	...	$\vdots$	$\vdots$
$x_{i_q}$	$y_{q1}^s$	...	$y_{qk}^s$	...	$y_{qp}^s$	$(B_s^{-1}b)_q$
$\vdots$	$\vdots$	...	$\vdots$	...	$\vdots$	$\vdots$
$x_{i_m}$	$y_{m1}^s$	...	$y_{mk}^s$	...	$y_{mp}^s$	$(B_s^{-1}b)_m$
	$z_1 - c_{j_1}$	...	$z_k - c_{j_k}$	...	$z_p - c_{j_p}$	$c_{B_s} B_s^{-1} b$

podemos fazer a actualização do quadro, de acordo com o seguinte procedimento:

1. Determina-se a linha pivotar, calculando

$$\frac{(B_s^{-1}b)_q}{y_{qk}^s} = \min_{1 \leq i \leq m} \left\{ \frac{(B_s^{-1}b)_i}{y_{ik}^s} : y_{ik}^s > 0 \right\}.$$

2. Divide-se a  $q$ -ésima linha do quadro pelo "pivot", ou seja,

$$y_{qj}^{s+1} = \frac{y_{qj}^s}{y_{qk}^s}, \quad \forall j \in \{1, 2, \dots, p\} \setminus \{k\}$$

$$(B_{s+1}^{-1}b)_q = \frac{(B_s^{-1}b)_q}{y_{qk}^s}.$$

3. Substituem-se os elementos da  $k$ -ésima coluna pelos simétricos dos quocientes das suas componentes pelo "pivot", ou seja,

$$y_{ik}^{s+1} = \frac{-y_{ik}^s}{y_{qk}^s} \quad \forall i \in \{1, 2, \dots, m\} \setminus \{q\};$$

$$z_k^{s+1} - c_{j_k} = -\frac{z_k^s - c_{j_k}}{y_{qk}^s}.$$

4. Substitui-se o "pivot" pelo seu inverso, i.e.,  $y_{qk}^{s+1} = \frac{1}{y_{qk}^s}$ .

5. Actualizam-se os restantes elementos do quadro, fazendo

$$y_{ij}^{s+1} = y_{ij}^s - \frac{y_{qj}^s}{y_{qk}^s} y_{ik}^s \quad \forall i \in \{1, \dots, m\} \setminus \{q\} \quad \text{e} \quad \forall j \in \{1, \dots, p\} \setminus \{k\};$$

$$(B_{s+1}^{-1}b)_i = (B_s^{-1}b)_i - \frac{(B_s^{-1}b)_q}{y_{qk}^s} y_{ik}^s \quad \forall i \in \{1, 2, \dots, m\} \setminus \{q\};$$

$$z_r^{s+1} - c_{j_r} = z_r^s - c_{j_r} - \frac{y_{qr}^s}{y_{qk}^s} (z_k^s - c_{j_k}) \quad \forall j_r \in J_{N_s} \setminus \{j_k\}.$$

6. Actualiza-se o valor da função objectivo para a nova solução, fazendo

$$c_{B_{s+1}} B_{s+1}^{-1} b = c_{B_s} B_s^{-1} b - \frac{(B_s^{-1} b)_q}{y_{qk}^s} (z_k^s - c_{j_k}).$$

7. Finalmente, a variável básica  $x_{i_q}$  passa a ser não básica e é substituída pela variável  $x_{j_k}$ .

As operações efectuadas no passo 5 são conhecidas por operações de *pivotação* e a regra de cálculo associada por regra em estrela.

Em termos matriciais, dada a partição da matriz  $A = [B_s, N_s]$ , onde  $B_s$  denota a submatriz básica corrente, tal que  $B_s^{-1} b \geq 0$ , e a correspondente partição do gradiente da função objectivo,  $c^T = [c_{B_s}, c_{N_s}]$ , o quadro simplex reduzido associado a  $x^s = \begin{bmatrix} B_s^{-1} b \\ 0 \end{bmatrix}$ , corresponde ao quadro

$$\begin{array}{c|cc} & x_{N_s} & \\ \hline x_{B_s} & B_s^{-1} N_s & B_s^{-1} b \\ \hline & c_{B_s} B_s^{-1} N_s - c_{N_s} & c_{B_s} B_s^{-1} b \end{array}$$

**Teorema 3.1.** *Aplicando o método simplex a um programa linear de minimização na forma padrão,  $PL(A, b, c)$ , seja  $x^s = \begin{bmatrix} x_{B_s}^s \\ x_{N_s}^s \end{bmatrix}$  a solução básica admissível corrente,  $c_{B_s}$  e  $c_{N_s}$  os subvectores linha de  $c^T = [c_{B_s}, c_{N_s}]$  que correspondem aos subvectores de componentes básicas  $x_{B_s}^s$  e não básicas  $x_{N_s}^s$  e  $z_j = c_{B_s} B_s^{-1} N_s e_j$  para  $j = 1, \dots, p$ . Então, podemos tirar as seguintes conclusões (que constituem a base de todo o processo de pesquisa utilizado no método simplex):*

1. Se  $z_j - (c_{N_s})_j \leq 0 \forall j \in \{1, \dots, p\}$ , então  $x^s$  é uma solução óptima.
2. Se  $x^s$  é uma solução óptima e  $x_{B_s}^s > 0$ , então  $z_j - (c_{N_s})_j \leq 0 \forall j \in \{1, \dots, p\}$ .
3. Se  $\exists j \in \{1, \dots, p\}$  tal que  $z_j - (c_{N_s})_j > 0$  e  $B_s^{-1} N_s e_j \leq 0$ , então o programa linear de minimização na forma padrão  $PL(A, b, c)$  não tem óptimo finito.
4. Se  $\exists j \in \{1, \dots, p\}$  tal que  $z_j - (c_{N_s})_j > 0$  e pelo menos uma componente do vector coluna  $B_s^{-1} N_s e_j$  é positiva, então fazendo

$$x^{s+1} = \begin{bmatrix} x_{B_s}^s \\ x_{N_s}^s \end{bmatrix} - \theta_j \begin{bmatrix} B_s^{-1} N_s e_j \\ -e_j \end{bmatrix},$$

onde  $\theta_j = \min_{1 \leq i \leq m} \left\{ \frac{(x_{B_s}^s)_i}{(B_s^{-1} N_s e_j)_i} : (B_s^{-1} N_s e_j)_i > 0 \right\}$ , obtém-se uma nova solução básica admissível,  $x^{s+1}$ , tal que  $c^T x^{s+1} \leq c^T x^s$ .

**Demonstração.** 1.  $\forall x \in S, \exists \theta_1, \theta_2, \dots, \theta_p \geq 0$  tais que

$$x = \begin{bmatrix} x_{B_s}^s \\ x_{N_s}^s \end{bmatrix} - \theta_1 \begin{bmatrix} B_s^{-1} N_s e_1 \\ -e_1 \end{bmatrix} - \theta_2 \begin{bmatrix} B_s^{-1} N_s e_2 \\ -e_2 \end{bmatrix} - \dots - \theta_p \begin{bmatrix} B_s^{-1} N_s e_p \\ -e_p \end{bmatrix}.$$

Logo, tendo em conta que  $c^T x = [c_{B_s}, c_{N_s}] x$ , obtém-se

$$\begin{aligned} c^T x &= c_{B_s} x_{B_s}^s - \sum_{j=1}^p \theta_j (c_{B_s} B_s^{-1} N_s e_j - (c_{N_s})_j) \\ &= c_{B_s} x_{B_s}^s - \sum_{j=1}^p \theta_j (z_j - (c_{N_s})_j). \end{aligned}$$

Consequentemente, uma vez que  $x \geq 0 \Rightarrow (\theta_j \geq 0 \ \forall j \in \{1, \dots, p\})$  e, de acordo com a hipótese,  $z_j - (c_{N_s})_j \leq 0 \ \forall j \in \{1, \dots, p\}$ ,  $c^T x \geq c^T x^s$ . Assim, tendo em conta que  $x$  é um ponto arbitrário de  $S$ ,  $x^s = \begin{bmatrix} x_{B_s}^s \\ x_{N_s}^s \end{bmatrix}$  é uma solução óptima para o programa linear de minimização na forma padrão  $PL(A, b, c)$ .

2. Se  $x^s = \begin{bmatrix} x_{B_s}^s \\ x_{N_s}^s \end{bmatrix}$  é uma solução óptima para o programa linear  $PL(A, b, c)$  tal que  $x_{B_s}^s > 0$ , então  $\forall j \in \{1, \dots, p\}$ ,  $\exists \theta_j > 0$  tal que

$$x = \begin{bmatrix} x_{B_s}^s \\ x_{N_s}^s \end{bmatrix} - \theta_j \begin{bmatrix} B_s^{-1} N_s e_j \\ -e_j \end{bmatrix} \geq 0,$$

i.e., tal que  $x \in S$ , pelo que

$$c^T x = c^T x^s - \theta_j (c_{B_s} B_s^{-1} N_s e_j - (c_{N_s})_j) = c^T x^s - \theta_j (z_j - (c_{N_s})_j).$$

Tendo em conta que, de acordo com a hipótese,  $c^T x \geq c^T x^s$ , obtém-se

$$z_j - (c_{N_s})_j \leq 0 \ \forall j \in \{1, \dots, p\}.$$

3. Supondo que  $\exists j \in \{1, \dots, p\}$  tal que  $z_j - (c_{N_s})_j > 0$  e  $B_s^{-1} N_s e_j \leq 0$ ,  $\forall \theta_j \geq 0$

$$x = \begin{bmatrix} x_{B_s}^s \\ x_{N_s}^s \end{bmatrix} - \theta_j \begin{bmatrix} B_s^{-1} N_s e_j \\ -e_j \end{bmatrix} \geq 0 \text{ e } c^T x = c^T x^s - \theta_j (z_j - (c_{N_s})_j).$$

Se  $\theta_j \rightarrow \infty$ , então  $c^T x = c^T x^s - \theta_j (z_j - (c_{N_s})_j) \rightarrow -\infty$  e o programa linear de minimização na forma padrão  $PL(A, b, c)$  não tem óptimo finito.

4. Seja  $j \in \{1, \dots, p\}$ , tal que  $z_j - (c_{N_s})_j > 0$ , e suponha que  $B_s^{-1} N_s e_j$  tem pelo menos uma das suas componentes positiva, então, para

$$\theta_j = \min_{1 \leq i \leq m} \left\{ \frac{(x_{B_s})_i}{(B_s^{-1} N_s e_j)_i} : (B_s^{-1} N_s e_j)_i > 0 \right\} = \frac{(x_{B_s})_q}{(B_s^{-1} N_s e_j)_q}$$

e  $x = \begin{bmatrix} x_{B_s}^s \\ x_{N_s}^s \end{bmatrix} - \theta_j \begin{bmatrix} B_s^{-1} N_s e_j \\ -e_j \end{bmatrix}$ , obtém-se  $x \geq 0$  (porquê?),  $c^T x \geq c^T x^s$  (porquê?) e  $x$  continua a ser uma solução básica admissível, uma vez que a  $q$ -ésima componente de  $x$  é nula (porquê?) e as restantes componentes não básicas de  $x^s$ , com excepção da  $j$ -ésima que agora tem o valor  $\theta_j$ , continuam a ser não básicas para  $x$ . Logo, substituindo na submatriz básica  $B_s$  a  $q$ -ésima coluna pela  $j$ -ésima coluna de  $N_s$ , obtém-se uma matriz  $B_{s+1}$  que continua a ser básica (como exercício, faça a prova desta última afirmação). □

Segue-se a descrição formal (algorítmica) do método simplex.

- Algoritmo (método simplex).

1. Determinar uma solução básica admissível inicial<sup>4</sup>,  $x^0 = \begin{bmatrix} x_B^0 \\ x_N^0 \end{bmatrix}$  (e, por conseguinte, uma base inicial  $B_0$ , à qual está associado o conjunto de índices básicos  $I_{B_0}$  e não básicos  $J_{N_0} = \{1, \dots, n\} \setminus I_{B_0} = \{j_1, \dots, j_p\}$ ), construir o respectivo quadro simplex reduzido

<sup>4</sup>Os métodos usualmente utilizados para a determinação de uma solução básica admissível inicial (método das duas fases e método do "big-M") serão alvo de estudo mais adiante.

	$x_{j_1}$	$\dots$	$x_{j_k}$	$\dots$	$x_{j_p}$	
$x_{B_0}$	$B_0^{-1}N_0$					$B_0^{-1}b$
	$z_1 - c_{j_1}$	$\dots$	$z_k - c_{j_k}$	$\dots$	$z_p - c_{j_p}$	$c_{B_0}B_0^{-1}b$

onde  $[z_1 - c_{j_1}, \dots, z_p - c_{j_p}] = c_{B_0}B_0^{-1}N_0 - c_{N_0}$ , fazer  $s = 0$  e passar a 2.

2. Se o problema é de

(a) *maximização*, determinar  $z_k - c_{j_k} = \min\{z_r - c_{j_r} : j_r \in J_{N_s}\}$  e se  $z_k - c_{j_k} \geq 0$  então STOP (a solução básica admissível corrente é ótima) senão passar a 3.

(b) *minimização*, determinar  $z_{j_k} - c_{j_k} = \max\{z_{j_i} - c_{j_i} : j_i \in J_{N_s}\}$  e se  $z_{j_k} - c_{j_k} \leq 0$  então STOP (a solução básica admissível corrente é ótima) senão passar a 3.

3. Se  $B_s^{-1}a_{j_k} \leq 0$  então STOP (o problema não tem ótimo finito) senão passar a 4.

4. Fazendo  $y_k = B_s^{-1}a_{j_k}$ , determinar  $q$  tal que

$$\frac{(B_s^{-1}b)_q}{y_{qk}} = \min_{1 \leq i \leq m} \left\{ \frac{(B_s^{-1}b)_i}{y_{ik}} : y_{ik} > 0 \right\}$$

e passar a 5.

5. Actualizar o quadro simplex, tomando para “pivot”  $y_{qk}$ , fazer  $s = s + 1$  e voltar a 2.

• Fim do algoritmo.

O critério de escolha da coluna pivotal adoptado neste algoritmo (o qual decorre dos cálculos do passo 2) é conhecido por critério de Dantzig e corresponde à escolha do vector de deslocamento com taxa de variação da função objectivo mais favorável.

**Exemplo 3.1.** Para ilustrar a aplicação deste algoritmo, vamos resolver o seguinte programa linear de minimização na forma padrão:

$$\begin{aligned} \min \quad & c^T x \\ \text{s.a.} \quad & Ax = b \\ & x \geq 0, \end{aligned}$$

$$\text{onde } c^T = [15, 30, 40, 10, 10, 10], A = \begin{bmatrix} 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 1 & -1 \end{bmatrix} \text{ e } b = \begin{bmatrix} 10.000 \\ 20.000 \\ 25.000 \end{bmatrix}.$$

**Solução.** Considerando  $x^0 = (10.000, 20.000, 25.000, 0, 0, 0)^T$ , como solução básica admissível inicial (à qual está associada a submatriz básica  $B = I$ ), obtém-se o quadro simplex reduzido

	$x_4$	$x_5$	$x_6$	
$x_1$	-1	0	0	10.000
$x_2$	1	-1	0	20.000
$x_3$	0	1	-1	25.000
	5	0	-50	1.750.000

Uma vez que, por aplicação do passo 2 (determinação do  $\max\{z_1 - c_4, z_2 - c_5, z_3 - c_6\}$ ), se obtém  $5 = \max\{5, 0, -50\}$  e, dado que as condições do passo 3 não se verificam (i.e., nem todas as componentes da 1ª coluna do quadro são não positivas), por aplicação do passo 4, a escolha do “pivot” recai sobre o coeficiente assinalado com um bordo, ou seja, o que se obtém do cruzamento da 2ª linha (linha pivotal) com a 1ª coluna (coluna pivotal) do quadro. Como consequência, após a actualização, obtém-se o quadro simplex associado à solução básica admissível  $x^1 = (30.000, 0, 25.000, 20.000, 0, 0)^T$ ,

	$x_2$	$x_5$	$x_6$	
$x_1$	1	-1	0	30.000
$x_4$	1	-1	0	20.000
$x_3$	0	1	-1	25.000
	-5	5	-50	1.650.000

Aplicando novamente os passos 2, 3 e 4, obtém-se a 2ª coluna como coluna pivotal e a 3ª linha como linha pivotal. Procedendo à actualização determinada pelo passo 5, obtém-se o quadro simplex óptimo

	$x_2$	$x_3$	$x_6$	
$x_1$	1	1	-1	55.000
$x_4$	1	1	-1	45.000
$x_5$	0	1	-1	25.000
	-5	-5	-45	1.525.000

o qual determina, como solução óptima para o programa linear em estudo, a solução

$$x^* = \begin{bmatrix} 55.000 \\ 0 \\ 0 \\ 45.000 \\ 25.000 \\ 0 \end{bmatrix}.$$

Se no quadro simplex óptimo existir, na linha de custos reduzidos, um (ou mais) coeficiente(s) nulo(s), então a solução óptima não é única (existem soluções óptimas alternativas). Com efeito, suponha que o quadro a seguir se refere a um problema de minimização e que a solução básica corrente é óptima. Adicionalmente, denotando os conjuntos de índices básicos e não básicos por  $J_B$  e  $J_N = \{j_1, \dots, j_p\}$ , respectivamente, considere que  $\exists j_k \in J_N$ , tal que  $z_k - c_{j_k} = 0$ .

	$x_{j_1}$	...	$x_{j_k}$	...	$x_{j_p}$	
$x_B$	$B^{-1}N$					$B^{-1}b$
	$z_1 - c_{j_1} \dots z_k - c_{j_k} \dots z_p - c_{j_p}$					$c_B B^{-1}b$

Se  $B^{-1}b > 0$ , então fazendo  $x = \begin{bmatrix} B^{-1}b \\ 0 \end{bmatrix} - \theta_k \begin{bmatrix} y_k \\ -e_k \end{bmatrix}$ , onde  $y_k = B^{-1}a_{j_k}$  e

$$\theta_k = \begin{cases} \mu, & \text{se } \{y_{ik} : y_{ik} > 0 \wedge 1 \leq i \leq m\} = \emptyset \\ \min_{1 \leq i \leq m} \left\{ \frac{(B^{-1}b)_i}{y_{ik}} : y_{ik} > 0 \right\}, & \text{no caso contrário,} \end{cases}$$

onde  $\mu$  é um escalar positivo arbitrariamente grande, vem que  $x$  é uma solução distinta da fornecida pelo quadro simplex óptimo e é tal que

$$c^T x = c^T \begin{bmatrix} B^{-1}b \\ 0 \end{bmatrix} - \theta_k c^T \begin{bmatrix} y_k \\ -e_k \end{bmatrix} = c_B B^{-1}b - \theta_k(z_k - c_{j_k}) = c_B B^{-1}b$$

(uma vez que  $z_k - c_{j_k} = 0$ ), ou seja, é também óptima.

Verifica-se ainda que qualquer combinação linear convexa de diferentes soluções óptimas é uma solução óptima.

### 3.1.2 Determinação de uma solução básica admissível inicial

Vamos agora analisar algumas formas de ultrapassar o (em geral) desconhecimento de uma solução básica admissível inicial. Com efeito, embora em qualquer programa linear a região admissível possa tomar a forma  $S' = \{x \in \mathbb{R}^n : Ax \leq b, x \geq 0\}$  e a partir dela, acrescentando as variáveis de desvio, a forma

$$S'' = \left\{ \begin{bmatrix} x \\ s \end{bmatrix} \in \mathbb{R}^{n+m} : [A, I] \begin{bmatrix} x \\ s \end{bmatrix} = b, x, s \geq 0 \right\}$$

da qual, com facilidade, se retira a solução básica definida por  $x = 0$  e  $s = b$ , existem muitas situações onde algumas das componentes de  $b$  são negativas, pelo que  $\begin{bmatrix} 0 \\ b \end{bmatrix}$  não é admissível para  $S'$ .

As técnicas que vamos analisar para ultrapassar o desconhecimento de uma solução básica admissível para um programa linear de minimização na forma padrão, i.e., para o programa linear

$$PL(A, b, c) \quad \min\{c^T x : x \in S\},$$

com  $S = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$ ,  $A \in \mathbb{R}^{m \times n}$  e  $r(A) = m$ , designam-se por *método das duas fases* e *método do "big-M"*.

Qualquer destes dois métodos exige que se transforme o sistema  $Ax = b$  para que se tenha  $b \geq 0$  (para o que basta multiplicar, adequadamente, cada uma das equações do sistema por 1 ou -1) e que se acrescentem variáveis artificiais,  $x_a$ , de modo a obter-se a região

$$S_{x_a} = \{x \in \mathbb{R}^n, x_a \in \mathbb{R}^m, Ax + x_a = b, x, x_a \geq 0\}.$$

O recurso a variáveis artificiais para a obtenção de uma solução admissível inicial foi publicado, pela primeira vez, em (Dantzig, 1951).

#### Método das duas fases

Com este método, a resolução do programa linear de minimização na forma padrão  $PL(A, b, c)$  divide-se em duas fases, a saber, a fase I e a fase II.

- A **fase I** consiste na resolução do programa linear auxiliar:

$$\min\{e^T x_a : \begin{bmatrix} x \\ x_a \end{bmatrix} \in S_{x_a}\},$$

onde  $e^T x_a = \sum_{j=1}^m x_{a_j}$ . Se o valor óptimo é positivo, isso significa que não existe  $x \geq 0$  tal que  $Ax = b$  (dado que se um tal  $x$  existisse, então  $\begin{bmatrix} x \\ 0 \end{bmatrix}$  seria uma solução admissível para este programa auxiliar, para a qual a função objectivo tomaria o valor nulo). Se o valor óptimo é nulo, então o programa linear de minimização na forma padrão  $PL(A, b, c)$  admite uma solução básica óptima da forma  $\begin{bmatrix} x^1 \\ 0 \end{bmatrix}$ , pelo que este programa é consistente.

- A **fase II** consiste em resolver o programa linear de minimização na forma padrão  $PL(A, b, c)$ , tomando para solução básica admissível inicial a solução  $x^1$  determinada na fase I, podendo aproveitar-se o quadro simplex óptimo da fase I para prosseguir com a fase II.

No caso de, no final da fase I, apesar do valor óptimo ser nulo, existir uma (ou mais) variável (variáveis) artificial (artificiais) na base, para a(s) retirar (no caso de serem várias o procedimento deve repetir-se para cada uma delas) basta proceder às operações de pivotação, tomando para linha pivotal a linha que lhe corresponde e para coluna pivotal uma de entre as colunas associadas a variáveis não artificiais e não básicas com componente não nula na linha pivotal. Neste caso, a eventualidade do “pivot” ser negativo não causa qualquer problema, uma vez que as variáveis básicas não nulas permanecem inalteradas e as que vierem a entrar para a base tomam o valor nulo. No caso de existir uma ou mais linhas associadas a variáveis artificiais básicas (com valor nulo) relativamente às quais não é possível determinar pivôs (coeficientes não nulos) associados a variáveis não artificiais e não básicas, tal significa que as restrições que lhes correspondem são redundantes (ou seja, são combinação linear das outras), pelo que podem ser retiradas do problema sem que isso afecte a respectiva região de admissibilidade. Com efeito, suponha-se que  $\xi_B = [\xi_1, \xi_2, \dots, \xi_q]^T$  é o  $q$ -uplo de valores das variáveis básicas não artificiais (as quais, sem perda de generalidade, se supõem ser as  $q$  primeiras, com  $q < m$ ), e que  $B_s$  é a base óptima obtida no final da fase I. Uma vez que

$$B_s^{-1}[A, b] = \begin{bmatrix} I_q & R & \xi_B \\ 0 & 0 & 0 \end{bmatrix} \Leftrightarrow [A, b] = B_s \begin{bmatrix} I_q & R & \xi_B \\ 0 & 0 & 0 \end{bmatrix},$$

onde  $I_q$  denota a matriz identidade de ordem  $q$  e  $R$  identifica uma matriz com  $q$  linhas e  $(n - q)$  colunas. Então, podemos concluir que a  $i$ -ésima linha de  $[A, b]$  vem dada por  $e_i^T [A, b] = e_i^T B_s \begin{bmatrix} I_q & R & \xi_B \\ 0 & 0 & 0 \end{bmatrix} = \sum_{j=1}^q (B_s)_{ij} (e_j^T [I_q R \xi_B])$ , para  $i = 1, \dots, m$ . Como consequência, todas as linhas de  $[A, b]$  são geradas pelas  $q$  linhas de  $[I_q R \xi_B]$ , pelo que  $[A, b]$  tem  $m - q$  linhas linearmente dependentes. Nestas condições, também se pode concluir que as  $m - q$  últimas linhas de  $[A, b]$  são combinação linear das  $q$  primeiras.

Segue-se um exemplo que ilustra a aplicação do método das duas fases.

**Exemplo 3.2.** *Vamos resolver o programa linear*

$$\begin{array}{rcll} \max & 3x_1 & - & x_2 \\ s. a. & 2x_1 & + & x_2 \geq 2 \\ & x_1 & + & 3x_2 \leq 3 \\ & & & x_2 \leq 4 \\ & x_1, & & x_2 \geq 0 \end{array} \quad (3.2)$$

pele método das duas fases.

**Solução.** Introduzindo as variáveis de desvio, obtém-se o programa linear na forma padrão:

$$\begin{array}{rcccccccc} \max & 3x_1 & - & x_2 & & & & & \\ \text{s. a.} & 2x_1 & + & x_2 & - & x_3 & & & = & 2 \\ & x_1 & + & 3x_2 & & & + & x_4 & = & 3 \\ & & & x_2 & & & & & + & x_5 = & 4 \\ x_1, & x_2, & x_3, & x_4, & x_5 & \geq & 0 & & & & \end{array}$$

• Fase I

$$\begin{array}{rcccccccc} \min & & & & & & & & & x_6 & \\ \text{s. a.} & 2x_1 & + & x_2 & - & x_3 & & & & + & x_6 = & 2 \\ & x_1 & + & 3x_2 & & & + & x_4 & & & = & 3 \\ & & & x_2 & & & & & + & x_5 & = & 4 \\ x_1, & x_2, & x_3, & x_4, & x_5, & x_6 & \geq & 0 & & & & \end{array}$$

1.

	$x_1$	$x_2$	$x_3$	
$x_6$	2	1	-1	2
$x_4$	1	3	0	3
$x_5$	0	1	0	4
	2	1	-1	2

2.

	$x_6$	$x_2$	$x_3$	
$x_1$	1/2	1/2	-1/2	1
$x_4$	-1/2	5/2	1/2	2
$x_5$	0	1	0	4
	-1	0	0	0

• Fase II

1.

	$x_2$	$x_3$	
$x_1$	1/2	-1/2	1
$x_4$	5/2	1/2	2
$x_5$	1	0	4
	5/2	-3/2	3

2.

	$x_2$	$x_4$	
$x_1$	3	1	3
$x_3$	5	2	4
$x_5$	1	0	4
	10	3	9

Tendo em conta o quadro óptimo obtido, conclui-se que  $x^* = \begin{bmatrix} 3 \\ 0 \\ 4 \\ 0 \\ 4 \end{bmatrix}$  é a solução

ótima do problema (3.2). ■

### Método do "big"-M

O método das duas fases obriga à resolução de dois programas lineares distintos (o da fase I e o da fase II). No método do "big-M" resolve-se um único problema e procura-se que a solução básica admissível a determinar esteja tão próxima da solução ótima quanto possível.

Com o método do big-M, a função objectivo do programa linear de minimização na forma padrão  $PL(A, b, c)$  é modificada com a introdução de um parâmetro de penalização  $M$ , ou seja, resolve-se o programa linear de minimização na forma padrão

modificado  $PL([A, I], b, \begin{bmatrix} c \\ M \\ \vdots \\ M \end{bmatrix})$ :

$$\begin{aligned} \min \quad & c^T x + M \sum_{j=1}^m x_{a_j} \\ \text{s. a.} \quad & Ax + x_a = b \\ & x, x_a \geq 0, \end{aligned}$$

onde  $M$  é uma constante positiva arbitrariamente grande.

Assim, sendo  $x^*$  uma solução ótima para o problema original, com  $M$  convenientemente escolhido,  $\begin{bmatrix} x^* \\ 0 \end{bmatrix}$  é uma solução ótima para o problema modificado. Com efeito, suponha que  $\begin{bmatrix} x' \\ x'_a \end{bmatrix}$ , com  $x'_a \neq 0$ , é uma solução ótima para o programa linear

de minimização na forma padrão  $PL([A, I], b, \begin{bmatrix} c \\ M \\ \vdots \\ M \end{bmatrix})$ , então  $c^T x' + M e^T x'_a \leq c^T x^*$ ,

com  $e^T = [1, \dots, 1]$ , o que é impossível,<sup>5</sup> para  $M > \frac{c^T x^* - c^T x'}{e^T x'_a}$ .

Conclui-se assim que se o programa linear de minimização na forma padrão  $PL(A, b, c)$  é consistente e tem ótimo finito, então, desde que  $M$  tome um valor suficientemente grande, o problema modificado apresenta o mesmo conjunto de soluções ótimas e o mesmo valor ótimo.

Deve observar-se que no caso do programa linear ser de maximização, com o método do big-M, o problema inicial é modificado subtraindo à respectiva função objectivo a soma das variáveis artificiais multiplicadas pelo parâmetro de penalização  $M$ . O exemplo a seguir ilustra, precisamente, este caso.

**Exemplo 3.3.** *Vamos resolver o programa linear de maximização na forma padrão do Exemplo 3.2, desta vez por aplicação do método do big-M.*

**Solução.** Introduzindo a variável artificial e o parâmetro de penalização obtém-se o

<sup>5</sup>Uma vez que  $x'_a \neq 0$ , podemos concluir que  $e^T x'_a > 0$ .

programa linear modificado:

$$\begin{array}{rcccccccc} \max & 3x_1 & - & x_2 & & & & - & Mx_6 \\ \text{s. a.} & 2x_1 & + & x_2 & - & x_3 & & + & x_6 & = & 2 \\ & x_1 & + & 3x_2 & & & + & x_4 & & = & 3 \\ & & & x_2 & & & & + & x_5 & = & 4 \\ & x_1, & & x_2, & & x_3, & & x_4, & & x_5, & & x_6 & \geq & 0 \end{array}$$

1.

	$x_1$	$x_2$	$x_3$	
$x_6$	2	1	-1	2
$x_4$	1	3	0	3
$x_5$	0	1	0	4
	-2M-3	-M+1	M	-2M

2.

	$x_6$	$x_2$	$x_3$	
$x_1$	1/2	1/2	-1/2	1
$x_4$	-1/2	5/2	1/2	2
$x_5$	0	1	0	4
	M+3/2	5/2	-3/2	3

3.

	$x_6$	$x_2$	$x_4$	
$x_1$	0	3	1	3
$x_3$	-1	5	2	4
$x_5$	0	1	0	4
	M	10	3	9

■

### 3.1.3 Determinação da inversa da base corrente

Supondo que a matriz básica admissível inicial (associada a  $x^0$ ) é a matriz identidade, ou seja,  $A = [I, N_0]$  (onde  $I = B_0$ ), e supondo que o quadro corrente determina a solução básica admissível  $x^s$  associada a base  $B_s$ , dado que

$$B_s^{-1}Ax^s = B_s^{-1}b \Leftrightarrow [B_s^{-1}, B_s^{-1}N_0]x^s = B_s^{-1}b,$$

a partir do correspondente quadro simplex completo

	$x$	
$x_{B_s}$	$B_s^{-1}A$	$B_s^{-1}b$
	$c_{B_s}B_s^{-1}A - c^T$	$c_{B_s}B_s^{-1}b$

com facilidade se determina  $B_s^{-1}$  (i.e., a inversa da base associada a  $x^s$ ), a qual ocupa a posição das colunas associadas à base inicial. É claro que, a partir do quadro simplex reduzido igualmente se consegue determinar a inversa da base associada a  $x^s$ , uma vez que todas as colunas que não constam neste quadro e fazem parte do quadro simplex completo são conhecidas (são as colunas da matriz identidade que correspondem às respectivas componentes básicas).

**Exemplo 3.4.** Para ilustrar a determinação da inversa da base corrente, a partir do respectivo quadro simplex, vamos determinar as inversas das bases associadas às soluções básicas admissíveis  $x^1$  e  $x^2$ , determinadas no Exemplo 3.1.

**Solução.** Tendo em conta que, no Exemplo 3.1, a base associada a  $x^0$  é a matriz identidade e as variáveis básicas iniciais são  $x_1$ ,  $x_2$  e  $x_3$ , do quadro simplex completo

associado a  $x^1$  (i.e., relativo a  $B_1$ ), vem

$$B_1^{-1}A = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \\ 1 & 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 1 & -1 \end{bmatrix},$$

donde  $B_1$  é constituída pelas 1<sup>a</sup>, 4<sup>a</sup> e 3<sup>a</sup> (por esta ordem) colunas de  $A$  e  $B_1^{-1} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ . Com efeito, denotando a  $j$ -ésima coluna de  $A$  por  $a_j$ , deve observar-se que

$$B_1^{-1}a_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad B_1^{-1}a_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{e} \quad B_1^{-1}a_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

ou seja,

$$B_1^{-1}[a_1 \ a_4 \ a_3] = I \quad \Leftrightarrow \quad [a_1 \ a_4 \ a_3] = B_1.$$

Por outro lado, tendo em conta que  $A = [I, N_0]$ , onde  $N_0$  é a submatriz não básica inicial (constituída pelas colunas de  $A$ ,  $a_4$ ,  $a_5$  e  $a_6$ ), vem  $B_1^{-1}A = [B_1^{-1}, B_1^{-1}N_0]$ . Relativamente ao quadro associado à solução óptima  $x^2$  (quadro óptimo), vem

$$B_2^{-1}A = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \\ 1 & 1 & 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 & 1 & -1 \end{bmatrix}$$

e, conseqüentemente, que  $B_2^{-1} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ . ■

Se em vez do quadro simplex completo dispusermos, apenas, do quadro simplex reduzido, a determinação da inversa da base corrente faz-se tendo em conta que as colunas associadas à base inicial (matriz identidade) que não constam do quadro corrente continuam básicas, pelo que são vectores da base canónica determinados pela posição, no quadro, das componentes básicas que lhes estão associadas. No Exemplo 3.1, a partir do quadro simplex reduzido óptimo, embora a coluna associada a  $x_1$  (única variável de entre as variáveis básicas iniciais que permanece básica) não conste no quadro, pela posição que ocupa, conclui-se que a coluna que lhe corresponde em  $B^{-1}$  é precisamente o  $1^0$  vector da base canónica.

## 3.2 Método simplex revisto

Embora a aplicação do método simplex, confinada à actualização dos quadros simplex reduzidos, proporcione alguma economia de esforço (nomeadamente em relação à utilização do quadro simplex completo), é ainda possível obter alguma redução no esforço computacional associado às operações de pivotação e, bem ainda, restringir o espaço de memória reservado aos sucessivos quadros simplex que se vão actualizando, recorrendo ao que se designa por método simplex revisto que foi introduzido em (Dantzig, 1953) e que se resume, essencialmente, à actualização da parte do quadro simplex onde se encontra a inversa da base corrente. Com efeito, conforme se referiu anteriormente, se

a submatriz básica inicial é a matriz identidade (o que é sempre possível de acontecer), então, após qualquer número de actualizações do quadro, nas entradas relativas a essa matriz identidade, encontramos a inversa da base corrente. Logo, actualizando essa parte do quadro, disponibilizamos a inversa da base corrente, com a qual determinamos os elementos estritamente essenciais ao desenvolvimento do método simplex, tal como a seguir se descreve.

Supondo que se pretende resolver o programa linear de minimização na forma padrão  $PL(A, b, c)$ , com o método simplex revisto, para além dos dados associados a este problema (como sejam, a matriz  $A$ , o vector de termos independentes  $b$  e o gradiente da função objectivo  $c$ ), para o  $s$ -ésimo quadro, apenas se torna necessário armazenar os elementos que constam no quadro a seguir

$$\begin{array}{c|c|c} \hline & & \\ \hline x_{B_s} & B_s^{-1} & B_s^{-1}b \\ \hline & w^s & c_{B_s} B_s^{-1}b \\ \hline \end{array}$$

onde  $x_{B_s}$  denota o  $m$ -uplo de componentes básicas,  $B_s^{-1}$  a inversa da base corrente e  $w^s = c_{B_s} B_s^{-1}$  (note-se ainda que  $c_{B_s} B_s^{-1}b = w^s b$ ). A partir deste quadro, com facilidade se obtém toda a informação associada ao quadro simplex:

$$\begin{array}{c|c|c} \hline & x_{N_s} & \\ \hline x_{B_s} & Y^s & B_s^{-1}b \\ \hline & z_1^s - c_{j_1} \quad \dots \quad z_p^s - c_{j_p} & c_{B_s} B_s^{-1}b \\ \hline \end{array}$$

com  $J_{N_s} = \{j_1, \dots, j_p\}$  e  $Y^s = B_s^{-1}N_s$ .

Com o método simplex revisto, a escolha da coluna pivotal pode fazer-se com recurso ao critério de Dantzig, determinando-se, para o  $s$ -ésimo quadro simplex,

$$z_k^s - c_{j_k} = \max\{z_r^s - c_{j_r} = w^s a_{j_r} - c_{j_r} : j_r \in J_{N_s}\},$$

ou, simplesmente, adoptando para coluna pivotal a primeira coluna  $\begin{bmatrix} y_k^s \\ z_k^s - c_{j_k} \end{bmatrix} = \begin{bmatrix} B_s^{-1}a_{j_k} \\ w^s a_{j_k} - c_{j_k} \end{bmatrix}$  que verifica a inequação  $z_k^s - c_{j_k} > 0$ , que é, geralmente, o que se faz em problemas com muitas variáveis.

Se  $\max\{z_r^s - c_{j_r} = w^s a_{j_r} - c_{j_r} : j_r \in J_{N_s}\} \leq 0$ , então podemos concluir que  $w^s A - c^T \leq 0$  e, conseqüentemente, que a solução básica admissível corrente é óptima.

Após a escolha da coluna pivotal, anteriormente considerada, se

$$\frac{(B_s^{-1}b)_q}{y_{qk}^s} = \min_{1 \leq i \leq m} \left\{ \frac{(B_s^{-1}b)_i}{y_{ik}^s} : y_{ik}^s > 0 \right\},$$

então, destacando a  $k$ -ésima coluna não básica

$$\begin{bmatrix} y_k^s \\ z_k^s - c_{j_k} \end{bmatrix} = \begin{bmatrix} B_s^{-1}a_{j_k} \\ w^s a_{j_k} - c_{j_k} \end{bmatrix}, \quad (3.3)$$

a determinação de  $B_{s+1}^{-1}$ ,  $B_{s+1}^{-1}b$ ,  $w^{s+1}$  e  $c_{B_{s+1}} B_{s+1}^{-1}b$  faz-se aplicando as operações de pivotação usuais (regra em estrela), considerando a coluna (3.3) como pivotal.

$$x_{B_s} \left[ \begin{array}{c|c} B_s^{-1} & B_s^{-1}b \\ \hline w^s & c_{B_s} B_s^{-1}b \end{array} \right] \left[ \begin{array}{c} y_{1k}^s \\ \vdots \\ y_{qk}^s \\ \vdots \\ y_{mk}^s \\ \hline z_k^s - c_{jk} \end{array} \right]$$

Note-se que da actualização da linha de custos reduzidos para os índices,  $i$ , associados às colunas que actualmente coincidem com as de  $B_s^{-1}$  (e que, por conseguinte, inicialmente coincidem com a matriz identidade que, por simplicidade consideramos posicionada nas primeiras colunas de  $A$ ), para  $i = 1, \dots, m$ , vem

$$\begin{aligned} w_i^{s+1} - c_i &= c_{B_{s+1}} B_{s+1}^{-1} e_i - c_i && \text{(dado que } e_i \text{ é a } i\text{-ésima coluna de } A) \\ &= z_i^{s+1} - c_i \\ &= z_i^s - c_i - (z_k^s - c_{jk}) \frac{y_{qi}^s}{y_{qk}^s} && \text{(aplicação da regra em estrela)} \\ &= c_{B_s} B_s^{-1} e_i - c_i - (z_k^s - c_{jk}) \frac{y_{qi}^s}{y_{qk}^s} \\ &= w_i^s - c_i - (z_k^s - c_{jk}) \frac{y_{qi}^s}{y_{qk}^s}, \end{aligned}$$

donde  $w_i^{s+1} = w_i^s - (z_k^s - c_{jk}) \frac{y_{qi}^s}{y_{qk}^s}$ , ou seja, a actualização da última linha do quadro do método simplex revisto faz-se (como no caso do quadro reduzido) de acordo com a regra em estrela.

Por outro lado, deve observar-se ainda que a actualização deste quadro corresponde à sua multiplicação, à esquerda, pela matriz quadrada de ordem  $m$ :

$$G^s = \begin{bmatrix} 1 & 0 & \dots & g_{1q}^s & \dots & 0 & 0 \\ 0 & 1 & \dots & g_{2q}^s & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \vdots & g_{qq}^s & \vdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & g_{(m-1)q}^s & \dots & 1 & 0 \\ 0 & 0 & \dots & g_{mq}^s & \dots & 0 & 1 \end{bmatrix}$$

ou seja, pela matriz que difere da identidade, unicamente, na  $q$ -ésima coluna, para a qual se tem

$$g_{iq}^s = \frac{-y_{ik}^s}{y_{qk}^s} \quad \forall i \in \{1, \dots, m\} \setminus \{q\} \text{ e } g_{qq}^s = \frac{1}{y_{qk}^s}.$$

Com efeito, a determinação da coluna  $y_j^{s+1}$  ( $j \neq k$ ) vem dada por

$$y_j^{s+1} = G^s y_j^s = \begin{bmatrix} 1 & 0 & \dots & \frac{-y_{1k}^s}{y_{qk}^s} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \frac{1}{y_{qk}^s} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \frac{-y_{mk}^s}{y_{qk}^s} & \dots & 1 \end{bmatrix} \begin{bmatrix} y_{1j}^s \\ \vdots \\ y_{qj}^s \\ \vdots \\ y_{mj}^s \end{bmatrix} = \begin{bmatrix} y_{1j}^s \\ \vdots \\ 0 \\ \vdots \\ y_{mj}^s \end{bmatrix} + y_{qj}^s \begin{bmatrix} \frac{-y_{1k}^s}{y_{qk}^s} \\ \vdots \\ \frac{1}{y_{qk}^s} \\ \vdots \\ \frac{-y_{mk}^s}{y_{qk}^s} \end{bmatrix}.$$

Consequentemente, denotando por  $G^1, G^2, \dots, G^s$ , as matrizes relativas às sucessivas operações de pivotação, a partir da inversa da matriz básica inicial  $B_0^{-1}$  (que em geral é a matriz identidade) obtém-se  $w^0 = c_{B_0} B_0^{-1}$  e a sequência,

$$\begin{array}{ll} B_1^{-1} = G^1 B_0^{-1} & \text{e } w^1 = c_{B_1} G^1 B_0^{-1}, \\ B_2^{-1} = G^2 G^1 B_0^{-1} & \text{e } w^2 = c_{B_2} G^2 G^1 B_0^{-1}, \\ \vdots & \vdots \\ B_s^{-1} = G^s \dots G^2 G^1 B_0^{-1} & \text{e } w^s = c_{B_s} G^s \dots G^2 G^1 B_0^{-1}. \end{array}$$

Nestas condições, basta armazenar as matrizes  $G^r$  (as quais ficam completamente caracterizadas pelo vector que as distingue da matriz identidade e pela sua posição enquanto vector coluna) para, em cada actualização, se determinar o quadro associado ao método simplex revisto.

**Exemplo 3.5.** *Vamos aplicar o método simplex revisto na resolução do programa linear do Exemplo 3.1.*

**Solução.** Aplicando o método simplex revisto, a partir da matriz básica inicial associada às variáveis  $x_1, x_2$  e  $x_3$  (matriz identidade) obtém-se a sequência de quadros que a seguir se indica (com \* assinalando o custo reduzido que determina a variável a entrar para a base e o "pivot" circunscrito por um bordo quadrangular):

1.	$x_1$ $x_2$ $x_3$	<table style="width: 100%; border-collapse: collapse;"> <tr><td style="border-right: 1px solid black; padding: 2px 5px;">1</td><td style="padding: 2px 5px;">0</td><td style="padding: 2px 5px;">0</td></tr> <tr><td style="border-right: 1px solid black; padding: 2px 5px;">0</td><td style="padding: 2px 5px;">1</td><td style="padding: 2px 5px;">0</td></tr> <tr><td style="border-right: 1px solid black; padding: 2px 5px;">0</td><td style="padding: 2px 5px;">0</td><td style="padding: 2px 5px;">1</td></tr> <tr style="border-top: 1px solid black;"><td style="border-right: 1px solid black; padding: 2px 5px;">15</td><td style="padding: 2px 5px;">30</td><td style="padding: 2px 5px;">40</td></tr> </table>	1	0	0	0	1	0	0	0	1	15	30	40	<table style="width: 100%; border-collapse: collapse;"> <tr><td style="padding: 2px 5px;">10.000</td><td style="padding: 2px 5px;"><math>w^1 a_4 - c_4 = 5^*</math></td></tr> <tr><td style="padding: 2px 5px;">20.000</td><td style="padding: 2px 5px;"><math>w^1 a_5 - c_5 = 0</math></td></tr> <tr><td style="padding: 2px 5px;">25.000</td><td style="padding: 2px 5px;"><math>w^1 a_6 - c_6 = -50</math></td></tr> <tr style="border-top: 1px solid black;"><td style="padding: 2px 5px;">1.750.000</td><td></td></tr> </table>	10.000	$w^1 a_4 - c_4 = 5^*$	20.000	$w^1 a_5 - c_5 = 0$	25.000	$w^1 a_6 - c_6 = -50$	1.750.000		<table style="width: 100%; border-collapse: collapse;"> <tr><td style="border-bottom: 1px solid black; padding: 2px 5px;"><math>x_4</math></td></tr> <tr><td style="padding: 2px 5px;">-1</td></tr> <tr><td style="border: 1px solid black; padding: 2px 5px; text-align: center;">1</td></tr> <tr><td style="padding: 2px 5px;">0</td></tr> <tr style="border-top: 1px solid black;"><td style="padding: 2px 5px;">5</td></tr> </table>	$x_4$	-1	1	0	5
1	0	0																											
0	1	0																											
0	0	1																											
15	30	40																											
10.000	$w^1 a_4 - c_4 = 5^*$																												
20.000	$w^1 a_5 - c_5 = 0$																												
25.000	$w^1 a_6 - c_6 = -50$																												
1.750.000																													
$x_4$																													
-1																													
1																													
0																													
5																													
2.	$x_1$ $x_4$ $x_3$	<table style="width: 100%; border-collapse: collapse;"> <tr><td style="border-right: 1px solid black; padding: 2px 5px;">1</td><td style="padding: 2px 5px;">1</td><td style="padding: 2px 5px;">0</td></tr> <tr><td style="border-right: 1px solid black; padding: 2px 5px;">0</td><td style="padding: 2px 5px;">1</td><td style="padding: 2px 5px;">0</td></tr> <tr><td style="border-right: 1px solid black; padding: 2px 5px;">0</td><td style="padding: 2px 5px;">0</td><td style="padding: 2px 5px;">1</td></tr> <tr style="border-top: 1px solid black;"><td style="border-right: 1px solid black; padding: 2px 5px;">15</td><td style="padding: 2px 5px;">25</td><td style="padding: 2px 5px;">40</td></tr> </table>	1	1	0	0	1	0	0	0	1	15	25	40	<table style="width: 100%; border-collapse: collapse;"> <tr><td style="padding: 2px 5px;">30.000</td><td style="padding: 2px 5px;"><math>w^2 a_2 - c_2 = -5</math></td></tr> <tr><td style="padding: 2px 5px;">20.000</td><td style="padding: 2px 5px;"><math>w^2 a_5 - c_5 = 5^*</math></td></tr> <tr><td style="padding: 2px 5px;">25.000</td><td style="padding: 2px 5px;"><math>w^2 a_6 - c_6 = -50</math></td></tr> <tr style="border-top: 1px solid black;"><td style="padding: 2px 5px;">1.650.000</td><td></td></tr> </table>	30.000	$w^2 a_2 - c_2 = -5$	20.000	$w^2 a_5 - c_5 = 5^*$	25.000	$w^2 a_6 - c_6 = -50$	1.650.000		<table style="width: 100%; border-collapse: collapse;"> <tr><td style="border-bottom: 1px solid black; padding: 2px 5px;"><math>x_5</math></td></tr> <tr><td style="padding: 2px 5px;">-1</td></tr> <tr><td style="padding: 2px 5px;">-1</td></tr> <tr><td style="border: 1px solid black; padding: 2px 5px; text-align: center;">1</td></tr> <tr style="border-top: 1px solid black;"><td style="padding: 2px 5px;">5</td></tr> </table>	$x_5$	-1	-1	1	5
1	1	0																											
0	1	0																											
0	0	1																											
15	25	40																											
30.000	$w^2 a_2 - c_2 = -5$																												
20.000	$w^2 a_5 - c_5 = 5^*$																												
25.000	$w^2 a_6 - c_6 = -50$																												
1.650.000																													
$x_5$																													
-1																													
-1																													
1																													
5																													
3.	$x_1$ $x_4$ $x_5$	<table style="width: 100%; border-collapse: collapse;"> <tr><td style="border-right: 1px solid black; padding: 2px 5px;">1</td><td style="padding: 2px 5px;">1</td><td style="padding: 2px 5px;">1</td></tr> <tr><td style="border-right: 1px solid black; padding: 2px 5px;">0</td><td style="padding: 2px 5px;">1</td><td style="padding: 2px 5px;">1</td></tr> <tr><td style="border-right: 1px solid black; padding: 2px 5px;">0</td><td style="padding: 2px 5px;">0</td><td style="padding: 2px 5px;">1</td></tr> <tr style="border-top: 1px solid black;"><td style="border-right: 1px solid black; padding: 2px 5px;">15</td><td style="padding: 2px 5px;">25</td><td style="padding: 2px 5px;">35</td></tr> </table>	1	1	1	0	1	1	0	0	1	15	25	35	<table style="width: 100%; border-collapse: collapse;"> <tr><td style="padding: 2px 5px;">55.000</td><td style="padding: 2px 5px;"><math>w^3 a_2 - c_2 = -5</math></td></tr> <tr><td style="padding: 2px 5px;">45.000</td><td style="padding: 2px 5px;"><math>w^3 a_3 - c_3 = -5</math></td></tr> <tr><td style="padding: 2px 5px;">25.000</td><td style="padding: 2px 5px;"><math>w^3 a_6 - c_6 = -45</math></td></tr> <tr style="border-top: 1px solid black;"><td style="padding: 2px 5px;">1.525.000</td><td></td></tr> </table>	55.000	$w^3 a_2 - c_2 = -5$	45.000	$w^3 a_3 - c_3 = -5$	25.000	$w^3 a_6 - c_6 = -45$	1.525.000							
1	1	1																											
0	1	1																											
0	0	1																											
15	25	35																											
55.000	$w^3 a_2 - c_2 = -5$																												
45.000	$w^3 a_3 - c_3 = -5$																												
25.000	$w^3 a_6 - c_6 = -45$																												
1.525.000																													

O método simplex revisto apresenta vantagens "computacionais" relevantes, relativamente à versão original, quando o número de colunas da matriz de decisão,  $A$ , é significativamente superior ao número de linhas (situação que, na prática, é muito comum). Com efeito, no caso revisto, para cada actualização do quadro, as operações de pivotação são executadas, unicamente, sobre uma matriz com  $m$  linhas e  $m + 1$  colunas e, a estas operações de pivotação, apenas se adiciona o esforço "computacional" relativo à determinação da coluna pivotal (cálculo dos custos reduzidos  $wa_j - c_j$ ), à actualização do vector  $w$  ( $c_B B^{-1}$ ) e ao cálculo (opcional) do valor corrente da função objectivo ( $wb$ ). Por sua vez, na versão original do método simplex, cada actualização do quadro (reduzido) implica a execução de operações de pivotação sobre uma matriz com  $m + 1$  linhas e  $(n - m) + 1$  colunas. ■

### 3.3 Convergência do método simplex

Tendo presente o algoritmo simplex, no pressuposto de que não existem soluções básicas degeneradas, é fácil provar a convergência finita deste método, no sentido em que após um número finito de operações de pivotação, ou se atinge uma situação que nos permite concluir a não existência de ótimo finito ou se determina uma solução básica ótima. Com efeito, na ausência de degenerescência, após cada actualização do quadro, obtém-se uma nova solução básica admissível (vértice) relativamente à qual a função objectivo toma um valor mais favorável (maior se o problema é de maximização ou menor se o problema é de minimização). Nestas condições, vão-se obtendo soluções básicas necessariamente distintas (uma vez que devido à estrita monotonicidade da função objectivo, em relação às sucessivas soluções básicas admissíveis que se vão determinando, o método não poderá produzir duas soluções básicas admissíveis iguais). Consequentemente, como o número de vértices é finito, ao fim de um número finito de quadros simplex ou se obtém uma solução ótima ou a conclusão de que o problema não tem ótimo finito.

Resta pois estudar os problemas levantados pelos casos degenerados, para os quais o método simplex pode entrar em ciclo, bem como o modo de os ultrapassar.

O primeiro exemplo de entrada em ciclo do método simplex foi publicado em (Hoffman, 1953) e pode ser consultado em (Dantzig, 1963). Trata-se de um programa linear com 11 variáveis e 3 equações. Mais tarde, em (Beale, 1955), surgiu o exemplo que a seguir se apresenta, com apenas 7 variáveis e 3 equações.

**Exemplo 3.6** (Beale, 1953). *Considerando o programa linear*

$$\begin{aligned} \min \quad & -\frac{3}{4}x_1 + 20x_2 - \frac{1}{2}x_3 + 6x_4 \\ \text{s. a} \quad & \frac{1}{4}x_1 - 8x_2 - x_3 + 9x_4 + x_5 = 0 \\ & \frac{1}{2}x_1 - 12x_2 - \frac{1}{2}x_3 + 3x_4 + x_6 = 0 \\ & x_3 + x_7 = 1 \\ & x_1, x_2, x_3, x_4, x_5, x_6, x_7 \geq 0, \end{aligned}$$

para o qual se obtém a solução ótima  $x^* = (1, 0, 1, 0, \frac{3}{4}, 0, 0)^T$  e cujo valor ótimo é igual a  $-\frac{5}{4}$ , vamos aplicar o método simplex revisto com o critério de Dantzig para escolha da coluna pivotal (com o qual sendo  $z_k = \max\{z_j = w^s a_j - c_j : 1 \leq j \leq n\} > 0$  se escolhe a  $k$ -ésima coluna para coluna pivotal), a partir do quadro associado à solução básica admissível inicial  $x^1 = (0, 0, 0, 0, 0, 0, 1)^T$  (cuja matriz básica é a matriz identidade).

**Solução.** Com o critério de Dantzig de escolha da coluna pivotal, a partir da solução básica referida, determina-se a seguinte sequência de quadros

$$x_{B_s} \quad \begin{array}{|c|} \hline B_s^{-1} \\ \hline w^s \\ \hline \end{array} \quad \begin{array}{|c|} \hline B_s^{-1}b \\ \hline w^s b \\ \hline \end{array}$$

1.

$x_5$	1	0	0	0
$x_6$	0	1	0	0
$x_7$	0	0	1	1
	0	0	0	0
	$\underbrace{\hspace{3cm}}$			
	$w^1$			

	$x_1$
	$\boxed{1/4}$
	1/2
	0
	3/4

onde  $\frac{3}{4} = \max\{z_1 - c_1 = \frac{3}{4}, z_2 - c_2 = -20, z_3 - c_3 = \frac{1}{2}, z_4 - c_4 = -6\}$ .

2.

$x_1$	4	0	0	0
$x_6$	-2	1	0	0
$x_7$	0	0	1	1
	-3	0	0	0

onde  $2 = \max\{z_1 - c_1 = \frac{7}{4}, z_2 - c_2 = -44, z_3 - c_3 = \frac{-1}{2}, z_6 - c_6 = 2\}$ .

7.

$x_5$	1	0	0	0
$x_6$	0	1	0	0
$x_7$	0	0	1	1
	0	0	0	0

Uma vez que o último dos quadros apresentados é igual ao primeiro, o método entra em ciclo sem nunca atingir a solução ótima. ■

Para evitar este tipo de problemas propuseram-se algumas regras, das quais vamos apresentar a mais simples de implementar, a regra de Bland (que condiciona a escolha dos índices da coluna e da linha pivotal).

### Regra de Bland

A regra de Bland parte de uma ordem inicialmente estabelecida para as variáveis e adopta como critério de escolha da variável a entrar e a sair da base o menor índice de entre os das respectivas candidatas.

**Exemplo 3.7.** Vamos aplicar a regra de Bland na resolução do exemplo de Beale, a partir do quadro simplex revisto

$x_5$	1	0	0	0
$x_6$	0	1	0	0
$x_7$	0	0	1	1
	0	0	0	0

**Solução.** Segue-se a sequência de quadros que se obtêm quando a escolha da coluna e da linha pivotal é feita de acordo com a regra de Bland.

1.

$x_5$	1	0	0	0
$x_6$	0	1	0	0
$x_7$	0	0	1	1
	0	0	0	0

$w^1$

$x_1$
$\frac{1}{4}$
$\frac{1}{2}$
0
$\frac{3}{4}$

2.

$x_1$	4	0	0	0
$x_6$	-2	1	0	0
$x_7$	0	0	1	1
	-3	0	0	0

$w^2$

$x_2$
-32
$\frac{4}{1}$
0
4

3.

$x_1$	-12	8	0	0
$x_2$	-1/2	1/4	0	0
$x_7$	0	0	1	1
		$\underbrace{-1 \quad -1 \quad 0}_{w^3}$		0

	$x_3$
	$\frac{8}{3/8}$
	1
	2

4.

$x_3$	-3/2	1	0	0
$x_2$	1/16	-1/8	0	0
$x_7$	3/2	-1	1	1
		$\underbrace{2 \quad -3 \quad 0}_{w^4}$		0

	$x_4$
	$\frac{-21/28}{3/16}$
	$\frac{21/2}{3}$

5.

$x_3$	2	-6	0	0
$x_4$	1/3	-2/3	0	0
$x_7$	-2	6	1	1
		$\underbrace{1 \quad -1 \quad 0}_{w^5}$		0

	$x_1$
	$\frac{-5/2}{-1/4}$
	$\frac{5/2}{1}$

6.

$x_3$	0	0	1	1
$x_4$	2/15	-1/15	1/10	1/10
$x_1$	-4/5	12/5	2/5	2/5
		$\underbrace{7/5 \quad 11/5 \quad -1/5}_{w^6}$		-1/5

	$x_6$
	$\frac{0}{2/15}$
	$\frac{-4/5}{7/5}$

7.

$x_3$	0	0	1	1
$x_5$	1	-1/2	3/4	3/4
$x_1$	0	2	1	1
	0	-3/2	-5/4	-5/4

Note-se que ao longo dos sucessivos quadros, a escolha da coluna pivotal caiu sobre a primeira coluna com coeficiente na linha de custos reduzidos positivo e a escolha da linha pivotal caiu sobre a de menor índice de entre as candidatas. Note-se ainda que o último quadro obtido é o quadro óptimo. ■

Da regra de Bland decorre uma propriedade de monotonicidade que implica a ausência de bases repetidas. Informalmente, denotando-se o conjunto de índices básicos e não básicos associados à  $s$ -ésima base, respectivamente, por  $J_{B_s}$  e  $J_{N_s}$ , esta propriedade consiste na impossibilidade de um índice  $q \in J_{N_s} \cap J_{B_{s+1}}$  (o índice que entra para a base  $B_{s+1}$ ) sair da base sem que um índice  $p \in J_{N_s}$ , tal que  $p > q$ , seja básico. Mais formalmente, sendo  $q$  um índice tal que

$$q \in J_{N_s} \cap J_{B_{s+1}} \cap \dots \cap J_{B_{s+r}} \cap J_{N_{s+r+1}},$$

i.e.,  $q$  entra na  $(s + 1)$ -ésima e sai da  $(s + r)$ -ésima base, então

$$q = \max J_{N_s} \cap J_{B_{s+1}} \leq \max J_{N_s} \cap J_{B_{s+2}} \leq \dots \leq \max J_{N_s} \cap J_{B_{s+r+1}} = p,$$

com  $p > q$ . Deve observar-se que  $J_{N_s} \cap J_{B_{s+1}}$  é um conjunto singular.

No Exemplo 3.8, pode evidenciar-se esta propriedade de monotonicidade, considerando a sequência dos pares de subconjuntos de índices, respectivamente, básicos e não básicos:

$$\begin{aligned} (J_{B_1}, J_{N_1}) &= (\{5, 6, 7\}, \{1, 2, 3, 4\}) & (J_{B_2}, J_{N_2}) &= (\{1, 6, 7\}, \{2, 3, 4, 5\}) \\ (J_{B_3}, J_{N_3}) &= (\{1, 2, 7\}, \{3, 4, 5, 6\}) & (J_{B_4}, J_{N_4}) &= (\{2, 3, 7\}, \{1, 4, 5, 6\}) \\ (J_{B_5}, J_{N_5}) &= (\{3, 4, 7\}, \{1, 2, 5, 6\}) & (J_{B_6}, J_{N_6}) &= (\{1, 3, 4\}, \{2, 5, 6, 7\}) \\ (J_{B_7}, J_{N_7}) &= (\{1, 3, 5\}, \{2, 4, 6, 7\}) \end{aligned}$$

e comparando cada índice  $q \in J_{N_s} \cap J_{B_{s+1}} \cap \dots \cap J_{B_{s+r}} \cap J_{N_{s+r+1}}$  com o maior índice de cada uma das intersecções  $J_{N_s} \cap J_{B_{s+1}}, \dots, J_{N_s} \cap J_{B_{s+r}}$  e  $J_{N_s} \cap J_{B_{s+r+1}}$ , conforme a seguir se explicita.

$$\begin{aligned} J_{N_1} \cap J_{B_2} \cap J_{B_3} \cap J_{N_4} &\ni 1 \leq 2 = \max J_{N_1} \cap J_{B_3} \leq 3 = \max J_{N_1} \cap J_{B_4}; \\ J_{N_2} \cap J_{B_3} \cap J_{B_4} \cap J_{N_5} &\ni 2 \leq 3 = \max J_{N_2} \cap J_{B_4} \leq 4 = \max J_{N_2} \cap J_{B_5}; \\ J_{N_4} \cap J_{B_5} \cap J_{B_6} \cap J_{N_7} &\ni 4 \leq 4 = \max J_{N_4} \cap J_{B_6} \leq 5 = \max J_{N_4} \cap J_{B_7}. \end{aligned}$$

**Teorema 3.2.** *Em situações degeneradas, o método simplex, com a aplicação da regra de Bland, não produz bases repetidas.*

**Demonstração.** Seja  $B_s$  uma base degenerada e suponha que, aplicando a regra de Bland,  $\exists r \geq 2$  tal que  $B_s = B_{s+r}$ <sup>6</sup>. Então o método simplex entra em ciclo e todas as bases, entre  $B_s$  e  $B_{s+r}$ , se repetem. Com esta repetição de bases, seja  $q$  o maior índice a entrar para base e suponha, sem perda de generalidade, que tal se verifica para  $B_{s+1}$ , i.e.,

$$\{q\} = J_{N_s} \cap J_{B_{s+1}}.$$

Uma vez que as bases se repetem e, em particular,  $J_{B_s} = J_{B_{s+r}}$ , podemos concluir que existe  $j < r$  tal que

$$q \in J_{N_s} \cap J_{B_{s+1}} \cap \dots \cap J_{B_{s+j}} \cap J_{N_{s+j+1}}.$$

Consequentemente, de acordo com a propriedade de monotonicidade da regra de Bland,  $\exists p > q$  tal que

$$q = \max J_{N_s} \cap J_{B_{s+1}} \leq \dots \leq \max J_{N_s} \cap J_{B_{s+j}} \leq \max J_{N_s} \cap J_{B_{s+j+1}} = p,$$

o que contraria a hipótese de  $q$  ser o maior índice a entrar para a base.  $\square$

Relativamente à existência de programas lineares de pequena dimensão susceptíveis de provocarem a entrada em ciclo quando resolvidos pelo métodos simplex, em (Marshall and Suurballe, 1969) provou-se que o  $PL(A, b, c)$  de menor dimensão tem  $m = 2$  e  $n = 6$  e apresenta um ciclo de comprimento 6.

<sup>6</sup>Deve observar-se que  $\forall s, B_s \neq B_{s+1}$ .

### 3.4 Variantes do método simplex

Em certas situações, a elevada dimensão do problema a resolver ou mesmo a sua estrutura particular, justificam que se introduzam algumas alterações tendo em vista aumentar a eficiência da aplicação do método simplex. Essas alterações dão origem a diferentes versões, das quais vamos analisar as que se designam, respectivamente, por método simplex para variáveis limitadas e o método de decomposição de Dantzig-Wolfe. Mais adiante, após a introdução da dualidade em programação linear, ainda abordaremos os métodos dual simplex e primal-dual.

#### 3.4.1 Método simplex para variáveis limitadas

Dado um programa linear de minimização na forma padrão,  $PL(A, b, c)$ , no qual as condições de não negatividade das variáveis de decisão são substituídas pelas **restrições de caixa**,  $l \leq x \leq u$ , ou seja, dado o programa linear:

$$\begin{aligned} \min \quad & c^T x \\ \text{s.a.} \quad & Ax = b \text{ com } A \in \mathbb{R}^{m \times n}, \\ & l \leq x \leq u, \end{aligned}$$

onde  $l$  e  $u$  denotam  $n$ -uplos, respectivamente, de limites inferiores e de limites superiores para o  $n$ -uplo de variáveis,  $x$ , vamos designá-lo por programa linear de minimização na forma padrão,  $PL(A, b, c)$ , com variáveis limitadas (ou restrições de caixa),  $l \leq x \leq u$ . A partir deste programa, fazendo  $y = x - l \Leftrightarrow x = y + l$  e procedendo à respectiva mudança de variáveis obtém-se

$$\begin{aligned} \min \quad & c^T y + c^T l \\ \text{s.a.} \quad & Ay = b - Al, \\ & 0 \leq y \leq u - l. \end{aligned}$$

Introduzindo variáveis de desvio, nas restrições relativas aos limites superiores, o programa toma a forma:

$$\begin{aligned} \min \quad & c^T y + c^T l \\ \text{s.a.} \quad & \begin{bmatrix} A & 0 \\ I_n & I_n \end{bmatrix} \begin{bmatrix} y \\ y_s \end{bmatrix} = \begin{bmatrix} b - Al \\ u - l \end{bmatrix}, \\ & y, y_s \geq 0. \end{aligned}$$

Com esta transformação, no entanto, aumenta-se o número de restrições de  $m$  para  $m + n$  e o número de variáveis de  $n$  para  $2n$ , o que, para valores elevados de  $n$ , não é muito recomendável. Por este motivo, desenvolveu-se uma variante do método simplex para problemas com variáveis limitadas que não recorre a esta mudança de variáveis.

Dado o sistema de equações e inequações  $\begin{cases} Ax = b \\ l \leq x \leq u \end{cases}$ , uma solução  $x'$  de  $Ax = b$  diz-se básica, se a matriz  $A$  pode ser partida em  $[B, N_l, N_u]$ , onde  $B$  é uma matriz quadrada invertível ( $r(B) = m$ ) tal que, sendo

$$x' = \begin{bmatrix} x'_B \\ x'_{N_l} \\ x'_{N_u} \end{bmatrix},$$

vem  $x'_{N_l} = l_{N_l}$ ,  $x'_{N_u} = u_{N_u}$  e  $x'_B = B^{-1}b - B^{-1}N_l l_{N_l} - B^{-1}N_u u_{N_u}$ .

Adicionalmente, se  $l_B \leq x'_B \leq u_B$ , então  $x'$  diz-se uma **solução básica admissível**. Nestas condições, a matriz  $B$  designa-se por **base de trabalho**,  $x_B$  denota o  $m$ -uplo de variáveis básicas e  $x_{N_l}$  e  $x_{N_u}$  denotam às variáveis não básicas, respectivamente, nos seus limites inferiores e superiores. Sendo  $x'$  uma solução básica admissível, se  $l_B < x'_B < u_B$ , então  $x'$  diz-se **não degenerada**, caso contrário diz-se **degenerada**.

Uma partição de  $A$ ,  $[B, N_l, N_u]$ , que corresponda a uma solução básica admissível, diz-se uma **partição básica admissível**. Denotando por  $J_B$ ,  $J_{N_l}$  e  $J_{N_u}$  os subconjuntos de índices, respectivamente, básicos, não básicos cujas variáveis atingiram os limites inferiores e não básicos cujas variáveis atingiram os limites superiores. Embora  $J_B$  fique completamente determinado a partir do conhecimento de  $J_{N_l}$  e  $J_{N_u}$ , vamos denotar uma partição básica admissível pelo respectivo terno de subconjuntos  $(J_B, J_{N_l}, J_{N_u})$ .

Duas partições básicas admissíveis,  $(J'_B, J'_{N_l}, J'_{N_u})$  e  $(J''_B, J''_{N_l}, J''_{N_u})$ , dizem-se adjacentes se, passando de uma para outra todas as variáveis não básicas, com excepção de uma, se mantêm não básicas e nos respectivos limites superiores e inferiores. Mais formalmente, as partições básicas admissíveis  $(J'_B, J'_{N_l}, J'_{N_u})$  e  $(J''_B, J''_{N_l}, J''_{N_u})$  dizem-se adjacentes se  $|J'_{N_l} \cap J''_{N_l}| + |J'_{N_u} \cap J''_{N_u}| = |J'_{N_l}| + |J'_{N_u}| - 1$ .

Seja  $A = [B, N_l, N_u]$ ,  $x = \begin{bmatrix} x_B \\ x_{N_l} \\ x_{N_u} \end{bmatrix}$ ,  $c^T = [c_B, c_{N_l}, c_{N_u}]$ ,  $\hat{b}$  o  $m$ -uplo de componen-

tes básicas  $x_B$  e  $\zeta$  o valor corrente da função objectivo  $c^T x$ . Então vem que

$$\begin{aligned} \hat{b} &= B^{-1}b - B^{-1}Nx_N \\ &= B^{-1}b - B^{-1}N_l x_{N_l} - B^{-1}N_u x_{N_u}, \\ \zeta &= c_B(B^{-1}b - B^{-1}Nx_N) + c_N x_N \\ &= c_B(B^{-1}b - B^{-1}N_l x_{N_l} - B^{-1}N_u x_{N_u}) + c_{N_l} x_{N_l} + c_{N_u} x_{N_u} \\ &= c_B B^{-1}b - (c_B B^{-1}N_l - c_{N_l})x_{N_l} - (c_B B^{-1}N_u - c_{N_u})x_{N_u}, \end{aligned}$$

onde  $N = [N_l, N_u]$  e  $x_N = \begin{bmatrix} x_{N_l} \\ x_{N_u} \end{bmatrix}$ . Nestas condições, o quadro simplex toma o aspecto:

	$x_{N_l}$	$x_{N_u}$	
$x_B$	$B^{-1}N_l$	$B^{-1}N_u$	$\hat{b}$
	$c_B B^{-1}N_l - c_{N_l}$	$c_B B^{-1}N_u - c_{N_u}$	$\zeta$

- Tendo em conta que  $\zeta = c_B B^{-1}b - \sum_{j \in J_{N_l}} (z_j - c_j)x_j - \sum_{j \in J_{N_u}} (z_j - c_j)x_j$ , podemos concluir o seguinte:

- Se  $\exists j \in J_{N_l}$  tal que  $z_j - c_j > 0$ , então um aumento de  $x_j$  ( $x_j = l_j + \Delta_j$ , com  $\Delta_j > 0$ ) beneficia a função objectivo.
- Se  $\exists j \in J_{N_u}$  tal que  $z_j - c_j < 0$ , então a diminuição de  $x_j$  ( $x_j = u_j - \Delta_j$ , com  $\Delta_j > 0$ ) beneficia a função objectivo.

Assim, a escolha da  $k$ -ésima coluna para coluna pivotada pode fazer-se, por aplicação do critério de Dantzig, a partir da determinação de

$$\delta_k = \max\left\{\max_{j \in J_{N_l}} (z_j - c_j), \max_{j \in J_{N_u}} -(z_j - c_j)\right\}.$$

Se  $\delta_k \leq 0$ , então conclui-se que a solução básica corrente é ótima (a prova desta afirmação é proposta no Exercício 3.12, no final deste capítulo). Caso contrário,  $k$  é o índice de uma variável não básica candidata a entrar para a base.

- Suponha que a solução básica corrente é a determinada pela partição básica admissível  $(J_B, J_{N_l}, J_{N_u})$  e que a  $k$ -ésima coluna é candidata a pivotada. Então ocorre uma das situações:

1. A variável candidata a entrar para a base está no seu limite inferior ( $x_k = l_k$ ).
2. A variável candidata a entrar para a base está no seu limite superior ( $x_k = u_k$ ).

Vamos analisar cada um destes casos, considerando  $J_N = J_{N_l} \cup J_{N_u}$ .

1. Supondo que o índice  $k$  é tal que  $x_k = l_k$ , pelo que  $z_k - c_k > 0$ , fazendo  $x_k = l_k + \Delta_k$ , com  $\Delta_k > 0$ , vem que

$$\begin{aligned} \zeta &\leftarrow c_B B^{-1} b - \sum_{j \in J_N \setminus \{k\}} (z_j - c_j) x_j - (z_k - c_k)(l_k + \Delta_k) \\ &= c_B B^{-1} b - \sum_{j \in N_l} (z_j - c_j) l_j - \sum_{j \in N_u} (z_j - c_j) u_j - (z_k - c_k) \Delta_k \\ &= \zeta - (z_k - c_k) \Delta_k. \end{aligned}$$

As restantes variáveis não básicas (com índice distinto de  $k$ ) permanecem com os mesmos valores e as básicas são actualizadas conforme se indica:

$$\begin{aligned} x_B &\leftarrow B^{-1} b - \sum_{j \in J_N \setminus \{k\}} x_j B^{-1} a_j - (l_k + \Delta_k) B^{-1} a_k \\ &= B^{-1} b - \sum_{j \in J_{N_l}} l_j B^{-1} a_j - \sum_{j \in J_{N_u}} u_j B^{-1} a_j - \Delta_k B^{-1} a_k \\ &= \hat{b} - \Delta_k y_k, \end{aligned}$$

onde  $y_k = B^{-1} a_k$  e  $\hat{b} = B^{-1} b - \sum_{j \in J_{N_l}} l_j B^{-1} a_j - \sum_{j \in J_{N_u}} u_j B^{-1} a_j$ . De modo a garantir a admissibilidade da nova solução, i.e., de modo a garantir as desigualdades  $l_k \leq l_k + \Delta_k \leq u_k$  e  $l_B \leq \hat{b} - \Delta_k y_k \leq u_B$  (onde  $l_B$  e  $u_B$  denotam os  $m$ -uplos dos limites, respectivamente, inferiores e superiores associados às variáveis básicas correntes), a escolha de  $\Delta_k$  deve satisfazer as desigualdades:

$$0 \leq \Delta_k \leq u_k - l_k,$$

e  $\forall i \in \{1, \dots, m\}$ ,

$$(l_B)_i \leq \hat{b}_i - \Delta_k y_{ik} \leq (u_B)_i \Leftrightarrow \begin{cases} \Delta_k \leq \frac{\hat{b}_i - (l_B)_i}{y_{ik}} & \text{se } y_{ik} > 0 \\ \Delta_k \leq \frac{(u_B)_i - \hat{b}_i}{-y_{ik}} & \text{se } y_{ik} < 0 \end{cases},$$

onde  $(l_B)_i$  e  $(u_B)_i$  denotam a  $i$ -ésima componente do  $m$ -uplo de limites, respectivamente, inferiores e superiores das variáveis básicas. Assim, o valor mais favorável para  $\Delta_k$  vem dada pelo ínfimo do conjunto

$$\left\{ \frac{\hat{b}_i - (l_B)_i}{y_{ik}} : 1 \leq i \leq m, y_{ik} > 0 \right\} \cup \left\{ \frac{(u_B)_i - \hat{b}_i}{-y_{ik}} : 1 \leq i \leq m, y_{ik} < 0 \right\} \cup \{u_k - l_k\},$$

o qual, no caso de não existirem limites superiores para determinadas variáveis, pode tomar valores arbitrariamente grandes e, em tal situação, diz-se que  $\Delta_k = \infty$ .

2. Supondo que o índice  $k$  é tal que  $x_k = u_k$ , pelo que  $z_k - c_k < 0$ , fazendo  $x_k = u_k - \Delta_k$ , com  $\Delta_k > 0$ , vem que

$$\begin{aligned} \zeta &\leftarrow c_B B^{-1} b - \sum_{j \in J_N \setminus \{k\}} (z_j - c_j) x_j - (z_k - c_k)(u_k - \Delta_k) \\ &= c_B B^{-1} b - \sum_{j \in N_l} (z_j - c_j) l_j - \sum_{j \in N_u} (z_j - c_j) u_j + (z_k - c_k) \Delta_k \\ &= \zeta + (z_k - c_k) \Delta_k. \end{aligned}$$

As restantes variáveis não básicas (com índice distinto de  $k$ ) permanecem com os mesmos valores e as básicas são actualizadas conforme se indica:

$$\begin{aligned} x_B &\leftarrow B^{-1} b - \sum_{j \in J_N \setminus \{k\}} x_j B^{-1} a_j - (u_k - \Delta_k) B^{-1} a_k \\ &= B^{-1} b - \sum_{j \in J_{N_l}} l_j B^{-1} a_j - \sum_{j \in J_{N_u}} u_j B^{-1} a_j + \Delta_k B^{-1} a_k \\ &= \hat{b} + \Delta_k y_k, \end{aligned}$$

onde  $y_k = B^{-1} a_k$  e  $\hat{b} = B^{-1} b - \sum_{j \in J_{N_l}} l_j B^{-1} a_j - \sum_{j \in J_{N_u}} u_j B^{-1} a_j$ . De modo a garantir a admissibilidade da nova solução, i.e., de modo a garantir as desigualdades  $l_k \leq u_k - \Delta_k \leq u_k$  e  $l_B \leq \hat{b} + \Delta_k y_k \leq u_B$ , a escolha de  $\Delta_k$  deve satisfazer as desigualdades:

$$0 \leq \Delta_k \leq u_k - l_k,$$

e  $\forall i \in \{1, \dots, m\}$ ,

$$(l_B)_i \leq \hat{b}_i + \Delta_k y_{ik} \leq (u_B)_i \Leftrightarrow \begin{cases} \Delta_k \leq \frac{\hat{b}_i - (l_B)_i}{-y_{ik}} & \text{se } y_{ik} < 0 \\ \Delta_k \leq \frac{(u_B)_i - \hat{b}_i}{y_{ik}} & \text{se } y_{ik} > 0 \end{cases}.$$

Logo, o valor mais favorável para  $\Delta_k$  vem dado pelo ínfimo do conjunto

$$\left\{ \frac{\hat{b}_i - (l_B)_i}{-y_{ik}} : 1 \leq i \leq m, y_{ik} < 0 \right\} \cup \left\{ \frac{(u_B)_i - \hat{b}_i}{y_{ik}} : 1 \leq i \leq m, y_{ik} > 0 \right\} \cup \{u_k - l_k\},$$

o qual, no caso de não existirem limites superiores para determinadas variáveis, pode tomar valores arbitrariamente grandes e, em tal situação, diz-se que  $\Delta_k = \infty$ .

- Mais geralmente (para 1 e 2), considerando a aplicação

$$\begin{aligned} \sigma : J_N &\mapsto \{1, -1\} \\ k &\rightsquigarrow \sigma(k) = \begin{cases} 1 & \text{se } k \in J_{N_l} \\ -1 & \text{se } k \in J_{N_u} \end{cases}, \end{aligned}$$

e denotando por  $\Omega(J_{N_l}, J_{N_u}, y_k)$  o conjunto

$$\left\{ \frac{\hat{b}_i - (l_B)_i}{\sigma(k)y_{ik}} : 1 \leq i \leq m, \sigma(k)y_{ik} > 0 \right\} \cup \left\{ \frac{(u_B)_i - \hat{b}_i}{-\sigma(k)y_{ik}} : 1 \leq i \leq m, \sigma(k)y_{ik} < 0 \right\} \cup \{u_k - l_k\},$$

vem

$$\Delta_k = \inf \Omega(J_{N_l}, J_{N_u}, y_k).$$

- Se  $\Delta_k = \infty$ , então o valor óptimo não é finito.
- Se  $\Delta_k = u_k - l_k < \infty$ , então a variável  $x_k$  mantém-se não básica, mudando apenas de um dos subconjuntos não básicos para o outro. Como consequência, no quadro simplex, tanto  $B^{-1}N$  como  $c_B B^{-1}N - c_N$  permanecem inalteráveis. Porém,  $x_k$ , o  $m$ -uplo de variáveis básicas  $x_B$  e  $\zeta$ , devem ser actualizados conforme se indica:

$$x_k \leftarrow x_k + \sigma(k)\Delta_k \quad (3.4)$$

$$x_B \leftarrow x_B - \sigma(k)\Delta_k y_k \quad (3.5)$$

$$\zeta \leftarrow \zeta - \sigma(k)\Delta_k(z_k - c_k) \quad (3.6)$$

- Se  $\Delta_k = \frac{\hat{b}_q - (l_B)_q}{\sigma(k)y_{qk}}$  ou  $\Delta_k = \frac{(u_B)_q - \hat{b}_q}{-\sigma(k)y_{qk}}$ , então o quadro simplex deve ser actualizado através das operações de pivotação, considerando  $y_{qk}$  como "pivô". No entanto, a actualização de  $x_k$ , a actualização das componentes básicas da solução básica corrente e a actualização do valor da função objectivo, deve continuar a fazer-se de acordo com (3.4), (3.5) e (3.6).
- A determinação da solução básica admissível inicial pode fazer-se recorrendo ao método das duas fases ou ao método do "big-M", atribuindo às variáveis originais (por exemplo) os respectivos limites inferiores e introduzindo variáveis artificiais com os valores obtidos para os termos independentes, depois de multiplicar por -1 as equações em que os termos independentes aparecem com valor negativo.

**Exemplo 3.8.** Vamos resolver o programa linear a seguir indicado, por aplicação do método simplex para problemas com variáveis limitadas.

$$\begin{array}{rllll} \min & -2x_1 & - & 4x_2 & - & x_3 \\ \text{s. a} & 2x_1 & + & x_2 & + & x_3 & \leq & 10, \\ & x_1 & + & x_2 & - & x_3 & \leq & 4, \\ 0 \leq & x_1 & & & & & \leq & 4, \\ 0 \leq & & & x_2 & & & \leq & 6, \\ 1 \leq & & & & & x_3 & \leq & 4. \end{array}$$

**Solução.** Introduzindo as variáveis de desvio,  $x_4$  e  $x_5$  (cujos limites inferior e superior são 0 e  $\infty$ , respectivamente), considerando estas variáveis como variáveis básicas e fazendo  $x_1 = x_2 = 0$  e  $x_3 = 1$ , obtém-se a solução básica admissível inicial,  $x_4 = 10 - 2x_1 - x_2 - x_3 = 9$  e  $x_5 = 4 - x_1 - x_2 + x_3 = 5$  e o quadro simplex

	1	1	1	
	$x_1$	$x_2$	$x_3$	
$x_4$	2	1	1	9
$x_5$	1	1	-1	5
	2	4	1	-1

no qual, as etiquetas  $l$  denotam que as correspondentes variáveis não básicas têm como valores correntes os limites inferiores (mais adiante, as variáveis não básicas com valores iguais aos limites superiores serão etiquetadas com  $u$ ).

$$1. \max_{j \in N_l} (z_j - c_j) = \max\{2, 4, 1\} = 4 = z_2 - c_2,$$

	1	1	1		
	$x_1$	$x_2$	$x_3$		$x_2$
$x_4$	2	1	1	9	1
$x_5$	1	1	-1	5	1
	2	4	1	-1	4

$$\begin{aligned} \Omega(J_{N_l}, J_{N_u}, y_2) &= \Omega(\{1, 2, 3\}, \emptyset, y_2) \\ &= \left\{ \frac{\hat{b}_1 - l_4}{\sigma(2)y_{12}}, \frac{\hat{b}_2 - l_5}{\sigma(2)y_{22}} \right\} \cup \emptyset \cup \{u_2 - l_2\} \\ &= \left\{ \frac{9-0}{1}, \frac{5-0}{1}, 6-0 \right\}. \end{aligned}$$

Uma vez que  $\Delta_2 = \inf \Omega(J_{N_l}, J_{N_u}, y_2) = \inf\{9, 5, 6\} = 5$ , a segunda linha do quadro é a pivotal e obtêm-se as actualizações:

$$\begin{aligned} x_2 &\leftarrow l_2 + \Delta_2 = 0 + 5 = 5; \\ \begin{bmatrix} x_4 \\ x_5 \end{bmatrix} &\leftarrow \hat{b} - \Delta_2 y_2 = \begin{bmatrix} 9 \\ 5 \end{bmatrix} - 5 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \end{bmatrix}; \\ \zeta &\leftarrow \zeta - \Delta_2 (z_2 - c_2) = -1 - 5 \times 4 = -21; \end{aligned}$$

	1	1	1	
	$x_1$	$x_5$	$x_3$	
$x_4$	1	-1	2	4
$x_2$	1	1	-1	5
	-2	-4	5	-21

$$2. \max_{j \in N_l} (z_j - c_j) = \max\{-2, -4, 5\} = 5 = z_3 - c_3,$$

	1	1	1		
	$x_1$	$x_5$	$x_3$		$x_3$
$x_4$	1	-1	2	4	2
$x_2$	1	1	-1	5	-1
	-2	-4	5	-21	5

$$\begin{aligned} \Omega(J_{N_l}, J_{N_u}, y_3) &= \Omega(\{1, 5, 3\}, \emptyset, y_3) \\ &= \left\{ \frac{\hat{b}_1 - l_4}{\sigma(3)y_{13}} \right\} \cup \left\{ \frac{u_2 - \hat{b}_2}{-\sigma(3)y_{23}} \right\} \cup \{u_3 - l_3\} \\ &= \left\{ \frac{4-0}{2}, \frac{6-5}{1}, 4-1 \right\}. \end{aligned}$$

Uma vez que  $\Delta_3 = \inf \Omega(J_{N_l}, J_{N_u}, y_3) = \inf\{2, 1, 3\} = 1$ , a segunda linha do quadro é a pivotal e obtêm-se as actualizações:

$$\begin{aligned} x_3 &\leftarrow l_3 + \Delta_3 = 1 + 1 = 2; \\ \begin{bmatrix} x_4 \\ x_2 \end{bmatrix} &\leftarrow \hat{b} - \Delta_3 y_3 = \begin{bmatrix} 4 \\ 5 \end{bmatrix} - \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ 6 \end{bmatrix}; \\ \zeta &\leftarrow \zeta - \Delta_3 (z_3 - c_3) = -21 - 1 \times 5 = -26; \end{aligned}$$

	1	1	u	
	$x_1$	$x_5$	$x_2$	
$x_4$	3	1	2	2
$x_3$	-1	-1	-1	2
	3	1	5	-26

$$3. \max \{ \max_{j \in N_l} (z_j - c_j), \max_{j \in N_u} -(z_j - c_j) \} = \max\{3, -5\} = 3 = z_1 - c_1,$$

	1	1	u		
	$x_1$	$x_5$	$x_2$		$x_1$
$x_4$	3	1	2	2	3
$x_3$	-1	-1	-1	2	-1
	3	1	5	-26	3

$$\begin{aligned} \Omega(J_{N_l}, J_{N_u}, y_1) &= \Omega(\{1, 5\}, \{2\}, y_1) \\ &= \left\{ \frac{\hat{b}_1 - l_4}{\sigma(1)y_{11}} \right\} \cup \left\{ \frac{u_3 - \hat{b}_2}{-\sigma(1)y_{21}} \right\} \cup \{u_1 - l_1\} \\ &= \left\{ \frac{2-0}{3}, \frac{4-2}{1}, 4-0 \right\}. \end{aligned}$$

Uma vez que  $\Delta_1 = \inf \Omega(J_{N_l}, J_{N_u}, y_1) = \inf \left\{ \frac{2}{3}, 2, 4 \right\} = \frac{2}{3}$ , a primeira linha do quadro é a pivotal e obtêm-se as actualizações:

$$\begin{aligned} x_1 &\leftarrow l_1 + \Delta_1 = 0 + \frac{2}{3} = \frac{2}{3}; \\ \begin{bmatrix} x_4 \\ x_3 \end{bmatrix} &\leftarrow \hat{b} - \Delta_1 y_1 = \begin{bmatrix} 2 \\ 2 \end{bmatrix} - \frac{2}{3} \begin{bmatrix} 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{8}{3} \end{bmatrix}; \\ \zeta &\leftarrow \zeta - \Delta_1 (z_1 - c_1) = -26 - \frac{2}{3} \times 3 = -28; \end{aligned}$$

	1	1	u	
	$x_4$	$x_5$	$x_2$	
$x_1$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{2}{3}$
$x_3$	$\frac{1}{3}$	$-\frac{2}{3}$	$-\frac{1}{3}$	$\frac{8}{3}$
	-1	0	3	-28

$$4. \max \{ \max_{j \in N_l} (z_j - c_j), \max_{j \in N_u} -(z_j - c_j) \} = \max\{0, -3\} = 0.$$

Nestas condições, a solução básica admissível corrente é óptima. Porém, uma vez que  $z_5 - c_5 = 0$ , existe uma solução óptima alternativa que pode ser determinada escolhendo  $y_5$  para coluna pivotal. Tendo em vista a sua determinação, procedendo como anteriormente, vem

	1	1	u		
	$x_4$	$x_5$	$x_2$		$x_5$
$x_1$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{1}{3}$
$x_3$	$\frac{1}{3}$	$-\frac{2}{3}$	$-\frac{1}{3}$	$\frac{8}{3}$	$-\frac{2}{3}$
	-1	0	3	-28	0

$$\begin{aligned} \Omega(J_{N_l}, J_{N_u}, y_5) &= \Omega(\{4, 5\}, \{2\}, y_5) \\ &= \left\{ \frac{\hat{b}_1 - l_1}{y_{15}} \right\} \cup \left\{ \frac{u_3 - \hat{b}_2}{-y_{25}} \right\} \cup \{u_5 - l_5\} \\ &= \left\{ \frac{\frac{2}{3} - 0}{\frac{1}{3}}, \frac{4 - \frac{8}{3}}{-\frac{2}{3}}, \infty - 0 \right\}. \end{aligned}$$

Uma vez que  $\Delta_5 = \inf \Omega(J_{N_l}, J_{N_u}, y_5) = \inf \{2, 2, \infty\} = 2$  e este ínfimo não é único,<sup>7</sup> escolhe-se (por exemplo) a primeira linha para pivotal e obtêm-se as actualizações:

$$\begin{aligned} x_5 &\leftarrow l_5 + \Delta_5 = 0 + 2 = 2; \\ \begin{bmatrix} x_1 \\ x_3 \end{bmatrix} &\leftarrow \hat{b} - \Delta_5 y_5 = \begin{bmatrix} \frac{2}{3} \\ \frac{2}{3} \end{bmatrix} - 2 \begin{bmatrix} \frac{1}{3} \\ -\frac{2}{3} \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \end{bmatrix}; \\ \zeta &\leftarrow \zeta - \Delta_5(z_5 - c_5) = -28 - 2 \times 0 = -28. \end{aligned}$$

■

### 3.4.2 Método de decomposição de Dantzig-Wolfe

Para ilustrar o método de decomposição de Dantzig-Wolfe, ao longo desta secção, vamos considerar o programa linear  $PL(A, b, c)$  de minimização:

$$\begin{aligned} \min \quad & c^T x \\ \text{s.a.} \quad & Ax = b, \\ & x \in X, \end{aligned}$$

onde  $X$  é um conjunto poliédrico definido à custa de restrições com determinada estrutura. Pelo teorema da representação (Teorema 2.5), sendo  $x^1, \dots, x^p$  os pontos extremos de  $X$  e sendo  $d^1, \dots, d^q$  as respectivas direcções extremas,  $\forall x \in X, \exists \lambda_1, \dots, \lambda_p \geq 0$ , com  $\sum_{j=1}^p \lambda_j = 1$ , e  $\exists \mu_1, \dots, \mu_q \geq 0$  tais que

$$x = \sum_{j=1}^p \lambda_j x^j + \sum_{i=1}^q \mu_i d^i.$$

Nestas condições o programa linear  $PL(A, b, c)$  pode reescrever-se do seguinte modo:

$$\begin{aligned} \min \quad & \sum_{j=1}^p \lambda_j c^T x^j + \sum_{i=1}^q \mu_i c^T d^i \\ \text{s. a} \quad & \sum_{j=1}^p \lambda_j A x^j + \sum_{i=1}^q \mu_i A d^i = b, \\ & \sum_{j=1}^p \lambda_j = 1, \\ & \lambda_1, \dots, \lambda_p, \quad \mu_1, \dots, \mu_q \geq 0. \end{aligned}$$

Assim, fazendo

$$\begin{aligned} \bar{A} &= \begin{bmatrix} Ax^1 & \dots & Ax^p & Ad^1 & \dots & Ad^p \\ 1 & \dots & 1 & 0 & \dots & 0 \end{bmatrix}, \\ \bar{b} &= \begin{bmatrix} b \\ 1 \end{bmatrix}, \\ \bar{c}^T &= [c^T x^1, \dots, c^T x^p, c^T d^1, \dots, c^T d^q], \end{aligned}$$

obtém-se o programa linear de minimização na forma padrão  $PL(\bar{A}, \bar{b}, \bar{c})$ , o qual designamos por **problema principal** de modo a distingui-lo de um outro problema, a introduzir mais adiante, que se designará por **problema auxiliar**.

Aplicando a  $PL(\bar{A}, \bar{b}, \bar{c})$  o método simplex revisto, para uma dada solução básica admissível,  $\begin{bmatrix} \lambda_{\bar{B}} \\ \mu_{\bar{B}} \end{bmatrix}$ , à qual está associada a matriz básica  $\bar{B}$ , obtém-se o quadro

<sup>7</sup>Devido à situação de empate existente na escolha da linha pivotal, a solução básica obtida é degenerada (com  $x_1$  e  $x_3$  a atingirem os respectivos limites, inferior e superior).

$\lambda_{\bar{B}}$	$\bar{B}^{-1}$	$\bar{B}^{-1}\bar{b}$
$\mu_{\bar{B}}$	$[w, \alpha]$	$\zeta$

onde  $[w, \alpha] = \bar{c}_{\bar{B}}\bar{B}^{-1}$  e  $\zeta = \bar{c}_{\bar{B}}\bar{B}^{-1}\bar{b} = wb + \alpha$ .

A determinação dos coeficientes da linha de custos reduzidos vem dada por:

$$\begin{aligned}\bar{z}_j - \bar{c}_j &= [w, \alpha] \begin{bmatrix} Ax^j \\ 1 \end{bmatrix} - c^T x_j = (wA - c^T)x^j + \alpha, \text{ para } j = 1, \dots, p, \\ \bar{z}_{p+i} - \bar{c}_{p+i} &= [w, \alpha] \begin{bmatrix} Ad^i \\ 0 \end{bmatrix} - c^T d^i = (wA - c^T)d^i, \text{ para } i = 1, \dots, q.\end{aligned}$$

Porém, tendo em conta o facto de, usualmente, desconhecermos a generalidade, quer dos pontos extremos, quer das direcções extremas de  $X$  (quando existem), a determinação (ou verificação da não existência) de uma coluna de  $A$  associada a um coeficiente positivo na linha de custos reduzidos será feita através da resolução do seguinte **problema auxiliar**:

$$\max\{(wA - c^T)x : x \in X\}. \quad (3.7)$$

No caso deste problema auxiliar ser ilimitado, então tal significa que existe  $d^{k-p}$  (direcção extrema de  $X$ ) tal que  $(wA - c^T)d^{k-p} > 0$ . Logo, escolhendo  $\mu_{k-p}$  como variável candidata a entrar para a base, a coluna candidata a pivotal (associada à  $k$ -ésima coluna de  $\bar{A}$ ) e o correspondente coeficiente da linha de custos reduzidos vêm dados por

$$y_k = \bar{B}^{-1} \begin{bmatrix} Ad^{k-p} \\ 0 \end{bmatrix} \text{ e } \bar{z}_k - \bar{c}_k = (wA - c^T)d^{k-p},$$

respectivamente (se  $y_k \leq 0$ , então podemos concluir que o problema original não tem óptimo finito).

No caso do problema auxiliar produzir uma solução básica óptima  $x^k$ , verifica-se uma das situações:

1.  $(wA - c^T)x^k + \alpha > 0$ .
2.  $(wA - c^T)x^k + \alpha \leq 0$ .

Vamos analisar cada um destes casos.

1. Se  $(wA - c^T)x^k + \alpha > 0$ , então escolhendo  $\lambda_k$  como variável candidata a entrar para a base, a coluna candidata a pivotal e o correspondente coeficiente da linha de custos reduzidos vêm dados por

$$y_k = \bar{B}^{-1} \begin{bmatrix} Ax^k \\ 1 \end{bmatrix} \text{ e } \bar{z}_k - \bar{c}_k = (wA - c^T)x^k + \alpha,$$

respectivamente (mais uma vez, se  $y_k \leq 0$ , então podemos concluir que o problema original não tem óptimo finito).

2. Se  $(wA - c^T)x^k + \alpha \leq 0$ , então

$$(wA - c^T)x + \alpha \leq (wA - c^T)x^k + \alpha \leq 0 \quad \forall x \in X,$$

ou seja,  $\bar{z}_j - \bar{c}_j \leq 0 \quad \forall j \in \{1, \dots, p+q\}$ , donde podemos concluir que a solução

básica corrente é óptima. Consequentemente, em tais condições, se  $\lambda_B = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_r \end{bmatrix}$

e  $\mu_B = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_s \end{bmatrix}$ , então

$$x^* = \sum_{j=1}^r \lambda_j x^j + \sum_{i=1}^s \mu_i d^i$$

constitui uma solução óptima para o problema original.

A determinação de uma base admissível inicial, para o problema principal, pode fazer-se à custa do conhecimento de um ponto extremo de  $X$ ,  $x^1$  (de modo a considerar-se  $\lambda_1 = 1$  como componente básica) e da introdução de variáveis de desvio (no caso das restrições serem da forma  $\sum_{j=1}^p \lambda_j Ax^j + \sum_{i=1}^q \mu_i Ad^i \leq b$  com  $b \geq 0$ ) ou artificiais (no caso contrário). Com efeito, supondo (sem perda de generalidade) que  $b \geq 0$ , modificando o problema principal, conforme se indica:

$$\begin{array}{ll} \min & \sum_{j=1}^p \lambda_j c^T x^j + \sum_{i=1}^q \mu_i c^T d^i \\ \text{s. a} & \sum_{k=1}^m s_k e_k + \sum_{j=1}^p \lambda_j Ax^j + \sum_{i=1}^q \mu_i Ad^i = b, \\ & \sum_{j=1}^p \lambda_j = 1, \\ & s_1, \dots, s_m, \quad \lambda_1, \dots, \lambda_p, \quad \mu_1, \dots, \mu_q \geq 0, \end{array}$$

onde  $e_k$  denota o  $k$ -ésimo vector da base canónica de  $\mathbb{R}^{m+1}$ , obtém-se

$$B = \begin{bmatrix} 1 & 0 & \dots & 0 & (Ax^1)_1 \\ 0 & 1 & \dots & 0 & (Ax^1)_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & (Ax^1)_m \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix},$$

o  $(m+1)$ -uplo de variáveis básicas  $\begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_m \\ \lambda_1 \end{bmatrix}$ ,  $\bar{b} = \bar{B}^{-1} \begin{bmatrix} b \\ 1 \end{bmatrix}$  e  $[w, \alpha] = \bar{c}_B^T \bar{B}^{-1} =$

$[0, \dots, 0, c^T x^1] \bar{B}^{-1}$ .

$$\text{Note-se que } \bar{B} = \begin{bmatrix} I & Ax^1 \\ 0 & 1 \end{bmatrix} \Rightarrow \bar{B}^{-1} = \begin{bmatrix} I & -Ax^1 \\ 0 & 1 \end{bmatrix}.$$

### Determinação de um minorante para o problema principal

No caso do problema auxiliar ter óptimo finito, a sua resolução permite-nos a determinação de um minorante para a função objectivo do problema principal. Com efeito, sendo  $x^k$  a solução óptima para o problema auxiliar,

$$\bar{z}_k - \bar{c}_k = \max\{(wA - c^T)x : x \in X\} + \alpha,$$

e ainda que

$$\begin{aligned} (wA - c^T)x + \alpha &\leq (wA - c^T)x^k + \alpha = \bar{z}_k - \bar{c}_k \\ &\Downarrow \\ wAx - c^T x + \alpha &\leq \bar{z}_k - \bar{c}_k. \end{aligned} \quad (3.8)$$

Conseqüentemente, sendo  $x$  uma solução admissível arbitrária para o problema principal, i.e., sendo  $x$  tal que  $Ax = b$  e  $x \in X$ , da desigualdade (3.8) decorre a desigualdade  $wb - c^T x + \alpha \leq \bar{z}_k - \bar{c}_k$ , ou seja,

$$\begin{aligned} c^T x &\geq wb + \alpha - (\bar{z}_k - \bar{c}_k) \\ &= [w, \alpha] \begin{bmatrix} b \\ 1 \end{bmatrix} - (\bar{z}_k - \bar{c}_k) \\ &= \bar{c}_B^T \bar{B}^{-1} \begin{bmatrix} b \\ 1 \end{bmatrix} - (\bar{z}_k - \bar{c}_k) \\ &= \zeta - (\bar{z}_k - \bar{c}_k). \end{aligned}$$

Dado que esta desigualdade é verdadeira para todo o  $x \in X$  tal que  $Ax = b$ , podemos concluir que

$$\min\{c^T x : Ax = b, x \in X\} \geq \zeta - (\bar{z}_k - \bar{c}_k),$$

ou seja, que  $wb + \alpha - (\bar{z}_k - \bar{c}_k)$  constitui um minorante para o valor óptimo do problema principal. Porém, deve observar-se que ao longo das sucessivas actualizações do quadro, os minorantes que se vão obtendo não têm (necessariamente) um comportamento monotonamente crescente, pelo que, durante este processo, deveremos guardar o melhor (maior) dos até então determinados.

No caso de aplicação do método de decomposição de Dantzig-Wolfe a problemas de maximização, com as modificações adequadas, a determinação de um majorante para o valor óptimo da função objectivo do problema original segue um procedimento semelhante ao anteriormente descrito.

### Descrição algorítmica

Estamos agora em condições de proceder à descrição algorítmica do método de decomposição de Dantzig-Wolfe.

- Algoritmo para o método de decomposição de Dantzig-Wolfe aplicado a problemas de minimização.
  1. Determinar uma solução básica admissível inicial para o problema principal, a correspondente base  $\bar{B}$  e o quadro do método simplex revisto que lhe está associado:

$s_{\bar{B}}$	$\bar{B}^{-1}$	$\bar{B}^{-1}\bar{b}$
$\lambda_{\bar{B}}$	$[w, \alpha]$	$\zeta$

2. Resolver o problema auxiliar  $\max\{(wA - c^T)x : x \in X\}$ .
  - (a) Se este problema tem óptimo finito e  $x^k$  é a solução básica óptima encontrada, **então** escolher  $y_k = \bar{B}^{-1} \begin{bmatrix} Ax^k \\ 1 \end{bmatrix}$  para coluna candidata a pivotar, fazer  $\bar{z}_k - \bar{c}_k = (wA - c^T)x^k + \alpha$  e continuar para o passo 3;

- (b) Se este problema não tem óptimo finito e  $d^k$  é uma direcção extrema de  $X$  tal que  $(wA - c^T)d^k > 0$ , então escolher  $y_k = \bar{B}^{-1} \begin{bmatrix} Ad^k \\ 0 \end{bmatrix}$  para coluna candidata a pivotar, fazer  $\bar{z}_k - \bar{c}_k = (wA - c^T)d^k$  e continuar para o passo 3;
3. Se  $\bar{z}_k - \bar{c}_k \leq 0$ , então **STOP** a solução corrente é óptima, caso contrário continuar para o passo 4;
4. Se  $y_k \leq 0$ , então **STOP** o problema não tem óptimo finito, caso contrário proceder à respectiva actualização do quadro e voltar ao passo 2;
- Fim do algoritmo.

Este algoritmo facilmente se adapta (com pequenas alterações) a casos de aplicação do método de decomposição de Dantzig-Wolfe a problemas de maximização.

**Exemplo 3.9.** Vamos resolver o programa linear a seguir indicado, por aplicação do método de decomposição de Dantzig-Wolfe.

$$\begin{array}{llllll} \min & -x_1 & - & 2x_2 & - & x_3 \\ \text{s. a} & x_1 & + & x_2 & + & x_3 & \leq & 12, \\ & -x_1 & + & x_2 & & & \leq & 2, \\ & -x_1 & + & 2x_2 & & & \leq & 8, \\ & & & & & x_3 & \leq & 3, \\ & x_1, & & x_2, & & x_3 & \geq & 0. \end{array}$$

**Solução.** Identificando-se a primeira restrição por  $Ax \leq b$  e assumindo-se que as restantes definem a região  $X$ , obtém-se o problema na forma:

$$\begin{array}{ll} \min & \sum_{j=1}^p \lambda_j c^T x^j + \sum_{i=1}^q \mu_i c^T d^i \\ \text{s. a} & s + \sum_{j=1}^p \lambda_j Ax^j + \sum_{i=1}^q \mu_i Ad^i = b, \\ & \sum_{j=1}^p \lambda_j = 1, \\ & \lambda_1, \dots, \lambda_p, \mu_1, \dots, \mu_q \geq 0, \end{array}$$

onde  $c^T = [-1, -2, -1]$ ,  $A = [1, 1, 1]$  e os  $x^j$ s e os  $d^i$ s denotam os pontos extremos e direcções extremas de  $X$ , respectivamente. Uma vez que  $x^1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \in X$  e  $Ax^1 \leq 12$ , introduzindo uma variável de desvio,  $s$ , de modo a obter-se  $s + Ax^1 = 12$ , podemos considerar para solução básica admissível inicial  $\begin{bmatrix} s \\ \lambda_1 \end{bmatrix} = \begin{bmatrix} 12 \\ 1 \end{bmatrix}$ , cuja base associada é a matriz identidade e o vector  $[w, \alpha]$  associado é o vector nulo. Consequentemente, obtém-se o quadro simplex revisto

$s$	1	0	12
$\lambda_1$	0	1	1
	0	0	0

a partir do qual se determina a sequência de problemas auxiliares e consequentes actualizações de quadros simplex revistos que se descreve.

1. O problema auxiliar  $\max\{(wA - c^T)x : x \in X\}$ , ou seja,

$$\begin{array}{rcll} \max & x_1 & + & 2x_2 & + & x_3 & & \\ \text{s. a} & -x_1 & + & x_2 & & & \leq & 2, \\ & -x_1 & + & 2x_2 & & & \leq & 8, \\ & & & & & x_3 & \leq & 3, \\ & x_1, & & x_2, & & x_3, & \geq & 0, \end{array}$$

decompõe-se em dois subproblemas (um associado a  $(x_1, x_2)$  e outro associado a  $x_3$ ). Uma vez que a solução ótima do problema associado a  $x_3$  é  $x_3^* = 3$ , vamos resolver, apenas, o problema associado a  $(x_1, x_2)$ , com o qual, denotando as variáveis de folga por  $y_1$  e  $y_2$ , se obtém a seguinte sequência de quadros reduzidos do método simplex:

(a)

$$\begin{array}{c|cc|c} & x_1 & x_2 & \\ \hline y_1 & -1 & \boxed{1} & 2 \\ y_2 & -1 & 2 & 8 \\ \hline & -1 & -2 & 0 \end{array}$$

(b)

$$\begin{array}{c|cc|c} & x_1 & y_1 & \\ \hline x_2 & -1 & 1 & 2 \\ y_2 & \boxed{1} & -2 & 4 \\ \hline & -3 & 2 & 4 \end{array}$$

(c)

$$\begin{array}{c|cc|c} & y_2 & y_1 & \\ \hline x_2 & 1 & -1 & 6 \\ x_1 & 1 & -2 & 4 \\ \hline & 3 & -4 & 16 \end{array}$$

A partir do último quadro, verifica-se que o problema auxiliar não tem ótimo finito e determina-se a direcção  $d^1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ , a qual, em  $X$  (espaço dos ternos

$(x_1, x_2, x_3)$ ) corresponde à direcção  $d^1 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$  e é tal que

$$(wA - c^T)d^1 = [1, 2, 1] \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = 4$$

Como consequência, a variável  $\mu_1$  é candidata a entrar para a base e, para o problema principal, vem que

$$\bar{z}_1 - \bar{c}_1 = 4,$$

$$Ad^1 = [1, 1, 1] \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = 3,$$

$$y_{p+1} = B^{-1} \begin{bmatrix} Ad^1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}.$$

Assim, executando as correspondentes operações de pivotação, obtém-se

$$\begin{array}{c|ccc|c} & & & \mu_1 \\ \hline s & 1 & 0 & 12 \\ \lambda_1 & 0 & 1 & 1 \\ \hline & 0 & 0 & 0 \end{array} \quad \begin{array}{c|c} \hline 3 \\ \hline 0 \\ \hline 4 \end{array} \quad \rightarrow \quad \begin{array}{c|ccc|c} & & & \mu_1 \\ \hline & \frac{1}{3} & 0 & 4 \\ \lambda_1 & 0 & 1 & 1 \\ \hline & -\frac{4}{3} & 0 & -16 \end{array}$$

2. Do último quadro simplex revisto para o problema principal decorre que  $[w, \alpha] = [-\frac{4}{3}, 0]$ <sup>8</sup> e que o gradiente da função objectivo do problema auxiliar é

$$wA - c^T = [-\frac{4}{3} + 1, -\frac{4}{3} + 2, -\frac{4}{3} + 1] = [-\frac{1}{3}, \frac{2}{3}, -\frac{1}{3}],$$

pelo que o problema auxiliar  $\max\{(wA - c^T)x + \alpha : x \in X\}$  toma o aspecto:

$$\begin{array}{rcll} \max & -\frac{1}{3}x_1 & + & \frac{2}{3}x_2 & - & \frac{1}{3}x_3 & & \\ \text{s. a} & -x_1 & + & x_2 & & & \leq & 2, \\ & -x_1 & + & 2x_2 & & & \leq & 8, \\ & & & & & x_3 & \leq & 3, \\ & x_1, & & x_2, & & x_3, & \geq & 0, \end{array}$$

Uma vez que o valor óptimo para  $x_3$  vem dado por  $x_3^* = 0$ , resta-nos resolver o problema em relação às variáveis  $(x_1, x_2)$ . Para a resolução deste problema, o quadro simplex inicial pode ser obtido a partir do último quadro encontrado durante a resolução do problema auxiliar anterior, depois de devidamente adaptado para o gradiente da nova função objectivo.

$$\begin{array}{c|cc|c} & y_2 & y_1 & \\ \hline x_2 & 1 & -1 & 6 \\ x_1 & 1 & -2 & 4 \\ \hline & \frac{1}{3} & 0 & \frac{8}{3} \end{array}$$

Da análise deste quadro verifica-se que a solução que lhe está associada é óptima (embora não seja única). Logo, uma solução óptima para o problema auxiliar

vem dada por  $x^2 = \begin{bmatrix} 4 \\ 6 \\ 0 \end{bmatrix}$ . Por outro lado,

$$\zeta - (\bar{z}_2 - \bar{c}_2) = -16 - \frac{8}{3} = -\frac{56}{3}$$

constitui um minorante para o valor óptimo do problema principal.

Voltando ao problema principal, sabendo que

$$\bar{z}_2 - \bar{c}_2 = \frac{8}{3},$$

$$Ax^2 = [1, 1, 1] \begin{bmatrix} 4 \\ 6 \\ 0 \end{bmatrix} = 10,$$

$$y_2 = B^{-1} \begin{bmatrix} Ax^2 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 10 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{10}{3} \\ 1 \end{bmatrix},$$

<sup>8</sup>Neste caso, uma vez que  $w < 0$ , a variável  $s$ , que não é considerada no problema auxiliar, não é candidata a entrar para a base, dado que  $\bar{z}_s - \bar{c}_s = c_B B^{-1} a_s - c_s = [w, \alpha] \begin{bmatrix} 1 \\ 0 \end{bmatrix} - 0 = w$ . Porém, nos casos em que se obtém, pelo menos um coeficiente de  $w$  positivo, deve ter-se em conta que a correspondente variável de desvio é candidata a entrar para a base.

actualizando o quadro simplex revisto do problema principal, obtém-se

$$\begin{array}{c|cc|c}
 \mu_1 & \frac{1}{3} & 0 & 4 \\
 \lambda_1 & 0 & 1 & 1 \\
 \hline
 & -\frac{4}{3} & 0 & -16
 \end{array}
 \quad
 \begin{array}{c|c}
 \lambda_2 \\
 \frac{10}{3} \\
 \hline
 1 \\
 \hline
 \frac{8}{3}
 \end{array}
 \rightarrow
 \begin{array}{c|cc|c}
 \mu_1 & \frac{1}{3} & -\frac{10}{3} & \frac{2}{3} \\
 \lambda_2 & 0 & 1 & 1 \\
 \hline
 & -\frac{4}{3} & -\frac{8}{3} & -\frac{56}{3}
 \end{array}$$

Uma vez que o minorante, anteriormente determinado, coincide com o valor que a função objectivo do problema principal toma para a solução admissível corrente, podemos concluir que ela é óptima.

Logo, a solução óptima do problema original vem dada por

$$x^* = \lambda_2 x^2 + \mu_1 d^1 \Leftrightarrow x^* = \begin{bmatrix} 4 \\ 6 \\ 0 \end{bmatrix} + \frac{2}{3} \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{16}{3} \\ \frac{20}{3} \\ 0 \end{bmatrix}.$$

### 3.5 Exercícios

**Exercício 3.1.** Aplique o método simplex ao programa linear

$$\begin{array}{rcll}
 \max & 2x_1 & + & x_2 \\
 \text{s. a.} & -x_1 & + & x_2 \leq 1 \\
 & x_1 & - & 2x_2 \leq 2 \\
 & x_1, & & x_2 \leq 0
 \end{array}$$

e verifique se o problema tem óptimo finito.

**Exercício 3.2.** Considerando o problema de optimização linear

$$\min\{c^T x : x = \begin{bmatrix} 0 \\ 0 \\ 4 \\ 1 \\ 6 \\ 0 \end{bmatrix} + \theta_1 \begin{bmatrix} 1 \\ 0 \\ 1 \\ -3 \\ 0 \\ 0 \end{bmatrix} + \theta_2 \begin{bmatrix} 0 \\ 1 \\ -1 \\ 1 \\ -2 \\ 0 \end{bmatrix} + \theta_3 \begin{bmatrix} 0 \\ 0 \\ -1 \\ 2 \\ -1 \\ 1 \end{bmatrix} \geq 0, \theta_1, \theta_2, \theta_3 \geq 0\},$$

onde  $c^T = [1, 1, 4, 0, 0, 0]$ , determine o quadro simplex reduzido relativo à solução básica admissível  $[0, 0, 4, 1, 6, 0]^T$ .

**Exercício 3.3.** Considere que durante a aplicação do método simplex na resolução de um dado programa linear de minimização na forma padrão,  $PL(A, b, c)$ , se obtiveram os quadros simplex reduzidos consecutivos:

$$\begin{array}{c|cc}
 & x_N & \\
 \hline
 x_B & B^{-1}N & B^{-1}b \\
 \hline
 & c_B B^{-1}N - c_N & c_B B^{-1}b
 \end{array}
 \quad
 \begin{array}{c|cc}
 & x_{N'} & \\
 \hline
 x_{B'} & B'^{-1}N' & B'^{-1}b \\
 \hline
 & c_{B'} B'^{-1}N' - c_{N'} & c_{B'} B'^{-1}b
 \end{array}$$

1. Prove que as operações de pivotação efectuadas no quadro associado à matriz básica  $B$  conduzem à obtenção da solução básica  $x'_B$  determinada pelo quadro associado à nova matriz básica  $B'$ .
2. Prove que, no caso das soluções básicas óptimas serem não degeneradas,

$$c_{B'}B'^{-1}N' - c_{N'} \leq 0$$

é uma condição necessária e suficiente de optimalidade para a solução básica admissível obtida no quadro simplex associado a  $B'$ .

**Exercício 3.4.** Considere o programa linear de maximização na forma canónica  $PL(A,b,c)$ , para o qual foi determinado o quadro simplex:

	$x_4$	$x_2$	$x_6$	
$x_1$	1	0	-2	$\alpha_1$
$x_5$	-2	-3	1	$\alpha_2$
$x_3$	1	1	1	$\alpha_3$
	3	4	0	$\delta$

onde  $c^T = [1, 2, 2]$ ,  $b^T = [2, 3, 1]$  e  $x_4$ ,  $x_5$  e  $x_6$  denotam as variáveis de desvio.

1. Determine a inversa da matriz básica e os valores de  $\alpha_1$ ,  $\alpha_2$ ,  $\alpha_3$  e  $\delta$ .
2. Determine a solução óptima do programa linear  $PL(A,b,c)$ .

**Exercício 3.5.** Considere o quadro simplex reduzido a seguir apresentado.

	$x_2$	$x_4$	
$x_1$	-2	1	1
$x_3$	-1	2	2
$x_5$	0	3	4
	$\alpha$	$\beta$	$\delta$

Sabendo que o programa linear original associado a este quadro simplex é

$$\min\{c^T x : Ax \leq b, x \leq 0\},$$

$b \geq 0$  e que se tomou para solução básica inicial  $x_B = \begin{bmatrix} x_3 \\ x_4 \\ x_5 \end{bmatrix} = b$ , onde  $x_3$ ,  $x_4$  e  $x_5$  são as variáveis de desvio, responda às seguintes questões:

1. Determine a inversa da matriz básica corrente.
2. Sabendo que  $c^T = [1, 1]$ , determina  $\alpha$  e  $\beta$  e  $\delta$ .
3. Verifique se o problema tem óptimo finito, justificando.

**Exercício 3.6.** Considere o programa linear:

$$\begin{array}{rcllcl}
 \min & -2x_1 & + & x_2 & & \\
 \text{s. a.} & x_1 & + & x_2 & \geq & 2 \\
 & x_1 & - & 2x_2 & \leq & 1 \\
 & x_1 & & & \leq & 3 \\
 & x_1, & & x_2 & \geq & 0
 \end{array}$$

1. Determine uma solução básica admissível para este problema (utilizando o método das duas fases).

2. A partir da solução básica encontrada na alínea anterior, determine a solução ótima pelo método simplex revisito.

**Exercício 3.7.** Resolva o programa linear

$$\begin{array}{rcllcl}
 \min & -x_1 & + & x_2 & & \\
 \text{s. a.} & 4x_1 & + & x_2 & \geq & 4 \\
 & x_1 & + & x_2 & \geq & 2 \\
 & x_1 & + & 3x_2 & \geq & 3 \\
 & x_1, & & x_2 & \geq & 0
 \end{array}$$

por aplicação do método do big-M.

**Exercício 3.8.** A partir do quadro simplex reduzido, a seguir apresentado,

	$x_1$	$x_2$	$x_3$	$x_4$	
$x_5$	0.5	-5.5	-2.5	9	0
$x_6$	0.5	-1.5	-0.5	1	0
$x_7$	1	0	0	0	1
	10	-57	-9	-24	

determine o quadro simplex ótimo (supondo que o problema é de minimização) por aplicação da regra de Bland.

**Exercício 3.9.** A partir da resolução do Exercício 3.8, evidencie a propriedade de monotonicidade da regra de Bland.

**Exercício 3.10.** Considere o programa linear

$$\begin{array}{rcllcl}
 \max & x_1 & + & x_2 & + & 3x_3 \\
 \text{s. a} & x_1 & + & x_2 & + & x_3 \leq 12, \\
 & x_1 & + & 2x_2 & & \leq 5, \\
 & & & x_2 & + & x_3 \leq 5, \\
 & 2 & \leq & x_1 & \leq & 6 \\
 & 1 & \leq & x_2 & \leq & 7 \\
 & & & & & x_3 \geq 0.
 \end{array}$$

e resolva-o aplicando o método simplex para variáveis limitadas.

**Exercício 3.11.** Resolva o programa linear a seguir indicado, por aplicação do método simplex para problemas com variáveis limitadas, a partir da solução básica inicial  $x_1 = x_2 = x_3 = 1$ .

$$\begin{array}{rllll} \min & x_1 & - & 3x_2 & - & 2x_3 \\ \text{s. a} & -x_1 & + & x_2 & + & x_3 & \leq & 4, \\ & x_1 & + & x_2 & + & 2x_3 & \leq & 10, \\ 1 \leq & x_1 & & & & & \leq & 4, \\ 1 \leq & & & x_2 & & & \leq & 6, \\ 1 \leq & & & & & x_3 & \leq & 4. \end{array}$$

**Exercício 3.12.** Dado um programa linear de minimização,  $PL(A, b, c)$ , com as restrições de caixa  $l \leq x \leq u$ , prove que uma partição básica admissível  $(J_B, J_{N_l}, J_{N_u})$  define uma solução ótima se

$$\max\left\{\max_{j \in J_{N_l}} (z_j - c_j), \max_{j \in J_{N_u}} -(z_j - c_j)\right\} \leq 0.$$

**Exercício 3.13.** Considere uma partição básica admissível,  $(J_B^0, J_{N_l}^0, J_{N_u}^0)$ , para um programa de minimização na forma padrão,  $PL(A, b, c)$ , com restrições de caixa  $l \leq x \leq u$ , e bem ainda que no quadro simplex que lhe está associado,  $\hat{b}$  denota o m-uplo dos respectivos termos independentes. Supondo que, neste quadro,  $\exists k \in J_{N_u}^0$  tal que  $z_k - c_k < 0$  e que  $\inf \Omega(J_{N_l}, J_{N_u}, y_k) = \frac{\hat{b}_q - (l_B)_q}{-y_{qk}}$ , com  $y_{qk} < 0$ , prove ou refute (justificadamente) cada uma das seguintes afirmações:

1. A  $k$ -ésima coluna é candidata a pivotal.
2. A atualização do quadro simplex, escolhendo a  $k$ -ésima coluna para pivotal, implica que, no próximo quadro, o índice  $q$  pertença ao conjunto  $J_{N_l}^1$ .
3. Se  $\frac{\hat{b}_q - (l_B)_q}{-y_{qk}} = u_k - l_k$  e a  $k$ -ésima coluna é escolhida para pivotal, então a próxima solução básica admissível (ou partição básica admissível) é degenerada.
4. Nas condições da alínea anterior, optando-se pela permanência da base corrente, se dos sucessivos quadros simplex decorrerem as partições básicas admissíveis  $(J_B^i, J_{N_l}^i, J_{N_u}^i)$  para  $i = 1, \dots, s$ , então  $\forall i \in \{1, \dots, s\}$ ,  $k \notin J_{N_u}^i$ .

**Exercício 3.14.** Considere o programa linear

$$\begin{array}{rllll} \max & -2x_1 & - & x_2 & - & x_3 & + & x_4 \\ \text{s. a} & x_1 & & & + & x_3 & & \leq & 2, \\ & x_1 & & & & & & \leq & 2, \\ & x_1 & + & x_2 & & & & \leq & 5, \\ & & & & - & x_3 & + & x_4 & \leq & 2, \\ & x_1, & x_2, & x_3, & x_4 & \geq & 0, \end{array}$$

e responda às seguintes questões:

1. Aplique o método de decomposição de Dantzig-Wolfe para a resolução deste problema, tomando para problema principal o que se define à custa da primeira restrição.
2. Indique, para cada um dos problemas auxiliares, resolvidos na alínea anterior, o valor do minorante determinado pelo método de decomposição de Dantzig-Wolfe.

**Exercício 3.15.** Supondo que, por aplicação do método de decomposição de Dantzig-Wolfe ao programa linear  $PL(A, b, c)$ ,  $\min\{c^T x : Ax = b, x \in X\}$ , onde  $X$  é um conjunto poliédrico não vazio com determinada estrutura, se obtém o quadro

$$\begin{array}{c|c|c} \hline & & \\ \hline \lambda_B & B^{-1} & B^{-1}b \\ \hline & [w, \alpha] & \zeta \\ \hline \end{array}$$

e que  $\tau = \max\{(wA - c^T)x : x \in X\}$ , prove ou refute (justificadamente) cada uma das afirmações:

1. Se  $\tau < \infty$ , então  $wb + \alpha - \tau$  constitui um minorante para  $PL(A, b, c)$ .
2. Supondo que  $x^k$  é uma solução ótima para o problema auxiliar e que  $y_k = B^{-1} \begin{bmatrix} Ax^k \\ 1 \end{bmatrix}$ , se  $\min\{\frac{(B^{-1}b)_i}{y_{ik}} : 1 \leq i \leq m, y_{ik} > 0\} = 1$ , então a solução básica determinada após as operações de pivotação (com  $y_k$  como coluna pivotal) é uma solução ótima para o problema principal.
3. Se  $\tau = \infty$ , então  $PL(A, b, c)$  não tem ótimo finito.

## Capítulo 4

# Dualidade e Análise Pós-optimal

O estudo da dualidade em Programação Matemática leva-nos à consideração de um problema (que, geralmente, se designa por problema dual) distinto daquele que pretendemos resolver, mas cuja abordagem nos permite obter algumas conclusões directamente relacionadas com o problema original (problema primal), nomeadamente, com as condições de optimalidade.

### 4.1 Resultados básicos sobre dualidade

**Definição 4.1.** *Dado o problema de programação linear de minimização na forma canónica*

$$\min c^T x \quad (4.1)$$

$$s. a \ Ax \geq b, \quad (4.2)$$

$$x \geq 0, \quad (4.3)$$

o qual consideramos ser o problema primal e denotamos por  $(P)$ , designa-se por dual de  $(P)$  e denota-se por  $(D)$ , o programa linear de maximização na forma canónica  $PL(A^T, c, b)$

$$\max b^T y$$

$$s. a \ A^T y \leq c,$$

$$y \geq 0,$$

o qual, muitas vezes, também se escreve na forma equivalente:

$$\max y^T b \quad (4.4)$$

$$s. a \ y^T A \leq c^T, \quad (4.5)$$

$$y \geq 0, \quad (4.6)$$

Esta definição de dual com facilidade se estende a problemas de programação linear na forma padrão ou outra. Por exemplo, considerando o problema de programação linear de minimização na forma padrão  $PL(A, b, c)$  e escrevendo-o na forma

$$\begin{array}{l} \min c^T x \\ \text{s. a } Ax \geq b, \\ \quad -Ax \geq -b, \\ \quad x \geq 0, \end{array} \Leftrightarrow \begin{array}{l} \min c^T x \\ \text{s. a } \begin{bmatrix} A \\ -A \end{bmatrix} x \geq \begin{bmatrix} b \\ -b \end{bmatrix}, \\ \quad x \geq 0, \end{array}$$

recorrendo à definição anterior, com facilidade se obtém o respectivo dual, ou seja,

$$\begin{array}{l} \max (y' - y'')^T b \\ \text{s. a } (y' - y'')^T A \leq c^T, \\ \quad y', y'' \geq 0, \end{array} \Leftrightarrow \begin{array}{l} \max y^T b \\ \text{s. a } y^T A \leq c^T, \\ \quad y \text{ (livre)}. \end{array}$$

**Teorema 4.1** (da dualidade fraca). *Considere o par de problemas primal-dual (P): (4.1)-(4.3) e (D): (4.4)-(4.6). Se  $x'$  é admissível para (P) e  $y'$  é admissível para (D), então*

$$y'^T b \leq c^T x'.$$

**Demonstração.** Uma vez que  $x' \geq 0$  e  $y'^T A \leq c^T$ , vem

$$y'^T A x' \leq c^T x'. \quad (4.7)$$

Por outro lado, dado que  $y' \geq 0$  e  $Ax' \geq b$ , vem

$$y'^T b \leq y'^T A x'. \quad (4.8)$$

Como consequência, tendo em conta (4.7) e (4.8), obtém-se  $y'^T b \leq c^T x'$ .  $\square$

Como corolário deste teorema decorrem as conclusões que a seguir se indicam e cuja prova fica como exercício.

1. Se ambos os problemas (primal e dual) têm soluções admissíveis, então ambos os problemas têm ótimo finito.
2. Se o problema primal admite uma sucessão de soluções admissíveis  $\{x^k\}_{k \in \mathbb{N}}$ , tal que  $c^T x^k \rightarrow -\infty$ , então o dual não tem qualquer solução admissível.
3. Se o problema dual admite uma sucessão de soluções admissíveis  $\{y^k\}_{k \in \mathbb{N}}$  tais que  $b^T y^k \rightarrow \infty$ , então o primal não tem qualquer solução admissível.

Das condições de optimalidade do método simplex decorre que, se o problema primal tem solução ótima (i.e., tem ótimo finito), então o respectivo dual também tem e os correspondentes valores ótimos coincidem. Com efeito, considere-se o quadro simplex reduzido ótimo associado ao problema de programação linear de minimização na forma padrão  $PL(A, b, c)$ :

	$x_N$	
$x_B$	$B^{-1}N$	$B^{-1}b$
	$c_B B^{-1}N - c_N$	$c_B B^{-1}b$

Tendo em conta a linha de custos reduzidos,  $c_B B^{-1}N - c_N \leq 0$ , supondo (sem perda de generalidade) que a matriz  $A$  se parte em  $A = [B, N]$ , fazendo  $y^T = c_B B^{-1}$ , vem

$$y^T A = c_B B^{-1} [B, N] = [c_B, c_B B^{-1} N] \leq [c_B, c_N] = c^T.$$

Logo, conclui-se que  $y^T = c_B B^{-1}$  é admissível para  $(D)$ . Por outro lado,  $y^T b = c_B B^{-1} b$  o que, mais uma vez, de acordo com o quadro, é precisamente o valor ótimo de  $(P)$ . Como consequência, do teorema da dualidade fraca (Teorema 4.1) decorre que os valores ótimos de  $(P)$  e  $(D)$  coincidem.

A teoria da programação linear permite obter resultados ainda mais fortes (do que os anteriormente referidos) sobre as relações primal-dual.

**Teorema 4.2** (da dualidade forte). *Considere o par de problemas primal-dual  $(P)$ : (4.1)-(4.3) e  $(D)$ : (4.4)-(4.6).*

1. *Se ambos os problemas,  $(P)$  e  $(D)$ , são consistentes (i.e., ambos têm soluções admissíveis), então ambos têm solução ótima e os respectivos valores ótimos coincidem.*
2. *Se  $(P)$  é inconsistente, então o dual é consistente sse não tem ótimo finito.*
3. *Se  $(D)$  é inconsistente, então o primal é consistente sse não tem ótimo finito.*
4. *Podem acontecer que nem  $(P)$  nem  $(D)$  seja consistente.*

**Demonstração.** 1. Considere-se o sistema de inequações

$$\begin{array}{rcl} Ax & & \geq b \\ & - & A^T y \geq -c \\ -c^T x + b^T y & \geq & 0 \\ x, & y & \geq 0 \end{array}$$

o qual se pode escrever na forma

$$\left\{ \begin{array}{l} \left[ \begin{array}{cc} A & 0 \\ 0 & -A^T \\ -c^T & b^T \end{array} \right] \begin{bmatrix} x \\ y \end{bmatrix} \geq \begin{bmatrix} b \\ -c \\ 0 \end{bmatrix} \\ x, y \geq 0 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} \bar{A} \bar{x} \geq \bar{b} \\ \bar{x} \geq 0 \end{array} \right. \quad (4.9)$$

onde  $\bar{A} = \begin{bmatrix} A & 0 \\ 0 & -A^T \\ -c^T & b^T \end{bmatrix}$ ,  $\bar{b} = \begin{bmatrix} b \\ -c \\ 0 \end{bmatrix}$  e  $\bar{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ .

Se este sistema admite uma solução  $(x', y')$ , então  $x'$  é admissível para  $(P)$  e  $y'$  é admissível para  $(D)$  e são tais que  $c^T x' \leq b^T y'$ . Logo, por aplicação do teorema da dualidade fraca (Teorema 4.1), conclui-se que  $c^T x' = b^T y'$  e, como consequência,  $x'$  é solução ótima para  $(P)$  e  $y'$  é solução ótima para  $(D)$ .

Suponha que o sistema (4.9) não admite qualquer solução. Então, pelo lema de Farkas (Lema 1.4), o sistema

$$\left\{ \begin{array}{l} \left[ \begin{array}{ccc} u^T & v^T & \lambda \end{array} \right] \bar{A} \leq 0 \\ \left[ \begin{array}{ccc} u^T & v^T & \lambda \end{array} \right] \bar{b} > 0 \\ u, v, \lambda \geq 0 \end{array} \right. \Leftrightarrow \begin{array}{rcl} A^T u & - & \lambda c \leq 0 \\ & - & Av + \lambda b \leq 0 \\ b^T u & - & c^T v > 0 \end{array}$$

tem uma solução  $(u', v', \lambda')$ .

Se  $\lambda' \neq 0$ , então (sem perda de generalidade) podemos admitir  $\lambda' = 1$ . Como

consequência,  $u'$  é admissível para  $(D)$ ,  $v'$  é admissível para  $(P)$  e  $b^T u' > c^T v'$ , o que entra em contradição com o teorema fraco da dualidade (Teorema 4.1). Logo  $\lambda' = 0$ .

Se  $c^T v' \geq 0$ , então  $b^T u > 0$  e, novamente, por aplicação do Lema 1.4,  $(P)$  não tem qualquer solução admissível, o que contraria a hipótese.

Assim, suponha  $c^T v' < 0$ . Então sendo  $x'$  uma solução admissível para  $(P)$   $\forall \theta \geq 0$ ,  $x_\theta = x' + \theta v'$  é admissível para  $(P)$  e é tal que  $c^T x_\theta = c^T x' + \theta c^T v' \rightarrow -\infty$  quando  $\theta \rightarrow \infty$ , logo, pelo teorema da dualidade fraca, somos levados a concluir que o dual não tem qualquer solução admissível, o que contraria a hipótese.

2. Dado que  $(P)$  é inconsistente, pelo lema de Farkas (Lema 1.4), o sistema

$$A^T y \leq 0 \quad (4.10)$$

$$b^T y > 0 \quad (4.11)$$

$$y \geq 0 \quad (4.12)$$

tem uma solução  $y'$ . Seja  $y^*$  uma solução admissível para  $(D)$  e, para cada  $\tau \geq 0$ , seja  $y_\tau = y^* + \tau y'$ . Então  $\forall \tau \geq 0$   $y_\tau^T \geq 0$  e  $A^T y_\tau = A^T y^* + \tau A^T y' \leq c$ , i. e.,  $y_\tau$  é admissível para o dual  $(D)$   $\forall \tau \geq 0$ . Logo, dado que  $b^T y_\tau = b^T y^* + \tau b^T y' \rightarrow \infty$  quando  $\tau \rightarrow \infty$ , pode concluir-se que  $(D)$  não tem óptimo finito.

3. Esta prova é idêntica à anterior.

4. Neste caso a prova pode fazer-se exibindo o seguinte exemplo de um par de problemas  $(P)$ - $(D)$ , em que ambos são inconsistentes:

$$(P) \quad \begin{array}{ll} \min & -x_1 \quad - \quad x_2 \\ \text{s. a} & x_1 \quad - \quad x_2 \geq 1 \\ & -x_1 \quad + \quad x_2 \geq 1 \\ & x_1, \quad x_2 \geq 0 \end{array}$$

$$(D) \quad \begin{array}{ll} \min & y_1 \quad + \quad y_2 \\ \text{s. a} & y_1 \quad - \quad y_2 \geq -1 \\ & -y_1 \quad + \quad y_2 \geq -1 \\ & y_1, \quad y_2 \geq 0 \end{array}$$

□

Como consequência, do Teorema 4.2 decorre o teorema que se segue, no qual se estabelece uma relação de complementaridade entre soluções óptimas primais e duais.

**Teorema 4.3** (da complementaridade). *Se  $x^*$  e  $y^*$  são soluções óptimas para o primal  $(P)$  e para o dual  $(D)$ , respectivamente, então verificam-se as seguintes igualdades:*

$$\begin{cases} (c_j - y^{*T} a_j) x_j^* = 0, & \text{para } j = 1, \dots, n \\ y_i^* (a^i x^* - b_i) = 0, & \text{para } i = 1, \dots, m \end{cases}$$

onde  $a_j$ , para  $j = 1, \dots, n$ , denota a  $j$ -ésima coluna de  $A$ ,  $a^i$ , para  $i = 1, \dots, m$ , denota a  $i$ -ésima linha de  $A$ ,  $c_j$ , para  $j = 1, \dots, n$ , denota o  $j$ -ésimo coeficiente do gradiente da função objectivo de  $(P)$  e  $b_i$ , para  $i = 1, \dots, m$ , denota o  $i$ -ésimo coeficiente do gradiente da função objectivo de  $(D)$ .

**Demonstração.** Dado que  $x^*$  e  $y^*$  são admissíveis para  $(P)$  e  $(D)$ , respectivamente, conclui-se que  $x^* \geq 0$  e  $Ax^* \geq b$ , e  $y^* \geq 0$  e  $y^{*T} A \leq c^T$ . Logo, vem que

$$\begin{cases} y^{*T}Ax^* \geq y^{*T}b \\ y^{*T}Ax^* \leq c^T x^* \end{cases},$$

ou seja,  $y^{*T}b \leq y^{*T}Ax^* \leq c^T x^*$ . Por outro lado, uma vez que  $y^*$  e  $x^*$  são soluções óptimas, pelo teorema da dualidade forte (Teorema 4.2), sabe-se que  $y^{*T}b = c^T x^*$ . Assim, conclui-se que  $y^{*T}b = y^{*T}Ax^* = c^T x^*$ , i. e.,

$$\begin{cases} y^{*T}(Ax^* - b) = 0 \\ (c^T - y^{*T}A)x^* = 0 \end{cases} \Leftrightarrow \begin{cases} \sum_{i=1}^m y_i^*(a^i x^* - b_i) = 0 \\ \sum_{j=1}^n (c_j - y^{*T}a_j)x_j^* = 0 \end{cases} \quad (4.13)$$

Consequentemente, dado que  $\forall i \in \{1, \dots, m\}$   $y_i^* \geq 0$  e  $a^i x^* - b_i \geq 0$ , e  $\forall j \in \{1, \dots, n\}$   $x_j^* \geq 0$  e  $c_j - y^{*T}a_j \geq 0$ , de (4.13), vem

$$\begin{cases} y_i^*(a^i x^* - b_i) = 0 & \text{para } i = 1, \dots, m \\ (c_j - y^{*T}a_j)x_j^* = 0 & \text{para } j = 1, \dots, n \end{cases} \quad (4.14)$$

□

Deste teorema decorrem as seguintes conclusões:

1.  $\forall i \in \{1, \dots, m\}$   $\begin{cases} y_i^* > 0 \Rightarrow a^i x^* = b_i \\ a^i x^* > b_i \Rightarrow y_i^* = 0 \end{cases}$
2.  $\forall j \in \{1, \dots, n\}$   $\begin{cases} x_j^* > 0 \Rightarrow y^{*T}a_j = c_j \\ y^{*T}a_j < c_j \Rightarrow x_j^* = 0 \end{cases}$ .

Assim, podemos afirmar que as variáveis de decisão duais positivas correspondem restrições primais activas (i.e., variáveis de desvio primais nulas) e as restrições primais não activas (i.e., variáveis de desvio primais positivas) correspondem variáveis de decisão duais nulas e reciprocamente, ou seja, as variáveis de decisão primais positivas correspondem restrições duais activas (i.e., variáveis de desvio duais nulas) e as restrições duais não activas (i.e., variáveis de desvio duais positivas) correspondem variáveis de decisão primais nulas. Devido a esta propriedade de complementaridade entre variáveis de decisão primais e variáveis duais de desvio e variáveis de decisão duais e variáveis primais de desvio, o teorema anterior é usualmente designado por teorema da complementaridade (ou por teorema da complementaridade de “slacks”). Quando o recíproco da propriedade de complementaridade, anteriormente referida, acontece, ou seja,

1.  $\forall i \in \{1, \dots, m\}$   $\begin{cases} a^i x^* = b_i \Rightarrow y_i^* > 0 \\ y_i^* = 0 \Rightarrow a^i x^* > b_i \end{cases}$
2.  $\forall j \in \{1, \dots, n\}$   $\begin{cases} y^{*T}a_j = c_j, \Rightarrow x_j^* > 0 \\ x_j^* = 0 \Rightarrow y^{*T}a_j < c_j \end{cases}$ ,

dizemos que o par de soluções  $(x^*, y^*)$  satisfaz a propriedade da complementaridade estrita.

Em programação linear, quando o par de problemas primal dual  $(P)$ - $(D)$  tem soluções óptimas, verifica-se que existe um par de soluções óptimas que tem a propriedade da complementaridade estrita, conforme o teorema (da complementaridade estrita) a seguir indicado estabelece.

**Teorema 4.4** (da complementaridade estrita). *Cada par de problemas primal-dual, com soluções óptimas, admite um par de soluções óptimas com a propriedade da complementaridade estrita.*

**Demonstração.** Primeiramente vamos provar a existência de um par de soluções  $(x', y')$ , óptimas para  $(P)$  e  $(D)$ , respectivamente, com a propriedade da complementaridade estrita entre as variáveis de desvio duais e as variáveis de decisão primais. Seja  $Y^* = \{\sum_{j=1}^r \lambda_j y^{*j} : \sum_{j=1}^r \lambda_j = 1, \lambda_j \geq 0, j = 1, \dots, r\}$  o conjunto de soluções óptimas para  $(D)$  e  $I = \{i : \exists y^{*k}, \text{ para o qual } s_i^* = c_i - (y^{*k})^T a_i > 0\}$ , então  $\exists y^* \in Y^*$  tal que  $\forall i \in I \ s_i^* = c_i - (y^*)^T a_i > 0$ . Com efeito, fazendo  $y^* = \sum_{j=1}^r \lambda_j y^{*j}$ , com  $\sum_{j=1}^r \lambda_j = 1$  e  $\lambda_j > 0, j = 1, \dots, r$ , vem que  $\forall i \in I$

$$\begin{aligned} s_i^* &= c_i - (y^*)^T a_i \\ &= c_i - \left( \sum_{j=1}^r \lambda_j (y^{*j})^T \right) a_i \\ &= \left( \sum_{j=1}^r \lambda_j \right) c_i - \left( \sum_{j=1}^r \lambda_j (y^{*j})^T a_i \right) \\ &= \sum_{j=1}^r \lambda_j [c_i - (y^{*j})^T a_i] > 0, \end{aligned}$$

dado que  $c_i - (y^{*j})^T a_i \geq 0 \ \forall j \in \{1, \dots, r\}$  e  $\exists y^{*k}$  tal que  $s_i^* = c_i - (y^{*k})^T a_i > 0$  e  $\lambda_k > 0$ . Consequentemente, para as variáveis de desvio que tomam valores positivos, para pelo menos uma solução óptima para  $(D)$ , a propriedade da complementaridade estrita é evidente.

Logo, resta fazer a prova para as componentes cujos índices estão no conjunto  $J = \{j : s_j^* = c_j - (y^*)^T a_j = 0 \ \forall y^* \in Y^*\}$ , onde, tal como anteriormente,  $Y^*$  denota o conjunto de soluções óptimas para  $(D)$ .

Sendo  $\nu^*$  o valor óptimo para  $(P)$  (e consequentemente para  $(D)$ ) o programa linear

$$\begin{array}{ll} \max & \sum_{j \in J} e_j^T s \\ \text{s. a} & A^T y + s = c, \\ & -b^T y \leq -\nu^*, \\ & y, s \geq 0, \end{array}$$

(onde  $e_j$  denota o  $j$ -ésimo vector da base canónica de  $\mathbb{R}^n$ ), tem valor óptimo nulo. Logo, o respectivo dual

$$\begin{array}{ll} \min & c^T x - \nu^* \lambda \\ \text{s. a} & Ax - b\lambda \geq 0, \\ & x_j \geq 1 \ \forall j \in J, \\ & x_j \geq 0 \ \forall j \notin J, \\ & \lambda \geq 0, \end{array}$$

tem solução óptima  $\begin{bmatrix} x' \\ \lambda' \end{bmatrix}$ , com valor óptimo  $c^T x' - \nu^* \lambda' = 0$ .

Supondo que  $x^*$  é uma solução óptima para  $(P)$ , obtém-se uma das seguinte situações:

1.  $\lambda' = 0$ , e então vem que

$$c^T x' = 0, \quad Ax' \geq 0, \quad x'_j \geq 1 \ \forall j \in J \text{ e } x'_j \geq 0 \ \forall j \notin J,$$

donde  $\forall \tau \geq 0$ ,  $x^* + \tau x'$  não só é uma solução óptima para  $(P)$ , uma vez que

$$x^* + \tau x' \geq 0, \quad A(x^* + \tau x') = Ax^* + \tau Ax' \geq b \text{ e } c^T(x^* + \tau x') = c^T x^* + \tau c^T x' = \nu^*,$$

como, para  $\tau > 0$ ,  $x^* + \tau x'$  é tal que  $(x^* + \tau x')_j > 0 \ \forall j \in J$ ;

2.  $\lambda' > 0$ , e então

$$c^T \frac{x'}{\lambda} = \nu^*, \quad \frac{x'}{\lambda} \geq 0, \quad A \frac{x'}{\lambda} \geq b \quad \text{e} \quad \frac{1}{\lambda} x'_j \geq \frac{1}{\lambda} > 0 \quad \forall j \in J.$$

A prova da existência de um par de soluções  $(x'', y'')$  ótima para  $(P)$  e  $(D)$ , respectivamente, com a propriedade da complementaridade estrita entre as variáveis de desvio primais e as variáveis de decisão duais, é idêntica à anterior.

Um vez provado que existe um par de soluções  $(x', y')$  ótimas para  $(P)$  e  $(D)$ , respectivamente, com a propriedade da complementaridade estrita entre as variáveis de desvio duais e as variáveis de decisão primais e que existe um par de soluções  $(x'', y'')$  ótima para  $(P)$  e  $(D)$ , respectivamente, com a propriedade da complementaridade estrita entre as variáveis de desvio primais e as variáveis de decisão duais,  $\forall \lambda \in ]0, 1[$ , fazendo  $x^* = \lambda x' + (1 - \lambda)x''$  e  $y^* = \lambda y' + (1 - \lambda)y''$ , fica provada a existência de um par de soluções  $(x^*, y^*)$  ótimas para  $(P)$  e  $(D)$ , respectivamente, com a propriedade da complementaridade estrita.  $\square$

Conforme decorre das relações primal-dual abordadas, o método simplex, ao longo das sucessivas actualizações do quadro, lida com soluções básicas, cujas bases definem (simultaneamente) soluções primais admissíveis e duais complementares, com vista à determinação de uma base que defina uma solução primal admissível e uma solução dual complementar que seja admissível. Quando tal acontece, significa que foram atingidas as condições de optimalidade. É, precisamente, com base nestas propriedades que se desenvolveu uma variante do método simplex, que se designa por método dual simplex.

## 4.2 Método dual simplex

Enquanto no método simplex clássico (também conhecido por método primal-simplex) a base corrente é primal admissível e as sucessivas actualizações do quadro (operações de pivotação) são executadas de modo a respeitar essa admissibilidade e a diminuir o valor corrente da função objectivo do problema primal, com o método dual simplex a base corrente satisfaz as condições de optimalidade na linha de custos reduzidos e o quadro é actualizado sem que estas condições sejam violadas e de modo a incrementar a função objectivo do problema dual. No primeiro caso (método primal simplex) pretendemos obter as condições de optimalidade na linha de custos reduzidos (as quais equivalem à admissibilidade dual) e no segundo (método dual simplex) tem-se em vista a obtenção da admissibilidade primal.

Considere-se o seguinte programa linear de minimização na forma canónica:

$$\begin{array}{ll} \min & c^T x \\ \text{s. a} & Ax \geq b, \\ & x \geq 0, \end{array}$$

e seja  $B$  uma submatriz básica de  $\bar{A} = [-I, A] \in \mathbb{R}^{m \times n}$  não, necessariamente, admissível. Considere-se ainda o correspondente quadro simplex

	$x_{N_1}$	...	$x_{N_k}$	...	$x_{N_p}$	
$x_{B_1}$	$y_{11}$	...	$y_{1k}$	...	$y_{1p}$	$b_1$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$x_{B_r}$	$y_{r1}$	...	$y_{rk}$	...	$y_{rp}$	$\bar{b}_r$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$x_{B_m}$	$y_{m1}$	...	$y_{mk}$	...	$y_{mp}$	$\bar{b}_m$
	$z_1 - c_{N_1}$	...	$z_k - c_{N_k}$	...	$z_p - c_{N_p}$	$c_B B^{-1} b$

no qual, sendo  $y^T = c_B B^{-1}$ , vem que  $z_j - c_{N_j} = y^T a_{N_j} - c_{N_j}$ , para  $j = 1, \dots, p$ . Se  $z_j - c_{N_j} \leq 0 \forall j \in \{1, \dots, p\}$ , então

$$y^T [B, N] - c^T = [c_B B^{-1} B, c_B B^{-1} N] - c^T = [c_B, c_B B^{-1} N] - [c_B, c_N] \leq 0 \Leftrightarrow y^T \bar{A} \leq c^T,$$

ou seja,  $y$  é uma solução dual admissível.

Supondo que  $\exists r \in \{1, \dots, m\}$  tal que  $\bar{b}_r < 0$ , então  $B^{-1}b$  não é uma solução primal admissível, pelo que podemos escolher, para linha pivotal, a  $r$ -ésima e proceder à escolha da coluna pivotal, de tal forma que a admissibilidade dual se mantenha (ou seja, de tal forma que, após a actualização do quadro, os coeficientes da linha de custos reduzidos se mantenham não positivos) e a função objectivo do problema dual seja incrementada. Logo, supondo que a coluna pivotal escolhida é a  $k$ -ésima (e, por conseguinte, o elemento escolhido para "pivot" é o coeficiente não nulo  $y_{rk}$ ), para que o valor corrente da função objectivo dual,  $\zeta = c_B B^{-1}b$ , seja incrementado é necessário (mas não suficiente) que  $y_{rk}$  seja negativo. Com efeito, supondo  $y_{rk} < 0$ , da actualização do quadro vem que

$$\zeta' = c_B B^{-1}b - \frac{z_k - c_{N_k} \bar{b}_r}{y_{rk}} \begin{cases} > \zeta & \text{se } z_k - c_{N_k} < 0 \\ = \zeta & \text{se } z_k - c_{N_k} = 0 \end{cases},$$

donde, em qualquer caso, se obtém  $\zeta' \geq \zeta$ , uma vez que  $\frac{z_k - c_{N_k}}{y_{rk}} \geq 0$  e  $\bar{b}_r < 0$ .

Com esta escolha de  $y_{rk} < 0$  para "pivot", o termo independente do novo quadro,  $\bar{b}_{r'} = \frac{\bar{b}_r}{y_{rk}}$ , vem positivo e os coeficientes da linha de custos reduzidos ( $z'_j - c_{N_j}$ ), para  $j = 1, \dots, p$  vêm actualizados do seguinte modo:

$$z'_j - c_{N_j} = \begin{cases} z_j - c_{N_j} - \frac{z_k - c_{N_k}}{y_{rk}} y_{rj}, & \text{se } j \neq k \\ - \frac{z_k - c_{N_k}}{y_{rk}}, & \text{se } j = k \end{cases}.$$

Consequentemente, vem que  $z'_k - c_{N_k} \leq 0$  e, para  $j \neq k$ ,

$$z_j - c_{N_j} - \frac{z_k - c_{N_k}}{y_{rk}} y_{rj} \leq 0 \Leftrightarrow \frac{z_j - c_{N_j}}{y_{rj}} \geq \frac{z_k - c_{N_k}}{y_{rk}} \quad (4.15)$$

se  $y_{rj} < 0$ . Tendo em conta (4.15) e a implicação

$$y_{rj} \geq 0 \Rightarrow z'_j - c_{N_j} = z_j - c_{N_j} - \frac{z_k - c_{N_k}}{y_{rk}} y_{rj} \leq 0,$$

a escolha da coluna pivotal deve ser feita determinando

$$\frac{z_k - c_{N_k}}{y_{rk}} = \min_{1 \leq j \leq p} \left\{ \frac{z_j - c_{N_j}}{y_{rj}} : y_{rj} < 0 \right\}$$

e, nestas condições, a  $k$ -ésima coluna do quadro reduzido é escolhida para pivotal sem que se destrua a admissibilidade dual da base associada ao próximo quadro.

Após a escolha do “pivot”, as operações de pivotação conducentes à actualização do quadro devem fazer-se como habitualmente.

Tendo em conta que as soluções duais obtidas são também básicas, desde que se previna a eventual entrada em ciclo nos casos degenerados (em que  $z_k - c_{N_k} = 0$ ), ao fim de um número finito de quadros, ou se conclui que o problema não tem óptimo finito, ou se conclui que não é possível incrementar a função objectivo de  $(D)$ , i.e.,  $\{i : 1 \leq i \leq m \text{ e } \bar{b}_i < 0\} = \emptyset$ . Esta última situação traduz, precisamente, a admissibilidade primal.

Se, para um dado quadro simplex,  $\bar{b}_r < 0$  e  $\{j : 1 \leq j \leq p \text{ e } y_{rj} < 0\} = \emptyset$  então, pode concluir-se que o primal é inconsistente. Com efeito, neste caso, sendo  $B$  a submatriz básica corrente e supondo que o sistema  $\bar{A}\bar{x} = b$  é equivalente ao sistema  $[B, N]\bar{x} = b$ , multiplicando ambos os membros por  $B^{-1}$  obtém-se  $B^{-1}\bar{A}\bar{x} = B^{-1}b \Leftrightarrow [I, B^{-1}N]\bar{x} = \bar{b}$ . Logo, considerando a  $r$ -ésima equação deste sistema e tendo em conta que os coeficientes que lhe estão associados são os coeficientes da  $r$ -ésima linha do respectivo quadro simplex, vem

$$x_r + \sum_{j=1}^p y_{rj}x_{m+j} = \bar{b}_r \Leftrightarrow x_r = \bar{b}_r - \sum_{j=1}^p y_{rj}x_{m+j}. \quad (4.16)$$

Da equação (4.16), tendo em conta que  $y_{rj} \geq 0 \forall j \in \{1, \dots, p\}$  e  $\bar{b}_r < 0$ , conclui-se que não existem valores não negativos para  $x_{m+j}$  ( $\forall j \in \{1, \dots, p\}$ ), com os quais se obtenha  $x_r \geq 0$ , pelo que  $(P)$  é inconsistente (e, conseqüentemente, pelo teorema da dualidade forte,  $(D)$  não tem óptimo finito).

Considerando para matriz básica inicial a submatriz constituída pelas  $m$  primeiras colunas de  $\bar{A}$ , a determinação de uma solução dual admissível inicial pode fazer-se acrescentando, ao conjunto de restrições do problema original, a restrição:

$$\sum_{j=m+1}^n x_j \leq M \Leftrightarrow x_0 + \sum_{j=m+1}^n x_j = M, \quad (4.17)$$

onde  $M$  é um escalar positivo suficientemente grande e  $x_0$  é uma variável de desvio (i.e.,  $x_0 \geq 0$ ). Posteriormente, a partir do quadro a seguir indicado, para se obter uma solução dual admissível, basta proceder à sua actualização (com as respectivas operações de pivotação) escolhendo para “pivot” o elemento determinado pela primeira linha e pela  $k$ -ésima coluna, desde que

$$z_{m+k} - c_{m+k} = \max_{m+1 \leq j \leq n} \{z_j - c_j\},$$

donde (no pressuposto de que a base corrente não define uma solução dual admissível) se conclui que  $z_{m+k} - c_{m+k} > 0$ .

	$x_{m+1}$	...	$x_{m+k}$	...	$x_n$	
$x_0$	1	...	1	...	1	M
$x_1$	$y_{11}$	...	$y_{1k}$	...	$y_{1(n-m)}$	$\bar{b}_1$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$x_r$	$y_{r1}$	...	$y_{rk}$	...	$y_{r(n-m)}$	$\bar{b}_r$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$x_m$	$y_{m1}$	...	$y_{mk}$	...	$y_{m(n-m)}$	$\bar{b}_m$
	$z_1 - c_{m+1}$	...	$z_{m+k} - c_{m+k}$	...	$z_n - c_n$	$c_B B^{-1} b$

Nestas condições, no quadro actualizado, vem

$$\begin{cases} z'_k - c_{m+k} = & - (z_k - c_{m+k}) < 0 \\ z'_j - c_{m+j} = z_j - c_{N_j} - (z_k - c_{m+k}) \leq 0, \text{ para } j \neq k \end{cases} .$$

Com esta modificação do problema original ( $P$ ), i.e., com a introdução da restrição (4.17), pode chegar-se a um dos seguintes casos:

1. o problema transformado é inconsistente;
2. obtém-se uma solução óptima,  $x^*$ , para a qual  $x_0^* = 0$  (ou seja, para a qual a variável de desvio associada à restrição introduzida tem valor nulo) e  $z_0 - c_0 < 0$  (i.e., o coeficiente que lhe está associado na linha de custos reduzidos é negativo);
3. obtém-se uma solução óptima,  $x^*$ , para a qual se verifica uma das situações:
  - (a)  $x_0^* > 0$ ;
  - (b)  $x_0^* = 0$  e  $z_0 - c_0 = 0$ .

Em 1 pode concluir-se que o problema original é inconsistente (uma vez que  $M$  é arbitrariamente grande).

Em 2 pode concluir-se que ( $P$ ) não tem óptimo finito, uma vez que  $x_0$  é não básica (e, conseqüentemente, a restrição introduzida é activa para valores de  $M$  arbitrariamente grandes) e qualquer deslocamento no sentido em que se viola a não negatividade da variável de desvio  $x_0$ , faz decrescer a função objectivo de ( $P$ )<sup>1</sup>.

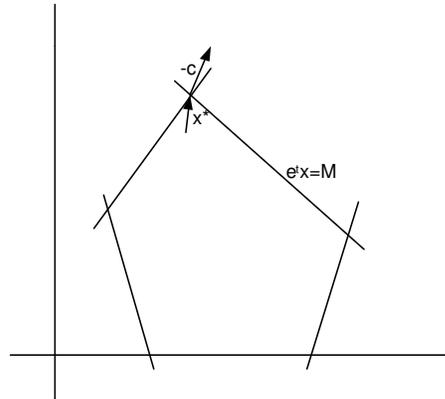


Figura 4.1: Exemplificação do caso descrito em 2.

Finalmente em 3 pode concluir-se que a solução encontrada é óptima para ( $P$ ), uma vez que em 3a ( $x_0^* > 0$ ) a restrição introduzida é redundante numa vizinhança da solução óptima encontrada e em 3b ( $x_0^* = 0$  e  $z_0 - c_0 = 0$ ) as variações em  $x_0$  não produzem variações na função objectivo.

Segue-se a descrição algorítmica do método dual simplex.

- Algoritmo para o método dual simplex

<sup>1</sup>Logo deslocando o hiperplano que define a restrição introduzida (aumentando  $M$ ), a função objectivo pode decrescer tanto quanto se queira.

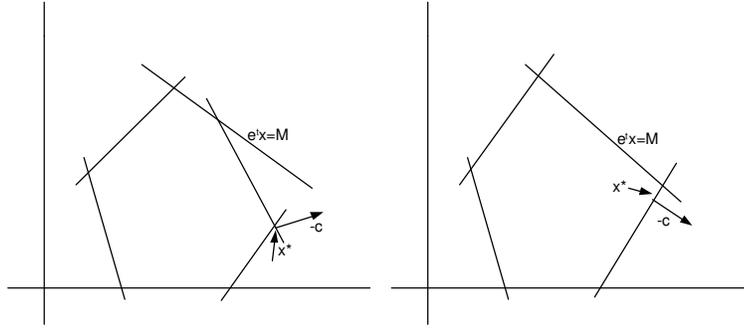


Figura 4.2: Exemplificação dos casos 3a e 3b (supondo-se em 3b que o gradiente da f.o. é ortogonal ao vector de componentes unitárias  $e$ ).

1. Determinar uma base  $B$  tal que  $c_B B^{-1} N - c_N \leq 0$  e passar a 2;
2. Se  $B^{-1} b \geq 0$  então **STOP**, a base corrente é ótima. Caso contrário determinar  $\bar{b}_r = \min\{\bar{b}_i : 1 \leq i \leq m\}$  e passar a 3;
3. Se  $\{y_{rj} : j \in J_N \text{ e } y_{rj} < 0\} = \emptyset$  (onde  $J_N$  denota o conjunto de índices não básicos correntes) então **STOP**,  $(P)$  é inconsistente. Caso contrário determinar

$$\frac{z_k - c_k}{y_{rk}} = \min_{j \in J_N} \left\{ \frac{z_j - c_j}{y_{rj}} : y_{rj} < 0 \right\}$$

e passar a 4;

4. Escolher  $y_{rk}$  para “pivot”, actualizar o quadro e voltar ao passo 2;

- Fim do algoritmo.

As regras a utilizar para a prevenção de entrada em ciclo para o método dual simplex, com ligeiras adaptações, são idênticas às utilizadas no método simplex (primal). No caso da regra de Bland, a sua adaptação ao método dual simplex é imediata, escolhendo-se primeiramente o menor dos índices linha, de entre os índices das linhas candidatas a pivotais e, posteriormente, o índice coluna, fazendo-se o desempate com a escolha do menor dos índices de entre os candidatos. A prova de que a versão da regra de Bland (para o método dual simplex) previne a entrada em ciclo pode ser consultada em (Schrijver, 1986).

Segue-se um exemplo de aplicação do método dual simplex.

**Exemplo 4.1.** *Vamos resolver o programa linear  $(P)$ :*

$$\begin{array}{rcll} \min & 2x_1 & + & 7x_2 & + & 6x_3 & + & 5x_4 \\ \text{s. a} & 2x_1 & - & 3x_2 & - & 5x_3 & + & 4x_4 & \geq & 20, \\ & 7x_1 & + & 2x_2 & + & 6x_3 & - & 2x_4 & \leq & 35, \\ & 4x_1 & + & 5x_2 & - & 3x_3 & - & 2x_4 & \geq & 15, \\ & x_1, & & x_2, & & x_3, & & x_4 & \geq & 0, \end{array}$$

*aplicando o método dual simplex.*



de ser óptima, qual o procedimento a adoptar para (com o menor esforço computacional possível) se atingir, de novo, a optimalidade. Este tipo de análise designa-se, habitualmente, por análise pós-óptimal (ou análise de sensibilidade).

Dado o programa linear  $PL(A, b, c)$ , as perturbações nos dados do problemas dividem-se nas seguintes:

1. Modificação dos coeficientes do gradiente da função objectivo.
2. Modificação dos termos independentes.
3. Modificação das entradas da matriz de decisão.<sup>2</sup>
4. Adição de uma nova actividade.<sup>3</sup>
5. Adição de uma nova restrição.

Para qualquer destas perturbações vamos supor conhecido o quadro simplex óptimo que a seguir se indica.

$$\begin{array}{c|c|c} & x_N & \\ \hline x_B & B^{-1}N & B^{-1}b \\ \hline & c_B B^{-1}N - c_N & c_B B^{-1}b \end{array}$$

### 4.3.1 Modificação dos coeficientes da função objectivo

Supondo que o vector  $c^T$  foi perturbado pela adição de  $\Delta c^T = [\Delta c_1, \dots, \Delta c_n]$ , i.e.,

$$c^T \leftarrow [c_1 + \Delta c_1, c_2 + \Delta c_2, \dots, c_n + \Delta c_n],$$

no quadro simplex óptimo, apenas se alteram os coeficientes da linha de custos reduzidos e o valor corrente da função objectivo, ou seja,

$$\begin{aligned} c_B B^{-1}N - c_N &\leftarrow c_B B^{-1}N - c_N + \Delta c_B B^{-1}N - \Delta c_N \\ c_B B^{-1}b &\leftarrow c_B B^{-1}b + \Delta c_B B^{-1}b. \end{aligned}$$

Consequentemente, se as desigualdades

$$c_B B^{-1}N - c_N + \Delta c_B B^{-1}N - \Delta c_N \leq 0$$

são satisfeitas, então a solução básica corrente permanece óptima. Para tal basta que qualquer que seja o índice  $j$  não básico se verifiquem as desigualdades

$$\Delta c_B B^{-1}a_j - \Delta c_j \leq c_j - c_B B^{-1}a_j = -(z_j - c_j).$$

No caso de alguma das desigualdades anteriores não se verificar, devemos prosseguir com o método simplex (primal) até se obter uma nova solução óptima para o problema  $PL(A, b, c + \Delta c)$ .

**Exemplo 4.2.** *Como exemplo de aplicação, vamos considerar o programa linear:*

$$\max\{c^T x : Ax \leq b, x \geq 0\},$$

cujo quadro simplex reduzido óptimo (no qual  $x_4$  e  $x_5$  são variáveis de desvio) é o que se indica

<sup>2</sup>Usualmente designadas por coeficientes tecnológicos.

<sup>3</sup>A qual envolve a consideração de uma nova variável, de uma nova coluna na matriz  $A$  e de um novo coeficiente no gradiente da função objectivo.

	$x_2$	$x_3$	$x_4$	
$x_1$	-2	-1	1	8
$x_5$	3	-1	1	12
	3	3	2	16

e analisar a perturbação do gradiente da função objectivo  $c^T$ , o qual (conforme decorre do quadro) é tal que  $c^T = [2, -7, -5]$ , pela adição do vector  $\Delta c^T = [-1, 2, 2]$ , i.e., de tal forma que  $c^T \leftarrow [1, -5, -3]$ .

**Solução.** Após a correspondente modificação do quadro simplex, obtém-se:

	$x_2$	$x_3$	$x_4$	
$x_1$	-2	-1	1	8
$x_5$	3	-1	1	12
	3	2	1	8

Uma vez que as condições de optimalidade se mantêm, podemos concluir que a solução básica óptima para o problema não perturbado se mantém óptima para o problema perturbado. ■

Esta análise pós-optimal (ou de sensibilidade) a este tipo de perturbação é usualmente utilizada na resolução de problemas de programação paramétrica, os quais se distinguem dos problemas de programação linear pelo facto das componentes do gradiente da função objectivo dependerem de um determinado parâmetro.

**Exemplo 4.3.** Como exemplo de aplicação, vamos resolver o problema de programação paramétrica

$$\begin{array}{llll} \min & (1 - \lambda)x_1 & + & (2 + \lambda)x_2 & + & (1 + \lambda)x_3 \\ \text{s. a} & x_1 & + & x_2 & + & x_3 \geq 20, \\ & 2x_1 & + & 2x_2 & + & x_3 \geq 30, \\ & x_1, & & x_2, & & x_3 \geq 0. \end{array}$$

**Solução.** Determinando-se o quadro simplex reduzido associado à solução básica  $(10, 0, 10)$ , obtém-se:

	$x_2$	$x_4$	$x_5$	
$x_1$	1	1	-1	10
$x_3$	0	-2	1	10
	$-1 - 2\lambda$	$-(1 + 3\lambda)$	$2\lambda$	20

Este quadro é óptimo, desde que  $\lambda$  satisfaça o seguinte sistema de desigualdades:

$$\begin{cases} -1 - 2\lambda \leq 0 \\ -(1 + 3\lambda) \leq 0 \\ 2\lambda \leq 0 \end{cases} \Leftrightarrow -\frac{1}{3} \leq \lambda \leq 0.$$

Para valores de  $\lambda \notin ]-\frac{1}{3}, 0[$ , deverá proceder-se às respectivas operações de pivotação, obtendo-se a sequência de quadros que a seguir se apresentam.

- $\lambda < -\frac{1}{3}$ ;

1°	$x_2$	$x_4$	$x_5$	
$x_1$	1	1	-1	10
$x_3$	0	-2	1	10
	$-1 - 2\lambda$	$-1 - 3\lambda$	$2\lambda$	20

2°	$x_2$	$x_1$	$x_5$	
$x_4$	1	1	-1	10
$x_3$	2	2	-1	30
	$\lambda$	$1 + 3\lambda$	$-1 - \lambda$	$30(1 + \lambda)$

Dado que  $\lambda < -\frac{1}{3}$ , no coeficiente da linha de custos reduzidos relativo a  $x_5$ , obtém-se  $-1 - \lambda \leq 0$ , para  $-1 \leq \lambda \leq -\frac{1}{3}$ , e  $-1 - \lambda > 0$ , para  $\lambda < -1$ . Nesta condições pode concluir-se que  $(0, 0, 30)$  é uma solução óptima para  $-1 \leq \lambda \leq -\frac{1}{3}$  e que o problema não tem óptimo finito para  $\lambda < -1$ .

- $\lambda > 0$ ;

1°	$x_2$	$x_4$	$x_5$	
$x_1$	1	1	-1	10
$x_3$	0	-2	1	10
	$-1 - 2\lambda$	$-1 - 3\lambda$	$2\lambda$	20
2°	$x_2$	$x_4$	$x_1$	
$x_5$	1	-1	1	20
$x_3$	0	-2	1	10
	$-1 - 2\lambda$	$-1 + \lambda$	$-2\lambda$	$20(1 - \lambda)$

A partir do último quadro obtido, concluí-se que  $(0, 0, 10)$  é solução óptima para  $0 < \lambda \leq 1$ .

Denotando por  $\mu(\lambda)$  a função que corresponde ao valor óptimo do programa paramétrico considerado, vem

$$\mu(\lambda) = \begin{cases} -\infty, & \text{se } \lambda < -1 \\ 30(1 + \lambda), & \text{se } -1 \leq \lambda \leq -\frac{1}{3} \\ 20, & \text{se } -\frac{1}{3} \leq \lambda \leq 0 \\ 20(1 - \lambda), & \text{se } 0 \leq \lambda \leq 1 \\ -\infty, & \text{se } \lambda > 1 \end{cases}$$



### 4.3.2 Modificação dos termos independentes

Supondo que o vector de termos independentes  $b$  é perturbado pelo vector  $\Delta b$ , i.e.,  $b \leftarrow b + \Delta b$ , o quadro simplex que lhe corresponde vem modificado da seguinte forma:

	$x_N$	
$x_B$	$B^{-1}N$	$B^{-1}b + B^{-1}\Delta b$
	$c_B B^{-1}N - c_N$	$c_B B^{-1}b + c_B B^{-1}\Delta b$

Consequentemente, o quadro simplex mantém-se óptimo sse as componentes do vector de termos independentes permanecerem não negativas, i.e.,

$$B^{-1}b + B^{-1}\Delta b \geq 0 \Leftrightarrow B^{-1}\Delta b \geq -B^{-1}b.$$

No caso de pelo menos uma das componentes passar a negativa, então a determinação da nova solução básica óptima do problema  $PL(A, b + \Delta b, c)$  pode fazer-se aplicando o método dual simplex (uma vez que as condições de optimalidade não são alteradas).

**Exemplo 4.4.** Voltando ao quadro simplex reduzido óptimo do Exemplo 4.2, do qual decorre que a inversa da base corrente é a matriz  $B^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$  e que  $b = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$  é o vector de termos independentes iniciais, vamos analisar a perturbação de  $b$  pela adição do vector  $\Delta b = \begin{bmatrix} -9 \\ 1 \end{bmatrix}$ .

**Solução.** O quadro simplex modificado pela perturbação vem dado por

	$x_2$	$x_3$	$x_4$	
$x_1$	-2	-1	1	-1
$x_5$	3	-1	1	4
	3	3	2	-2

Uma vez que as condições de optimalidade da linha de custos reduzidos se mantêm, aplica-se o método dual simplex, com o qual se obtêm a sequência de quadros simplex que a seguir se indica. Note-se que, neste caso, uma vez que o problema é de maximização, a escolha da coluna pivotal é determinada calculando  $\min_{2 \leq j \leq 4} \left\{ \frac{z_j - c_j}{-y_{1j}} : y_{1j} < 0 \right\} = \min \left\{ \frac{3}{2}, 3 \right\} = \frac{3}{2} = \frac{z_2 - c_2}{-y_{12}}$ .

1º	$x_2$	$x_3$	$x_4$	
$x_1$	-2	-1	1	-1
$x_5$	3	-1	1	4
	3	3	2	-2
2º	$x_1$	$x_3$	$x_4$	
$x_2$	-1/2	1/2	-1/2	1/2
$x_5$	3/2	-5/2	5/2	5/2
	3/2	3/2	7/2	-9/2

■

### 4.3.3 Modificação das entradas da matriz de decisão

Este tipo de perturbação pode dividir-se nos seguintes subcasos: 1– modificação de uma coluna não básica; 2– modificação de uma coluna básica.

#### 1. Modificação de uma coluna não básica.

Suponha que a coluna não básica  $a_j$  é perturbada pela adição do vector coluna  $\Delta a_j$ , ou seja,  $a_j \leftarrow a_j + \Delta a_j$ , então, no quadro, apenas se modificam a coluna  $B^{-1}a_j$  (que passa a ser  $B^{-1}a_j + B^{-1}\Delta a_j$ ) e o coeficiente da linha de custos reduzidos  $c_B B^{-1}a_j - c_j$  (que passa a ser  $c_B B^{-1}a_j - c_j + c_B B^{-1}\Delta a_j$ ). Nestas condições, desde que  $c_B B^{-1}a_j - c_j + c_B B^{-1}\Delta a_j$  se mantenha não positivo, i.e., desde que se verifique a desigualdade

$$c_B B^{-1}a_j - c_j \leq -c_B B^{-1}\Delta a_j,$$

o quadro corrente mantém-se óptimo.

#### 2. Modificação de uma coluna básica.

Se a coluna  $a_j$  perturbada é correntemente uma coluna básica, então é preferível acrescentar uma nova actividade  $x'_j$ , bem como a correspondente coluna  $a'_j = a_j + \Delta a_j$  e, posteriormente, após o cálculo de  $B^{-1}a'_j$  e  $c_B B^{-1}a'_j - c'_j$  (com

$c'_j = c_j$ ), estender o quadro simplex reduzido a esta nova variável, acrescentando a coluna

$$\begin{bmatrix} y'_j \\ z'_j - c'_j \end{bmatrix} = \begin{bmatrix} B^{-1}a'_j \\ c_B B^{-1}a'_j - c'_j \end{bmatrix}.$$

- Se  $y'_{jj} \neq 0$ , então deve escolher-se este elemento para “pivot” e proceder às operações de pivotação de modo a retirar da base a variável  $x_j$ . Uma vez que  $x_j$  não esteja na base, devemos ignorá-la (retirando-a do quadro) e
  - prosseguir com o método simplex (primal), se as condições de admissibilidade primal se mantiverem,
  - prosseguir com o método dual simplex, se as condições de admissibilidade dual se mantiverem

ou introduzindo variáveis artificiais no caso de nenhuma das situações se verificar.

- Se  $y'_{jj} = 0$ , então pode concluir-se que a nova coluna  $a'_j$  não pode substituir a coluna  $a_j$  na base (uma vez que, em tal situação, as demais não formam uma base juntamente com  $a'_j$ ). Nesta última situação, ou se retira  $x_j$  da base substituindo-a por qualquer outra (se tal for possível) ou se considera esta variável como artificial e se lhe associa um peso que prejudica a sua permanência na base (utilizando uma técnica do tipo “big-M”).

**Exemplo 4.5.** Vamos analisar o tipo de perturbação 1, considerando que na matriz de decisão do Exemplo 4.2, à coluna associada a  $x_2$  se adiciona o vector de perturbação  $\Delta a_2 = \begin{bmatrix} 4 \\ -4 \end{bmatrix}$ .

**Solução.** Nestas condições obtêm-se

$$B^{-1}a_2 \leftarrow B^{-1}a_2 + B^{-1}\Delta a_2 = \begin{bmatrix} 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ -4 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

e

$$\begin{aligned} c_B B^{-1}a_2 - c_2 &\leftarrow c_B B^{-1}a_2 - c_2 + c_B B^{-1}\Delta a_2 \\ &= 3 + [2, 0] \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ -4 \end{bmatrix} \\ &= 11. \end{aligned}$$

■

**Exemplo 4.6.** Vamos analisar o tipo de perturbação 2, supondo que no Exemplo 4.2 a coluna de  $A$  associada a  $x_1$ ,  $a_1$ , (para a qual se obtém  $B^{-1}a_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ ) é perturbada pela adição do vector  $\Delta a_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ .

**Solução.** Introduzindo a variável  $x_6$ , à qual se associa a coluna  $a_6 = a_1 + \Delta a_1$  e o coeficiente do gradiente da função objectivo  $c_6 = c_1 = 2$ , obtêm-se as igualdades

$$\begin{aligned} B^{-1}a_6 &= B^{-1}a_1 + B^{-1}\Delta a_1 \\ &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \end{aligned}$$

e

$$c_B B^{-1} a_6 - c_6 = [2, 0] \begin{bmatrix} 2 \\ 2 \end{bmatrix} - 2 = 2.$$

Como consequência, determinado o quadro simplex reduzido

	$x_2$	$x_3$	$x_4$	$x_6$	
$x_1$	-2	-1	1	2	8
$x_5$	3	-1	1	2	12
	3	3	2	2	16

e procedendo às operações de pivotação, conducentes a tirar da base a variável  $x_1$ , (que posteriormente é ignorada), obtém-se o quadro simplex reduzido óptimo.

	$x_2$	$x_3$	$x_4$	
$x_6$	-1	-1/2	1/2	4
$x_5$	5	0	0	4
	5	4	1	8

■

#### 4.3.4 Adição de uma nova actividade

Supondo que se pretende acrescentar uma nova actividade,  $x_{n+1}$ , à qual está associada a coluna de coeficientes tecnológicos  $a_{n+1}$  e o custo unitário  $c_{n+1}$ , deve acrescentar-se ao quadro a coluna

$$\begin{bmatrix} y_{p+1} \\ z_{p+1} - c_{n+1} \end{bmatrix} = \begin{bmatrix} B^{-1} a_{n+1} \\ c_B B^{-1} a_{n+1} - c_{n+1} \end{bmatrix}.$$

Se  $z_{n+1} - c_{n+1} \leq 0$ , então a solução corrente com  $x_{n+1}^* = 0$  continua óptima.<sup>4</sup>

Se  $z_{n+1} - c_{n+1} > 0$ , então  $x_{n+1}$  é candidata a entrar para a base e o método simplex (primal) deve prosseguir.

**Exemplo 4.7.** Voltando ao Exemplo 4.2, vamos analisar a perturbação provocada ao acrescentar-se uma nova variável ( $x_6$ ) cuja coluna, na matriz dos coeficientes tecnológicos  $A$ , é  $a_6 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$  e a respectiva componente do gradiente da função objectivo é  $c_6 = 4$ .

**Solução.** Tendo em conta que

$$\begin{aligned} B^{-1} a_6 &= \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \\ &= \begin{bmatrix} 1 \\ 3 \end{bmatrix} \\ c_B B^{-1} a_6 - c_6 &= c_B B^{-1} \begin{bmatrix} 1 \\ 2 \end{bmatrix} - 4 \\ &= [2, 0] \begin{bmatrix} 1 \\ 2 \end{bmatrix} - 4 \\ &= -2, \end{aligned}$$

<sup>4</sup>Note-se que embora o exemplo que temos vindo a considerar seja um problema de maximização, o estudo mais abstracto que estamos a fazer diz respeito a um problema de minimização.

a partir do quadro simplex reduzido, apresentado no Exemplo 4.2, determina-se o quadro simplex (para o problema modificado).

	$x_2$	$x_3$	$x_4$	$x_6$	
$x_1$	2	1	1	1	8
$x_5$	3	-1	1	3	12
	3	3	2	-2	16

Procedendo às respectivas operações de pivotação, com vista à determinação da nova solução óptima, determina-se o quadro óptimo

	$x_2$	$x_3$	$x_4$	$x_5$	
$x_1$	1	4/3	2/3	-1/3	4
$x_6$	1	-1/3	1/3	1/3	4
	5	7/3	8/3	2/3	24

do qual decorre a solução óptima  $x^* = (4, 0, 0, 0, 0, 4)^T$ . ■

### 4.3.5 Adição de uma nova restrição

A adição de uma nova restrição é, precisamente, a operação dual da adição de uma nova actividade.

Sendo  $a^{m+1}x \leq b_{m+1} \Leftrightarrow a^{m+1}x + x_{n+1} = b_{m+1}$ , com  $x_{n+1} \geq 0$ , a restrição acrescentada, onde  $a^{m+1}$  denota o vector linha dos respectivos coeficientes tecnológicos, as matrizes básica e não básica passam a ser

$$B' = \begin{bmatrix} B & 0 \\ a_B^{m+1} & 1 \end{bmatrix} \text{ e } N' = \begin{bmatrix} N \\ a_N^{m+1} \end{bmatrix},$$

onde  $a_B^{m+1}$  e  $a_N^{m+1}$  correspondem aos subvectores de  $a^{m+1}$  associados, respectivamente, às variáveis  $x_B$  e  $x_N$ .

Tendo em conta a forma de  $B'$ , facilmente se conclui que a sua inversa corresponde à matriz

$$B'^{-1} = \begin{bmatrix} B^{-1} & 0 \\ -a_B^{m+1}B^{-1} & 1 \end{bmatrix},$$

donde vem que

$$\begin{aligned} B'^{-1}N' &= \begin{bmatrix} B^{-1} & 0 \\ -a_B^{m+1}B^{-1} & 1 \end{bmatrix} \begin{bmatrix} N \\ a_N^{m+1} \end{bmatrix} \\ &= \begin{bmatrix} B^{-1}N \\ -a_B^{m+1}B^{-1}N + a_N^{m+1} \end{bmatrix}, \\ B'^{-1} \begin{bmatrix} b \\ b_{m+1} \end{bmatrix} &= \begin{bmatrix} B^{-1} & 0 \\ -a_B^{m+1}B^{-1} & 1 \end{bmatrix} \begin{bmatrix} b \\ b_{m+1} \end{bmatrix} \\ &= \begin{bmatrix} B^{-1}b \\ b_{m+1} - a_B^{m+1}B^{-1}b \end{bmatrix}, \\ c_{B'}B'^{-1} \begin{bmatrix} b \\ b_{m+1} \end{bmatrix} &= [c_B, 0] \begin{bmatrix} B^{-1}b \\ b_{m+1} - a_B^{m+1}B^{-1}b \end{bmatrix} \\ &= c_B B^{-1}b, \\ c_{B'}B'^{-1}N' - c_{N'} &= [c_B, 0] \begin{bmatrix} B^{-1}N \\ -a_B^{m+1}B^{-1}N + a_N^{m+1} \end{bmatrix} - c_N \\ &= c_B B^{-1}N - c_N. \end{aligned}$$

Como consequência, obtém-se o quadro simplex

	$x_N$	
$x_B$	$B^{-1}N$	$B^{-1}b$
$x_{n+1}$	$a_N^{m+1} - a_B^{m+1}B^{-1}N$	$b_{m+1} - a_B^{m+1}B^{-1}b$
	$c_B B^{-1}N - c_N$	$c_B B^{-1}b$

do qual se conclui, relativamente ao original, que a única modificação susceptível de pôr em causa a optimalidade da solução básica corrente diz respeito a

$$b_{m+1} - a^{m+1}B^{-1}b.$$

- Se  $b_{m+1} - a^{m+1}B^{-1}b \geq 0$ <sup>5</sup>, então a solução básica corrente é ótima.
- Se  $b_{m+1} - a^{m+1}B^{-1}b < 0$ , então recomenda-se a utilização do método dual simplex, com vista à obtenção da primal admissibilidade e, como consequência, de uma solução ótima para o problema modificado.

**Exemplo 4.8.** *Mais uma vez, recorrendo ao Exemplo 4.2, vamos analisar as alterações provocadas no quadro simplex quando se acrescenta a restrição*

$$x_1 + x_2 + x_3 \leq 3 \Leftrightarrow x_1 + x_2 + x_3 + x_6 = 3,$$

com  $x_6 \geq 0$ , e utilizar as técnicas pós-optimais recomendadas.

**Solução.** Tendo conta que  $b_3 = 3$  e  $a^3 = [1, 1, 1, 0, 0]$ , vem que

$$a_N^3 - a_B^3 B^{-1}N = [1, 1, 0] - [1, 0] \begin{bmatrix} -2 & -1 & 1 \\ 3 & -1 & 1 \end{bmatrix} = [3, 2, -1]$$

e ainda que  $b_3 - a_B^3 B^{-1}b = 3 - [1, 0] \begin{bmatrix} 8 \\ 12 \end{bmatrix} = 5$ .

Como consequência, obtém-se (para o novo problema) o quadro simplex reduzido a seguir indicado e, por aplicação do método dual simplex, o respectivo quadro ótimo.

1°	$x_2$	$x_3$	$x_4$	
$x_1$	-2	-1	1	8
$x_5$	3	-1	1	12
$x_6$	3	2	-1	-5
	3	3	2	16

2°	$x_2$	$x_3$	$x_6$	
$x_1$	1	1	1	3
$x_5$	6	1	1	7
$x_4$	-3	-2	-1	5
	9	7	2	6

Este tipo de análise pós-optimal (relativamente à adição de restrições) pode ser utilizada na resolução de problemas de programação inteira com uma técnica que usualmente se designa técnica dos cortes (fraccionais) de Gomory. ■

<sup>5</sup>Nestas condições, a solução básica ótima (anteriormente obtida) não viola a nova restrição, o que implica que se mantenha ótima para o problema modificado (porquê?).

## 4.4 Cortes de Gomory em programação inteira

Muitos problemas práticos são representados por modelos de programação linear com a restrição adicional de que as variáveis apenas podem tomar valores inteiros. Uma tentativa de resolver o problema de otimização

$$(P) \quad \min\{c^T x : Ax = b, x \geq 0, x \in \mathbb{Z}^n\},$$

consiste em ignorar a restrição que impõe que as variáveis só podem tomar valores inteiros, na expectativa de que a solução ótima a encontrar (para este problema relaxado) satisfaça essa restrição. No caso de tal não acontecer, supondo que o quadro ótimo é o que a seguir se apresenta,

	$x_{N_1}$	...	$x_{N_k}$	...	$x_{N_p}$	
$x_{B_1}$	$y_{11}$	...	$y_{1k}$	...	$y_{1p}$	$\bar{b}_1$
$\vdots$	$\vdots$	$\ddots$	...	$\ddots$	...	...
$x_{B_r}$	$y_{r1}$	...	$y_{rk}$	...	$y_{rp}$	$\bar{b}_r$
...	...	$\ddots$	...	$\ddots$	...	...
$x_{B_m}$	$y_{m1}$	...	$y_{mk}$	...	$y_{mp}$	$\bar{b}_m$
	$z_1 - c_{N_1}$	...	$z_k - c_{N_k}$	...	$z_p - c_{N_p}$	$c_B B^{-1} b$

sabe-se que pelo menos um dos termos independentes não é inteiro e que  $Ax = b$  é equivalente aos sistema de equações

$$x_{B_i} + \sum_{j=1}^p y_{ij} x_{N_j} = \bar{b}_i, \quad i = 1, \dots, m,$$

onde os coeficientes  $y_{ij}$  são os que constam no quadro. Denotando por  $[\alpha]$  o maior inteiro não superior a  $\alpha$  (ou seja,  $[\alpha]$  denota a parte inteira de  $\alpha$ ), supondo que  $\bar{b}_r$  não é inteiro (i.e., que a variável básica  $x_{B_r}$  não tem um valor inteiro) e reescrevendo a  $r$ -ésima equação, tal como se indica

$$\begin{aligned} x_{B_r} + \sum_{j=1}^p ([y_{rj}] + (y_{rj} - [y_{rj}])) x_{N_j} &= [\bar{b}_r] + (\bar{b}_r - [\bar{b}_r]) \\ &\Downarrow \\ x_{B_r} + \sum_{j=1}^p [y_{rj}] x_{N_j} - [\bar{b}_r] &= (\bar{b}_r - [\bar{b}_r]) - \sum_{j=1}^p (y_{rj} - [y_{rj}]) x_{N_j}, \end{aligned}$$

podemos concluir que o novo termo independente é inteiro, qualquer que seja a solução admissível com componentes inteiras.

Porém, uma vez que o termo independente

$$(\bar{b}_r - [\bar{b}_r]) - \sum_{j=1}^p (y_{rj} - [y_{rj}]) x_{N_j}$$

é inferior a 1 (dado que  $\bar{b}_r - [\bar{b}_r] < 1$ ,  $y_{rj} - [y_{rj}] \geq 0$  e  $x_{N_j} \geq 0$  para  $j = 1, \dots, p$ ), as soluções admissíveis (com componentes inteiras) devem satisfazer a nova restrição:

$$\begin{aligned} (\bar{b}_r - [\bar{b}_r]) - \sum_{j=1}^p (y_{rj} - [y_{rj}]) x_{N_j} &\leq 0 \\ &\Downarrow \\ -\sum_{j=1}^p (y_{rj} - [y_{rj}]) x_{N_j} &\leq -(\bar{b}_r - [\bar{b}_r]). \end{aligned}$$

Desta forma podemos concluir que a nova restrição corta a região admissível do programa relaxado, pelo que a sua adição ao conjunto de restrições iniciais a vai “apertar” e, eventualmente, alterar a solução óptima corrente. Por aplicação da técnica pós-otimal anteriormente referida (associada à adição de uma nova restrição) pode facilmente determinar-se uma solução óptima (para o problema relaxado) que satisfaça a restrição que acabamos de introduzir.

A aplicação sistemática deste procedimento designa-se por técnica dos cortes (fracionais) de Gomory e conduz-nos à obtenção de uma solução óptima para o problema de optimização (P).<sup>6</sup>

**Exemplo 4.9.** *Vamos resolver o problema de optimização:*

$$\min\{x_1 + x_2 : x_1 + 2x_2 \geq 3, 3x_1 + x_2 \geq 3, x_1, x_2 \geq 0 \text{ e } x_1, x_2 \in \mathbb{Z}\},$$

*utilizando a técnica dos cortes de Gomory.*

**Solução.** Resolvendo o problema relaxado, a partir da solução básica admissível inicial  $(x_1, x_2) = (3, 0)$ , obtém-se:

1°	$x_2$	$x_3$	
$x_1$	2	-1	3
$x_4$	5	-3	6
	1	-1	3

2°	$x_4$	$x_3$	
$x_1$	-2/5	1/5	3/5
$x_2$	1/5	-3/5	6/5
	-1/5	-2/5	9/5

Uma vez que a solução óptima encontrada não tem componentes inteiras, vamos seleccionar uma variável não inteira (por exemplo  $x_1$ ) e gerar o corte associado à equação do quadro

$$x_1 - \frac{2}{5}x_4 + \frac{1}{5}x_3 = \frac{3}{5} \Leftrightarrow x_1 - (0 + \frac{2}{5})x_4 + (0 + \frac{1}{5})x_3 = (0 + \frac{3}{5}) \Leftrightarrow x_1 = \frac{3}{5} + \frac{2}{5}x_4 - \frac{1}{5}x_3.$$

Como consequência obtém-se o corte

$$\frac{3}{5} + \frac{2}{5}x_4 - \frac{1}{5}x_3 \leq 0 \Leftrightarrow \frac{2}{5}x_4 - \frac{1}{5}x_3 + x_5 = -\frac{3}{5}, x_5 \geq 0.$$

Adicionando ao problema inicial a nova restrição, obtém-se o quadro

	$x_N$	
$x_B$	$B^{-1}N$	$B^{-1}b$
$x_5$	$a_N^3 - a_B^3 B^{-1}N$	$b_3 - a_B^3 B^{-1}b$
	$c_B B^{-1}N - c_N$	$c_B B^{-1}b$

onde  $a_N^3 = [\frac{2}{5}, -\frac{1}{5}]$ ,  $a_B^3 = [0, 0]$  e  $b_3 = -\frac{3}{5}$ , pelo que se obtém

$$a_N^3 - a_B^3 B^{-1}N = [\frac{2}{5}, -\frac{1}{5}] \text{ e } b_3 - a_B^3 B^{-1}b = -\frac{3}{5}.$$

Nestas condições, por aplicação do método dual simplex, obtém-se os quadros:

3°	$x_4$	$x_3$	
$x_1$	-2/5	1/5	3/5
$x_2$	1/5	-3/5	6/5
$x_5$	2/5	-1/5	-3/5
	-1/5	-2/5	9/5

4°	$x_4$	$x_5$	
$x_1$	0	1	0
$x_2$	-1	-3	3
$x_3$	-2	-5	3
	-1	-2	3

A partir do quadro simplex óptimo obtido (o 4°), conclui-se que  $(x_1, x_2) = (0, 3)$  é uma solução óptima para o problema inicial.

<sup>6</sup>Estes problemas são usualmente designados, na literatura, por problemas de programação linear inteira.

## 4.5 Exercícios

**Exercício 4.1.** Escreva os duais dos seguintes problemas de programação linear:

1.  $\max\{c^T x : A'x = b', A''x \geq b'', x \geq 0\}$ ;
2.  $\min\{c^T x : Ax \leq b\}$ ;
3.  $\min\{y^T b : y^T A \leq c^T, y \geq 0\}$ ;
4.  $\max\{c'^T x + c''^T x'' : A'x' + A''x'' \leq b, x' \geq 0\}$ .

**Exercício 4.2.** Prove que o dual do dual de um programa linear é o primal.

**Exercício 4.3.** Considerando o programa linear primal na forma canônica

$$\min\{c^T x : Ax \geq b, x \geq 0\},$$

prove que se existe um par de soluções  $x^*$  e  $y^*$  admissíveis, respectivamente, para o primal e para o dual, satisfazendo a complementaridade de “slacks”, i.e.,  $y^{*T}(Ax^* - b) = (c^T - y^{*T}A)x^* = 0$ , então estas soluções são ótimas, respectivamente, para o primal e para o dual.

**Exercício 4.4.** Considere o programa linear  $(P) \min\{c^T x : Ax = b, x \in X\}$ , que o conjunto  $X = \{x \in \mathbb{R}^n : Bx \geq d, x \geq 0\}$  é um politopo e que  $\{x \in \mathbb{R}^n : Ax = b, x \in X\} \neq \emptyset$ .

1. Prove que o programa

$$\max_w \{w^T b + \min\{(c^T - w^T A)x : x \in X\}\}$$

equivale ao dual de  $(P)$ <sup>7</sup>.

2. Denotando por  $x^1, \dots, x^p$ , os vértices de  $X$ , prove que o problema determinado na alínea anterior se pode transformar no problema<sup>8</sup>

$$\begin{aligned} & \max z \\ \text{s. a} \quad & z \leq w^T b + (c^T - w^T A)x^j, \quad \text{para } j = 1, \dots, p. \\ & z, w \text{ não restringidos.} \end{aligned}$$

**Exercício 4.5.** Considere o programa linear  $(P)$ :

$$\begin{aligned} \min \quad & 2x_1 + 2x_2 \\ \text{s. a} \quad & -2x_1 + x_2 \geq 1 \\ & x_1 - 2x_2 \geq 1 \\ & x_1, x_2 \geq 0 \end{aligned}$$

1. Determine o quadro simplex associado à solução básica não admissível  $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} =$

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

2. Formule o dual de  $(P)$ .
3. Aplicando o dual simplex ao quadro obtido em 1 verifique (justificando) se o dual de  $(P)$  tem solução ótima finita.

<sup>7</sup>Sugestão: construa o dual de  $(P)$ ,  $(D)$ , nas variáveis  $w$  e  $y$  e, para cada  $w$  (fixo), considere o dual de  $(D)$ .

<sup>8</sup>Este tipo de decomposição é conhecida por decomposição de Benders.

**Exercício 4.6.** Considere o programa linear  $\min\{c^T x : Ax \leq b, x \geq 0\}$ , onde  $c^T = [-2, 1, -1]$  e o quadro simplex reduzido óptimo que lhe está associado (no qual  $x_4$  e  $x_5$  são variáveis de desvio)

	$x_2$	$x_3$	$x_4$	
$x_1$	1	1	1	6
$x_5$	2	1	1	10
	-3	-1	-2	-12

1. Sabendo que pretende modificar  $c$ , fazendo

$$c^T = [-2, 1, -1] + \Delta[1, 1, 2],$$

determine o intervalo de valores de  $\Delta$  para os quais a solução básica corrente se mantém óptima.

2. Sabendo que a coluna de  $A$  associada a  $x_2$ ,  $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ , passou para  $\begin{bmatrix} 2 \\ 5 \end{bmatrix}$ , verifique se as condições de optimalidade associadas à solução básica corrente foram alteradas.

3. Considerando que ao problema original se acrescentou a restrição

$$x_1 + 4x_2 \leq 7,$$

determine a respectiva solução óptima para o problema modificado.

**Exercício 4.7.** Considere o programa linear (P):

$$\begin{array}{ll} \min & 5x_1 + 2x_2 + 4x_3 \\ \text{s. a.} & 3x_1 + x_2 + 2x_3 \geq 5 \\ & 6x_1 + 3x_2 + 5x_3 \geq 10 \\ & x_1, \quad x_2, \quad x_3 \geq 0. \end{array}$$

- Escreva o problema dual (D) e resolva-o, aplicando o método simplex.
- A partir da solução óptima obtida para (D), determine uma solução óptima para (P), utilizando o teorema da complementaridade entre as soluções óptimas do dual e do primal.

## Parte II

# Técnicas e Modelos de Otimização Combinatória



## Capítulo 5

# Redes e Grafos

A origem da *teoria dos grafos* é, em geral, associada ao problema das sete pontes de Königsberg (cidade da Prússia Oriental que agora se designa por Kaliningrad). Parte desta cidade localizava-se em duas ilhas do rio Pregel ligadas entre si por uma ponte e às margens por 6 pontes. Uma das ilhas estava ligada por duas pontes a cada uma das margens do rio e a outra tinha uma ponte para cada margem. Consta que os habitantes gostavam de dar passeios de modo a atravessar todas as pontes e que alguns andavam particularmente aborrecidos pelo facto de não encontrarem um trajecto (com partida e chegada a um mesmo lugar) que lhes permitisse atravessar apenas uma vez cada uma das pontes. O matemático suíço Leonhard Euler (1707-1783) tomou conhecimento deste problema e resolveu-o (indicando a impossibilidade de existência de um tal percurso), numa memória que publicou em San Petersburgo em 1736. Problemas com ingredientes semelhantes foram posteriormente abordados por outros matemáticos como é o caso do matemático irlandês Sir William Rowan Hamilton (1805-1865) que, por sua vez, propôs um problema que designou por *viagem à volta do mundo* que consiste em percorrer todos os vértices de um dodecaedro (poliedro regular com 20 vértices de grau 3 e faces pentagonais) passando uma única vez em cada vértice e com partida e chegada no mesmo vértice. Desde então, muitos problemas de natureza combinatória têm sido modelados por grafos com abordagens que, por sua vez, têm gerado novos conceitos e resultados combinatórios de aplicação generalizada, transformando a teoria dos grafos numa área da matemática contemporânea de intensa investigação.

Este capítulo inicia-se com o estudo dos conceitos e resultados fundamentais da teoria dos grafos, seguindo-se os problemas métricos em grafos, o estudo dos grafos bipartidos, árvores e florestas e o estudo dos conjuntos independentes, cliques e colorações.

### 5.1 Conceitos e resultados fundamentais

Um *grafo*  $G$  é um par de conjuntos  $(V, E)$ , tal que  $V = V(G) = \{v_1, \dots, v_n\}$  é o conjunto dos *vértices* e  $E = E(G)$  é o conjunto das *arestas*, a cada uma das quais corresponde um subconjunto de  $V(G)$  de cardinalidade 2, i.e.,  $E(G) = \{e_1, \dots, e_m\}$ , com  $e_k = \{v_{k_i}, v_{k_j}\}$ , para  $k \in \{1, \dots, m\}$ . Um grafo diz-se *simples* se não existem arestas paralelas (mais do que uma aresta entre os mesmos vértices) nem lacetes (arestas com ambos os extremos no mesmo vértice). No caso dos grafos simples, por simplicidade

de notação, uma aresta entre os vértices  $x$  e  $y$  será representada por  $xy$ .

Um *digrafo* ou *grafo orientado*  $\vec{G}$  é um par de conjuntos  $(V, A)$ , tal que  $V = V(\vec{G}) = \{x_1, \dots, x_n\}$  é o conjunto dos *vértices* e  $A = A(\vec{G}) \subset V \times V$  é o conjunto dos *arcos* (conjunto de pares ordenados de vértices). Dado um arco  $a = (x, y) \in A(\vec{G})$ , o vértice  $x$  designa-se *cauda* de  $a$  e denota-se  $cau(a)$  e o vértice  $y$  designa-se por *cabeça* de  $a$  e denota-se por  $cab(a)$ . Por exemplo, no caso do digrafo representado na Figura 5.1, considerando o arco  $a_4 = (x, z)$ ,  $cau(a_4) = x$  e  $cab(a_4) = z$ .

Um digrafo diz-se *simples* se não tem nem arcos múltiplos (mais do que um arco entre o mesmo par de vértices) nem arcos com ambos os extremos no mesmo vértice.

Neste texto, os grafos (digrafos) simples são designados apenas por *grafos* (*digrafos*) e os grafos (digrafos) com lacetes e/ou arestas paralelas por *multigrafos* (*multidigrafos*). O número de vértices e arestas de um grafo (digrafo) designa-se, respectivamente, por *ordem* e *dimensão* do grafo (digrafo).

Os grafos (digrafos) são muitas vezes representados por figuras planas constituídas por linhas (linhas orientadas) e pontos, as primeiras representando arestas (arcos) e os segundos vértices. Como exemplo, na Figura 5.1 representa-se o grafo  $G$  de ordem 6 e dimensão 7 e o digrafo  $\vec{H}$  de ordem 5 e dimensão 6.

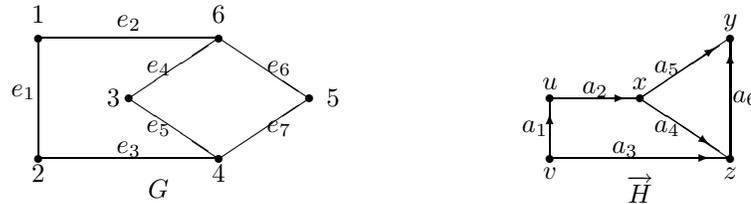


Figura 5.1: Exemplos de grafo e digrafo.

Dado um grafo (digrafo)  $G(\vec{G})$ , uma aresta (arco) diz-se incidente no vértice  $v$ , se  $v$  é um dos seus extremos (a cauda ou a cabeça do arco, no caso dos digrafos) e dois vértices,  $x$  e  $y$ , dizem-se *adjacentes* se

$$xy \in E(G) \quad (\{(x, y), (y, x)\} \cap A(\vec{G}) \neq \emptyset).$$

De agora em diante, sempre que os conceitos a introduzir sejam válidos tanto para grafos como digrafos, optaremos pela sua abordagem recorrendo, unicamente, à notação dos grafos (fazendo-se a respectiva passagem para digrafos, substituindo aresta por arco).

Dados dois grafos  $G$  e  $G'$ , diz-se que  $G'$  é um subgrafo de  $G$  e que  $G$  é um supergrafo de  $G'$  quando  $V(G') \subseteq V(G)$  e  $E(G') \subseteq E(G)$ . Por sua vez, designa-se por subgrafo de  $G$  *induzido* pelo subconjunto de vértices  $V'$  e denota-se por  $G[V']$  o subgrafo obtido de  $G$  ignorando o subconjunto de vértices  $V(G) \setminus V'$  e, conseqüentemente, as arestas que lhe são incidentes. Por exemplo, na Figura 5.2 representa-se o subgrafo do grafo  $G$ , representado na Figura 5.1, induzido pelo subconjunto de vértices  $\{3, 4, 5, 6\}$ , ou seja,  $G[\{3, 4, 5, 6\}]$ .

Designa-se por grafo *completo* (*nulo*) de ordem  $n$  e denota-se por  $K_n$  ( $N_n$ ) um grafo com  $n$  vértices dois a dois adjacentes (não adjacentes, ou seja, sem qualquer aresta). Por sua vez, designa-se por grafo *complementar* de um grafo  $G$  e denota-se por  $\bar{G}$  um

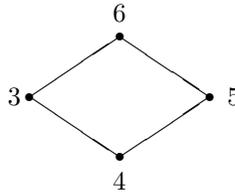


Figura 5.2: Subgrafo  $G[\{3, 4, 5, 6\}]$  do grafo representado na Figura 5.1.

grafo com o mesmo conjunto de vértices de  $G$  no qual dois vértices são adjacentes se e só se não são adjacentes em  $G$ .

Denotando por  $\mathcal{G}_n$  o conjunto de todos os grafos de ordem  $n$  e por  $\subseteq_P$  a relação binária definida em  $\mathcal{G}_n$  tal que  $G_1 \subseteq_P G_2$  se e só se  $G_1$  é um subgrafo de  $G_2$ , é imediato concluir que  $P = (\mathcal{G}_n, \subseteq_P)$  é um conjunto parcialmente ordenado. Neste conjunto parcialmente ordenado,  $K_n$  é o único elemento maximal e  $N_n$  o único elemento minimal. Adicionalmente, dado  $G \in \mathcal{G}_n \setminus \{N_n, K_n\}$ , é claro que  $G$  e  $\bar{G}$  não são comparáveis.

Dado um grafo  $G$  e um vértice  $v \in V(G)$  designa-se por *grau* ou *valência* de  $v$  e denota-se por  $d_G(v)$  o número de arestas de  $G$  incidentes em  $v$ . O máximo grau dos vértices de  $G$  denota-se por  $\Delta(G)$  e o mínimo grau por  $\delta(G)$ . Um grafo diz-se  $p$ -regular se todos os seus vértices têm grau  $p$  (no caso de  $p = 3$  estes grafos também se designam por *cúbicos*).

Tendo em conta que ao adicionarmos os graus de todos os vértices de um grafo arbitrário,  $G$ , cada aresta conta duas vezes, com facilidade se conclui que

$$\sum_{v \in V(G)} d_G(v) = 2|E(G)| \quad (5.1)$$

e, conseqüentemente, que  $\sum_{v \in V(G)} d_G(v) = 0 \pmod{2}$ . Da igualdade (5.1) decorre ainda que

$$\delta(G) \leq \lfloor \frac{2|E(G)|}{n} \rfloor \leq \Delta(G). \quad (5.2)$$

No caso de um digrafo,  $\vec{G}$ , podemos separar o grau de um vértice  $v \in V(\vec{G})$  em semi-grau incidente

$$d_{\vec{G}}^-(v) = |\{(x, v) \in A(\vec{G})\}|$$

e semi-grau emergente

$$d_{\vec{G}}^+(v) = |\{(v, x) \in A(\vec{G})\}|,$$

pelo que  $d_{\vec{G}}(v) = d_{\vec{G}}^-(v) + d_{\vec{G}}^+(v)$ , sendo também imediato concluir que

$$\sum_{v \in V(\vec{G})} d_{\vec{G}}^+(v) = \sum_{v \in V(\vec{G})} d_{\vec{G}}^-(v).$$

Como exemplo, o vértice  $x$  do digrafo  $\vec{H}$  representado na Figura 5.1 é tal que  $d_{\vec{G}}^-(x) = 1$  e  $d_{\vec{G}}^+(x) = 2$ .

O conjunto de vértices adjacentes a um vértice  $v$  designa-se por *vizinhança* de  $v$  (ou conjunto de vizinhos de  $v$ ) e denota-se por  $N_G(v)$ . Como consequência, é claro que  $d_G(v) = |N_G(v)|$ .

Dados dois grafos  $G$  e  $H$ , a *união* de  $G$  com  $H$  denota-se por  $G \cup H$  e corresponde ao grafo  $G \cup H = (V(G) \cup V(H), E(G) \cup E(H))$  e a *intersecção* de  $G$  com  $H$  denota-se por  $G \cap H$  e corresponde ao grafo

$$G \cap H = (V(G) \cap V(H), E(G) \cap E(H)).$$

Com base nesta definição de união e intersecção, podemos afirmar que um grafo  $G$  se parte nos grafos  $G_1 = (V_1, E_1), \dots, G_k = (V_k, E_k)$  se  $G = \bigcup_{1 \leq j \leq k} G_j$  e  $G_p \cap G_q = \emptyset \quad \forall p \neq q$ , onde  $\emptyset$  denota o grafo definido pelo par  $(\emptyset, \emptyset)$ .

Um grafo  $G$  diz-se *conexo* se não admite qualquer partição para além da trivial (i.e.,  $G = G \cup \emptyset$ ). Por sua vez, dado um grafo não conexo  $G$ , um seu subgrafo diz-se uma *componente conexa* (ou simplesmente *componente*) de  $G$ , se é um subgrafo conexo maximal, no sentido em que sendo induzido pelo subconjunto de vértices  $V'$ ,  $\forall x \in V(G) \setminus V'$ , o subgrafo  $G[V' \cup \{x\}]$  não é conexo.

Designa-se por *passeio* num grafo  $G$ , entre os vértices  $x$  e  $y$ , toda a sequência de vértices e arestas da forma

$$x = v_1, v_1v_2, v_2, \dots, v_{k-1}, v_{k-1}v_k, v_k = y,$$

com eventual repetição de vértices e arestas. Neste caso, os vértices  $x$  e  $y$  designam-se por vértices extremos do passeio ( $x$  por vértice inicial e  $y$  por vértice final). Um *trajecto* num grafo  $G$  entre os vértices  $x$  e  $y$  é um passeio entre  $x$  e  $y$  sem arestas repetidas (podendo, no entanto, existir vértices repetidos). Um *caminho* entre os vértices  $x$  e  $y$  é um trajecto entre  $x$  e  $y$  sem vértices repetidos. Os trajectos fechados (onde o vértice final coincide com o inicial) designam-se por *circuitos* e os trajectos fechados, onde os vértices inicial e final são os únicos que coincidem, designam-se por *ciclos*. No caso dos grafos simples, os caminhos, trajectos, passeios, circuitos e ciclos podem representar-se pela respectiva sequência de vértices.

A matriz de *adjacência* de um grafo  $G$  tal que  $V(G) = \{v_1, \dots, v_n\}$ , é uma matriz  $n \times n$ ,  $A_G$ , tal que

$$(A_G)_{ij} = \begin{cases} 1 & \text{se } v_i v_j \in E(G), \\ 0 & \text{caso contrário.} \end{cases}$$

Por sua vez, a matriz de *adjacência* de um digrafo  $\vec{G}$  de ordem  $n$  é uma matriz  $n \times n$ ,  $A_{\vec{G}}$ , tal que

$$(A_{\vec{G}})_{ij} = \begin{cases} 1 & \text{se } (v_i, v_j) \in A(DG), \\ 0 & \text{nos outros casos.} \end{cases}$$

**Exemplo 5.1.** Considerando o grafo  $G$  e o digrafo  $\vec{H}$ , representados na Figura 5.1, vamos determinar as respectivas matrizes de adjacência.

**Solução.**

$$A_G = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \end{pmatrix} \quad \text{e} \quad A_{\vec{H}} = \begin{pmatrix} 0 & -1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 & 1 \\ 0 & -1 & -1 & 0 & 1 \\ 0 & 0 & -1 & -1 & 0 \end{pmatrix}.$$

□

**Teorema 5.1.** *Dado um grafo  $G$  e a sua matriz de adjacência  $A_G$ , denotando as entradas da matriz  $A_G^k$  por  $a_{ij}^k$  e por  $p_{ij}(k)$  o número de passeios de comprimento  $k$  entre os vértices  $v_i$  e  $v_j$  de  $G$ , podemos concluir que  $a_{ij}^k = p_{ij}(k)$  (e nestas condições, note-se que as entradas da diagonal principal de  $A_G^k$  correspondem ao número de passeios fechados de comprimento  $k$  que começam e terminam no vértice  $v_i$ , ou seja,  $p_{ii}(k)$ ).*

**Demonstração.** Dado que o resultado é trivialmente verdadeiro para  $k = 0$  (admitindo que  $A_G^0 = I$ ) e para  $k = 1$ , vamos fazer a prova por indução sobre  $k$ . Assim, suponha que o resultado se verifica para  $k$ , com  $k \geq 1$ , e considere a matriz  $A_G^{k+1} = A_G^k A_G$ . Então, qualquer que seja  $v_i v_j \in E(G)$ ,

$$a_{ij}^{k+1} = \sum_{r=1}^n a_{ir}^k a_{rj} = \sum_{r=1}^n p_{ir}(k) a_{rj} = p_{ij}(k+1).$$

□

**Corolário 5.2.** *Um grafo  $G$  de ordem  $n$  é conexo se e somente se  $\sum_{k=0}^{n-1} A_G^k \geq J$ , onde  $J$  denota a matriz quadrada de ordem  $n$  com entradas unitárias.*

**Demonstração.** Supondo que  $G$  é conexo, então existe um caminho (de comprimento não superior a  $n - 1$ ) entre quaisquer pares de vértices, pelo que  $\forall ij \in E(G)$  existe  $k \leq n - 1$  tal que  $a_{ij}^k \geq 1$ . Consequentemente, podemos concluir a desigualdade  $\sum_{k=0}^{n-1} A_G^k \geq J$ .

Reciprocamente, supondo que a desigualdade  $\sum_{k=0}^{n-1} A_G^k \geq J$  se verifica. Então,  $\forall v_i v_j \in E(G)$  existe  $k \leq n - 1$  tal que  $a_{ij}^k \geq 1$  e, uma vez que  $p_{ij}^k = a_{ij}^k$ , podemos concluir que existe pelo menos um passeio de comprimento não superior a  $n - 1$  entre  $v_i$  e  $v_j$ . □

Denotando por  $t$  o número de triângulos, com base no Teorema 5.1, podemos concluir ainda o seguinte:

1.  $tr(A_G^1) = 0$ ,
2.  $tr(A_G^2) = 2|E(G)|$ ,
3.  $tr(A_G^3) = 6t$ ,

onde  $tr(A)$  denota o traço da matriz quadrada  $A$ , ou seja, a soma dos seus elementos diagonais (que, como se sabe, é igual à soma dos respectivos valores próprios).

**Lema 5.3.** *Qualquer que seja o valor próprio  $\lambda$  da matriz de adjacência  $A_G$ ,*

$$-\Delta(G) \leq \lambda \leq \Delta(G).$$

**Demonstração.** Seja  $\hat{u}$  o vector próprio (que vamos supor de norma unitária) associado ao valor próprio  $\lambda$  e  $\hat{u}_i$  a componente de  $\hat{u}$  com maior valor absoluto, então  $|\lambda| |\hat{u}_i| = |\lambda \hat{u}_i| = |\sum_{j \in N_G(i)} \hat{u}_j| \Leftrightarrow |\lambda| = |\sum_{j \in N_G(i)} \frac{\hat{u}_j}{|\hat{u}_i}||$ . Logo,

$$|\lambda| \leq \sum_{j \in N_G(i)} \left| \frac{\hat{u}_j}{\hat{u}_i} \right| \leq \Delta(G). \quad (5.3)$$

□

Pode acrescentar-se, ainda, que se o grafo  $G$  tem pelo menos uma aresta, então  $A_G$  tem valores próprios positivos e negativos. O menor valor próprio é não superior a  $-1$  e é igual a  $-1$  se e só se todas as componentes conexas de  $G$  são subgrafos completos. Se  $G$  é conexo, então o maior valor próprio tem multiplicidade um (ou seja, é simples) e o correspondente vector próprio tem componentes estritamente positivas.

A matriz de incidência aresta vértice de um grafo de ordem  $n$  e dimensão  $m$  é uma matriz  $n \times m$ ,  $B_G$ , tal que

$$(B_G)_{ij} = \begin{cases} 1 & \text{se } e_j = v_i v_k, \text{ para algum } v_k \in V(G), \\ 0 & \text{caso contrário.} \end{cases}$$

Por sua vez, a matriz de incidência arco vértice de um digrafo de ordem  $n$  e dimensão  $m$  é uma matriz  $n \times m$ ,  $B_{\vec{G}}$ , tal que

$$(B_{\vec{G}})_{ij} = \begin{cases} 1 & \text{se } a_j = (v_i, v_k), \text{ para algum } v_k \in V(\vec{G}), \\ -1 & \text{se } a_j = (v_k, v_i), \text{ para algum } v_k \in V(\vec{G}), \\ 0 & \text{se } a_j = (v_p, v_q), \text{ e } v_i \notin \{v_p, v_q\}. \end{cases}$$

**Exemplo 5.2.** Considerando o grafo  $G$  e o digrafo  $\vec{H}$ , representados na Figura 5.1, vamos determinar as respectivas matrizes de incidência.

**Solução.**

$$B_G = \begin{matrix} & e_1 & e_2 & e_3 & e_4 & e_5 & e_6 & e_7 \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} & \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{pmatrix} \end{matrix} \text{ e } B_{\vec{H}} = \begin{matrix} & a_1 & a_2 & a_3 & a_4 & a_5 & a_6 \\ \begin{matrix} u \\ v \\ x \\ y \\ z \end{matrix} & \begin{pmatrix} -1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & -1 & -1 & 0 & 1 \end{pmatrix} \end{matrix}$$

□

Dado um grafo  $G$ , designa-se por *grafo linha* de  $G$  e denota-se por  $L(G)$  o grafo que se obtém de  $G$ , considerando como vértices as arestas de  $G$  e como relação de adjacência entre os seus vértices a respectiva relação de adjacência entre arestas. Assim, dois vértices de  $L(G)$  são adjacentes se e só se as correspondentes arestas em  $G$  são adjacentes (ou seja, têm um vértice comum).

**Exemplo 5.3.** Vamos determinar a o grafo linha do grafo  $G$  representado na Figura 5.1.

**Solução.** Segue-se a representação de  $L(G)$ .

□

Como consequência da definição de grafo linha, com facilidade se conclui que, dado um grafo arbitrário  $G$  com  $m$  arestas,

$$B_G^T B_G = 2I_m + A_{L(G)},$$

onde  $I_m$  denota a matriz identidade de ordem  $m$ .

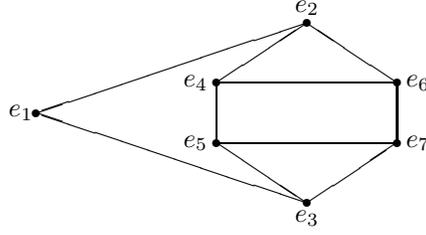


Figura 5.3: Grafo linha do grafo  $G$  representado na Figura 5.1.

**Teorema 5.4.** *Dado um grafo arbitrário  $G$ , com pelo menos uma aresta (caso contrário não faz sentido falar em grafo linha), os valores próprios de  $A_{L(G)}$  são não inferiores a  $-2$ . Adicionalmente,  $-2$  é o menor valor próprio de  $A_{L(G)}$  se e só se  $G$  tem uma componente com pelo menos um ciclo com um número par de arestas ou pelo menos dois ciclos com um número ímpar de arestas.*

**Demonstração.** Sendo  $B_G$  a matriz de incidência aresta vértice, vem

$$B_G^T B_G = 2I + A_{L(G)}.$$

Dado que  $\forall x \in \mathbb{R}^n \quad x^T B_G^T B_G x = \|B_G x\|^2 \geq 0$ , podemos concluir que  $2I + A_{L(G)}$  é semidefinida positiva e, conseqüentemente,  $\lambda_{\min}(A_{L(G)}) \geq -2$ . Supondo que  $G$  tem um ciclo comprimento par formado pelas arestas  $a_1, a_2, \dots, a_{2r}$ , aos quais correspondem as colunas de  $B_G$ ,  $\hat{b}_1, \hat{b}_2, \dots, \hat{b}_{2r}$ , considerando o vector  $\hat{u}$  definido por  $\hat{u} = -\hat{e}_{a_1} + (-1)^2 \hat{e}_{a_2} + \dots + (-1)^{2r-1} \hat{e}_{a_{2r-1}} + (-1)^{2r} \hat{e}_{a_{2r}}$ , obtém-se

$$B_G^T B_G \hat{u} = B_G^T \sum_{j=1}^{2r} (-1)^j \hat{b}_j = B_G^T \hat{0} = 0.$$

Note-se que sendo  $a_j = xy$  e  $a_{j+1} = yz$ ,

$$(-1)^j \hat{b}_j + (-1)^{j+1} \hat{b}_{j+1} = \begin{cases} \hat{e}_x + \hat{e}_y - \hat{e}_y - \hat{e}_z = \hat{e}_x - \hat{e}_z, & \text{se } j \text{ é par;} \\ -\hat{e}_x - \hat{e}_y + \hat{e}_y + \hat{e}_z = -\hat{e}_x + \hat{e}_z, & \text{caso contrário.} \end{cases}$$

Logo,  $\hat{u} \in \text{Ker}(2I + A_{L(G)})$  (onde  $\text{Ker}(A)$  denota o subespaço nulo da matriz  $A$ ) e, conseqüentemente,  $\hat{u}$  é vector próprio de  $A_{L(G)}$  associado ao valor próprio  $-2$ . Idêntica conclusão se obtém na presença de dois ciclos de comprimento ímpar, considerando o vector definido por um passeio fechado de comprimento par obtido percorrendo cada um dos ciclos, no caso de terem um vértice comum, ou cada um dos ciclos mais duas vezes um caminho que os ligue, no caso contrário.

Reciprocamente, se  $-2$  é valor próprio de  $A_{L(G)}$ , sendo  $\hat{u}$  um vector pertencente ao correspondente subespaço próprio, vem  $B_G^T B_G \hat{u} = 0$  e uma vez que as colunas de  $B_G^T$  são linearmente independentes, podemos concluir a igualdade  $B_G \hat{u} = 0$ . Logo o vector nulo é combinação linear não nula das colunas de  $B_G$ , o que corresponde à existência de um ciclo com um número par de arestas ou dois ciclos na mesma componente com um número ímpar de arestas.  $\square$

Designa-se por *matriz laplaciana* de  $G$  e denota-se por  $L_G$ , a matriz  $L_G = D_G - A_G$ , onde  $D_G$  denota a matriz diagonal  $\text{diag}(d_G(v_1), \dots, d_G(v_n))$ . Por sua vez, designa-se

por *matriz laplaciana sem sinal* e denota-se por  $|L_G|$ , a matriz  $|L_G| = D + A_G$ . É fácil concluir que  $B_G B_G^T = D_G + A_G$  e, atribuindo uma orientação arbitrária às arestas de  $G$ , obtendo-se  $\vec{G}$ ,  $B_{\vec{G}} B_{\vec{G}}^T = D_G - A_G$ .

Dado um grafo  $G$ , com  $cc$  componentes conexas, designa-se por *corte* de  $G$ , todo o subconjunto de arestas  $E' \subset E(G)$  tal que  $G - E'$  (subgrafo de  $G$  com o mesmo conjunto de vértices  $V(G)$ , mas cujo subconjunto de arestas é  $E(G) \setminus E'$ ) tem  $cc + 1$  componentes conexas. Adicionalmente, dados dois vértices de  $u, v \in V(G)$ , se existe um caminho em  $G$  entre  $u$  e  $v$  e não existe qualquer caminho entre  $u$  e  $v$  em  $G - E'$ , então  $E'$  diz-se um  $uv$ -corte.

Dado um subconjunto de vértices  $S \subset V(G)$ , denota-se por  $\partial(S)$  o corte definido pelas arestas com um extremo em  $S$  e outro extremo em  $V(G) \setminus S$ . Por exemplo, considerando o grafo  $G$  representado na Figura 5.1 e o subconjunto de vértices  $\{1, 2, 3\}$ ,  $\partial(\{1, 2, 3\}) = \{e_2, e_3, e_4, e_5\}$ . No caso de um digrafo,  $\vec{H}$ , dado um subconjunto de vértices  $S \subset V(\vec{G})$ , podemos dividir o corte  $\partial(S)$  no subconjunto de arcos com cauda em  $S$  e cabeça em  $V(G) \setminus S$  que designamos por *arcos para a frente* e denotamos por  $\partial^+(S)$  e subconjunto de arcos com cabeça em  $S$  e cauda em  $V(G) \setminus S$  que designamos por *arcos para trás* e denotamos por  $\partial^-(S)$ . Por exemplo, considerando o subconjunto de vértices  $S = \{u, x, y\}$  do digrafo  $\vec{H}$ , obtém-se  $\partial^+(S) = \{a_4\}$  e  $\partial^-(S) = \{a_1, a_6\}$  (é claro que  $\partial(S) = \partial^+(S) \cup \partial^-(S)$ ).

Se  $\hat{x}$  é o vector característico do subconjunto de vértices  $S \subset V(\vec{G})$ , ou seja,  $\hat{x} \in \{0, 1\}^n$  é tal que

$$\hat{x}_v = \begin{cases} 1, & \text{se } v \in S, \\ 0, & \text{no caso contrário,} \end{cases}$$

então  $\hat{x}^T B_{\vec{H}} = \hat{y}^T$  é um vector dos espaços dos arcos de  $\vec{H}$ , tal que

$$\hat{y}_a = \begin{cases} +1, & \text{se } a \in \partial^+(S); \\ -1, & \text{se } a \in \partial^-(S); \\ 0, & \text{nos restantes casos.} \end{cases}$$

**Exemplo 5.4.** Sendo  $\hat{x}$  o vector característico do subconjunto de vértices  $S = \{u, x, y\}$  do digrafo  $\vec{H}$  representado na Figura 5.1, vamos determinar  $\partial^+(S)$  e  $\partial^-(S)$ , a partir do vector  $\hat{x}^T B_{\vec{H}}$ .

**Solução.** Uma vez que  $\hat{x}^T = [1, 0, 1, 1, 0]$ , tendo em conta o Exemplo 5.2, obtém-se

$$\hat{x}^T B_{\vec{H}} = [1, 0, 1, 1, 0] \begin{pmatrix} -1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & -1 & -1 & 0 & 1 \end{pmatrix} = [-1, 0, 0, 1, 0, -1].$$

Logo,  $\partial^+(S) = \{a_4\}$  e  $\partial^-(S) = \{a_1, a_6\}$ . □

## 5.2 Problemas métricos em grafos

Dado um passeio  $P$  de um grafo  $G$ , designa-se por comprimento de  $P$  e denota-se por  $comp(P)$  o número de arestas que o constitui. Por exemplo, uma aresta é um caminho de comprimento 1 e um vértice é um caminho de comprimento 0. Por outro lado, um

triângulo é um ciclo de comprimento 3.

Dado um grafo  $G$  de ordem  $n$  e dois vértices  $x, y \in V(G)$ , denotando por  $\mathcal{P}_G(x, y)$  o conjunto de todos os caminhos de  $G$  entre  $x$  e  $y$ , designa-se por *distância* entre vértices de  $G$  a função

$$d_G : V(G) \times V(G) \mapsto \{0, \dots, n-1\} \cup \{\infty\} \quad (5.4)$$

$$(x, y) \rightsquigarrow d_G(x, y) = \begin{cases} \min\{\text{comp}(P) : P \in \mathcal{P}_G(x, y)\}, & \text{se } \mathcal{P}_G(x, y) \neq \emptyset, \\ \infty, & \text{se } \mathcal{P}_G(x, y) = \emptyset. \end{cases}$$

A partir das definições de grafo conexo e de componente conexa, podemos concluir que um grafo é conexo se e só se tem uma única componente ou, em alternativa, se e só se existe um caminho entre quaisquer dois dos seus vértices. A maior distância entre vértices de um grafo  $G$  designa-se por *diâmetro* de  $G$  e denota-se por  $\text{diam}(G)$ , por sua vez, designa-se por *cintura* de  $G$  e denota-se por  $g(G)$ , o comprimento do ciclo de menor comprimento contido em  $G$ . Quando  $G$  não é conexo, diz-se que tem diâmetro infinito e escreve-se  $\text{diam}(G) = \infty$  e quando é acíclico (i.e, não tem ciclos) diz-se que tem cintura infinita e escreve-se  $g(G) = \infty$ .

**Teorema 5.5.** *Dado um grafo  $G$ , se  $\delta(G) \geq 2$ , então  $G$  contém um caminho  $P$  e um ciclo  $C$  tais que  $\text{comp}(P) \geq \delta(G)$  e  $\text{comp}(C) \geq \delta(G) + 1$ .*

**Demonstração.** Seja  $P = v_0, \dots, v_k$  um caminho de maior comprimento em  $G$ . Então todos os vizinhos de  $v_k$  estão em  $P$  (caso contrário,  $P$  poderia ser estendido para um vizinho e não teria comprimento máximo). Logo,

$$\text{comp}(P) = k \geq d_G(v_k) \geq \delta(G).$$

Por outro lado, se  $v_i$  é o vértice de menor índice em  $P$  tal que  $v_i v_k \in E(G)$ , então  $v_i, v_{i+1}, \dots, v_k, v_i$  é um ciclo de comprimento pelo menos  $\delta(G) + 1$  (dado que  $\{v_i, v_{i+1}, \dots, v_k\}$  contém todos os vértices adjacentes a  $v_k$ , cujo número é não inferior a  $\delta(G)$ ).  $\square$

Um vértice  $v \in V(G)$  diz-se *central* em  $G$  se a maior das distâncias entre  $v$  e os restantes vértices é a menor possível. Esta distância designa-se por *raio* de  $G$  e denota-se por  $\text{raio}(G)$ , pelo que  $\text{raio}(G) = \min_{v \in V(G)} \max\{d_G(v, x) : x \in V(G)\}$ . Com facilidade se conclui que  $\text{raio}(G) \leq \text{diam}(G) \leq 2\text{raio}(G)$ .

### 5.2.1 Problemas de caminho mais curto

Para grafos com custos ou pesos nas arestas, o custo de um caminho é igual à soma dos custos ou pesos das suas arestas e o mesmo se passa no caso de grafos orientados com custos (ou pesos) nos arcos. Em certas aplicações, porém, os custos referem-se a distâncias entre vértices, pelo que, algumas vezes, designamos o custo de um caminho como sendo o seu comprimento. Assim, neste contexto, o caminho de custo mínimo ou o caminho mais curto não é necessariamente o que tem menor número de arestas (no caso dos grafos) ou arcos (no caso dos digrafos).

Por exemplo, no grafo representado na Figura 5.4, existem três caminhos entre os vértices  $v_1$  e  $v_2$ , cujos custos são os apresentados na Tabela 5.1. Conforme se pode verificar, de entre estes caminhos, o caminho de custo mínimo é o que tem maior número de arestas.

caminho	número de arestas	soma de pesos
$v_1v_2$	1	5
$v_1v_3v_2$	2	5
$v_1v_4v_3v_2$	3	4

Tabela 5.1: Comprimentos dos caminhos entre os vértices  $v_1$  e  $v_2$  do grafo da Figura 5.4.

Um grafo simples com custos nas arestas representa-se pelo terno  $G = (V, E, W)$ , onde  $W = (w_{ij})$  denota a matriz de custos. Nesta matriz, a entrada  $w_{ij}$  corresponde ao custo da aresta  $ij$ , se uma tal aresta existe, ou  $w_{ij} = \infty$  se  $ij \notin E(G)$ . Adicionalmente, assume-se que  $w_{ii} = 0$  para cada  $i$ . Deste modo, podemos lidar com o grafo como se tratasse de um grafo completo, no qual as arestas com custo infinito correspondem a arestas ausentes no grafo original. Por exemplo, o grafo representado na Figura 5.4 tem como matriz de custos a matriz:

$$W = \begin{pmatrix} 0 & 5 & 4 & 1 \\ 5 & 0 & 1 & \infty \\ 4 & 1 & 0 & 2 \\ 1 & \infty & 2 & 0 \end{pmatrix}.$$

Quando todos os custos das arestas são não negativos, o algoritmo que frequentemente mais se utiliza na determinação de um caminho de custo mínimo entre dois vértices  $s$  e  $t$  é o *algoritmo de Dijkstra*. A ideia básica deste algoritmo consiste em agrupar num conjunto, sucessivamente mais alargado, os vértices à menor distância de  $s$  até se incluir nesse conjunto o vértice  $t$ , marcando-se temporariamente os vértices que no passo corrente se consideram mais próximos de  $s$  e mudando-se a marca temporária de um vértice  $v$  para marca permanente quando se obtém o caminho mais curto entre  $s$  e  $v$ . Na descrição deste algoritmo vamos utilizar a seguinte notação:

- *Marca*[ $v$ ] – comprimento do caminho mais curto entre  $s$  e  $v$ , de entre os caminhos já determinados;
- *Antecessor*[ $v$ ] – antecessor do vértice  $v$  no caminho mais curto entre  $s$  e  $v$ , de entre os já determinados;
- *Temporários* – conjunto dos vértices com marca temporária;

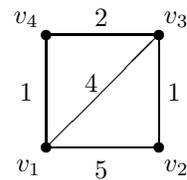


Figura 5.4: Grafo com custos nas arestas.

- $z$  – vértice com menor marca temporária corrente que passa a ser permanente.

### Descrição do algoritmo de Dijkstra.

1. Inicialização das variáveis. Marcar inicialmente o vértice  $s$  com a marca permanente 0 (uma vez que  $\text{dist}(s, s) = 0$ ) e marcar todos os restantes vértices com a marca temporária  $\infty$ , ou seja,  $\text{Marca}[s] \leftarrow 0$ ,  $\text{Marca}[v] \leftarrow \infty$ , para  $v \neq s$ ,  $\text{Temporários} \leftarrow V(G) \setminus \{s\}$  e  $z \leftarrow s$ .
2. Para cada vértice  $v$  que não tenha marca permanente, determinar uma nova marca, conforme se indica:

$$\text{Marca}[v] = \begin{cases} \text{Marca}[z] + w_{zv}, & \text{se } \text{Marca}[v] > \text{Marca}[z] + w_{zv}, \\ \text{Marca}[v], & \text{se } \text{Marca}[v] \leq \text{Marca}[z] + w_{zv}. \end{cases}$$

Observe-se que a desigualdade  $\text{Marca}[v] > \text{Marca}[z] + w_{zv}$ , significa que o caminho mais curto entre  $s$  e  $v$ , de entre os caminhos já determinados, passa pelo vértice  $z$ , ou seja, é o definido pelo  $(s, z)$ -caminho mais curto, ao qual se junta a aresta (ou arco)  $zv$ . Neste caso, o antecessor de  $v$  é  $z$ , ou seja,  $\text{Antecessor}[v] \leftarrow z$ .

3. Determinar um novo vértice  $z$  de marca mínima, de entre todos os vértices com marca temporária. Mudar a marca de  $z$  para marca permanente, ou seja,  $\text{Temporários} \leftarrow \text{Temporários} \setminus \{z\}$ .
4. Se  $z \neq t$ , então voltar ao passo 2.
5. FIM – o comprimento de um caminho mais curto entre  $s$  e  $t$  ( $\text{dist}(s, t)$ ) é igual a  $\text{Marca}[t]$  e este caminho é o definido pela sequência de vértices

$$(s, \dots, \text{Antecessor}[\text{Antecessor}[t]], \text{Antecessor}[t], t).$$

**Exemplo 5.5.** Utilizando algoritmo de Dijkstra, vamos determinar um caminho mais curto (e a respectiva distância) entre os vértices  $v_5$  e  $v_8$  do grafo definido pela matriz de distâncias:

$$W = \begin{matrix} & \begin{matrix} v_1 & v_2 & v_3 & v_4 & v_5 & v_6 & v_7 & v_8 \end{matrix} \\ \begin{matrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \\ v_8 \end{matrix} & \begin{pmatrix} 0 & 12 & \infty & \infty & 12 & \infty & \infty & \infty \\ 12 & 0 & 13 & \infty & 12 & 14 & 15 & \infty \\ \infty & 13 & 0 & 13 & \infty & \infty & 11 & 15 \\ \infty & \infty & 13 & 0 & \infty & \infty & \infty & 11 \\ 12 & 12 & \infty & \infty & 0 & 15 & \infty & \infty \\ \infty & 14 & \infty & \infty & 15 & 0 & 13 & \infty \\ \infty & 15 & 11 & \infty & \infty & 13 & 0 & 12 \\ \infty & \infty & 15 & 11 & \infty & \infty & 12 & 0 \end{pmatrix} \end{matrix}$$

**Solução.** A Tabela 5.2 apresenta os valores obtidos durante a aplicação do algoritmo de Dijkstra em cada um dos respectivos passos. Note-se que nesta tabela, para cada vértice  $v$ , em cada passo determinamos um par

$$(Marca[v], Antecessor[v]),$$

onde  $Marca[v]$  corresponde à distância corrente ao vértice inicial que aparece a negrito quando passa a permanente. Analisando a tabela, conclui-se que  $dist(v_5, v_8) = 39$  e que o  $(v_5, v_8)$ -caminho mais curto (que possivelmente não é único) é o determinado pela sequência de vértices  $v_5v_2v_7v_8$ .  $\square$

$v_5$	$v_1$	$v_2$	$v_3$	$v_4$	$v_6$	$v_7$	$v_8$
<b>(0, -)</b>	( $\infty, -$ )						
	<b>(12, <math>v_5</math>)</b>	(12, $v_5$ )	( $\infty, -$ )	( $\infty, -$ )	(15, $v_5$ )	( $\infty, -$ )	( $\infty, -$ )
		<b>(12, <math>v_5</math>)</b>	( $\infty, -$ )	( $\infty, -$ )	(15, $v_5$ )	( $\infty, -$ )	( $\infty, -$ )
			(25, $v_2$ )	( $\infty, -$ )	<b>(15, <math>v_5</math>)</b>	(27, $v_2$ )	( $\infty, -$ )
			<b>(25, <math>v_2</math>)</b>	( $\infty, -$ )		(27, $v_2$ )	( $\infty, -$ )
				(38, $v_3$ )		<b>(27, <math>v_2</math>)</b>	(40, $v_3$ )
				<b>(38, <math>v_3</math>)</b>			(39, $v_7$ )
							<b>(39, <math>v_7</math>)</b>

Tabela 5.2: Determinação de um caminho mais curto entre os vértices  $v_5$  e  $v_8$ , por aplicação do algoritmo de Dijkstra.

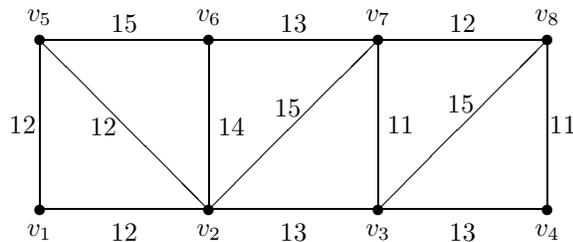


Figura 5.5: Grafo do Exemplo 5.5.

A determinação do caminho mais curto entre dois vértices pode ser estendida à determinação do ciclo de comprimento mínimo de um grafo, ou seja, à determinação da sua cintura  $g(G)$ . Com efeito, para qualquer aresta  $uv \in E(G)$  podemos determinar o caminho mais curto entre  $u$  e  $v$  no grafo  $G - \{u, v\}$  (caso exista). Juntando a este caminho a aresta  $uv$ , obtém-se um ciclo de  $G$  de menor comprimento de entre os que contêm  $uv$ . Assim, denotando por  $C_{uv}$  um ciclo de menor comprimento de entre os ciclos que contêm  $uv$ ,

$$g(G) = \min\{comp(C_{uv}) : uv \in E(G)\}.$$

### 5.3 Grafos bipartidos, árvores e florestas

Um grafo  $G$  diz-se *bipartido* se existe uma partição do seu conjunto de vértices em  $V'$  e  $V''$  tal que não existem arestas entre qualquer par de vértices de  $V'$  nem entre qualquer par de vértices de  $V''$ . Um tal grafo bipartido é usualmente representado por  $G = (V', V'', E)$ , onde  $E$  denota o respectivo conjunto de arestas. Quando  $|V'| = m$ ,  $|V''| = n$  e  $\forall x \in V' \forall y \in V''$ ,  $xy \in E(G)$  este grafo designa-se por *grafo bipartido completo* e denota-se por  $K_{mn}$ .

**Teorema 5.6.** *Um grafo é bipartido se e só se não tem circuitos de comprimento ímpar.*

**Demonstração.** Se  $G = (V', V'', E(G))$  é um grafo bipartido, então é claro que todos os circuitos têm comprimento par. Com efeito, uma vez que tanto em  $V'$  como em  $V''$  não existem vértices adjacentes, partindo-se, por exemplo, de um vértice em  $V'$ , de cada vez que se passa para  $V''$ , para se obter um circuito, tem de se voltar a  $V'$  na aresta seguinte, pelo que qualquer circuito tem comprimento par.

Reciprocamente, suponha que  $G$  não tem circuitos de comprimento ímpar. Uma vez que um grafo é bipartido se e só se cada uma das suas componentes constitui um subgrafo bipartido, podemos supor, sem perda de generalidade, que  $G$  é conexo. Considere um vértice arbitrário  $z \in V(G)$  e seja  $V' = \{w \in V(G) : d_G(z, w) \text{ é ímpar}\}$ . Nestas condições não existem arestas que liguem vértices de  $V'$  (caso contrário existiriam circuitos de comprimento ímpar). Por outro lado, como todos os vértices de  $V(G) \setminus V'$  estão a uma distância par de  $z$  (em particular  $z$  está a uma distância 0 dele próprio), não existem vértices adjacentes em  $V(G) \setminus V'$  (uma vez que, por razões idênticas às anteriores, em tais condições, existiriam circuitos de comprimento ímpar). Logo, fazendo  $V'' = V(G) \setminus V'$  obtém-se uma bipartição para  $G$ , dada por  $G = (V', V'', E(G))$ .  $\square$

Os grafos conexos acíclicos (i.e, sem ciclos) designam-se por *árvores* e constituem uma classe especial de grafos bipartidos. Por sua vez, designa-se por *floresta* todo o grafo acíclico, pelo que uma floresta é um grafo cujas componentes são árvores.

**Teorema 5.7.** *Sendo  $G$  um grafo, são equivalentes as seguintes proposições:*

1.  $G$  é uma árvore.
2.  $G$  é conexo e tem  $|V(G)| - 1$  arestas.
3.  $G$  não tem circuitos, mas acrescentando-se uma aresta a  $G$  resulta um único circuito.

**Demonstração.** Para provar a implicação  $1 \Rightarrow 2$ , dado que uma árvore é um grafo conexo, basta provar que se  $G$  é uma árvore, então tem  $|V(G)| - 1$  arestas, para o que vamos utilizar indução sobre o número de arestas de grafos que definem árvores. Seja  $G_k$  uma árvore com  $k$  arestas. Para  $k = 1$ ,  $|V(G_1)| = 2$ , pelo que o resultado se verifica. Suponha que o resultado é verdadeiro para  $k \geq 1$

e considere a árvore  $G_{k+1}$ . Dado que  $G_{k+1}$  não tem circuitos, existe pelo menos um vértice,  $v$ , com grau 1. Considerando o subgrafo obtido de  $G_{k+1}$ , retirando-se o vértice  $v$  (e consequentemente a aresta que lhe é incidente), determinamos um grafo  $G_k$  que continua a ser uma árvore (uma vez que permanece conexo e sem circuitos). Logo, por hipótese de indução,  $|E(G_k)| = |V(G_k)| - 1$ . Dado que  $|V(G_{k+1})| = |V(G_k)| + 1$  conclui-se que  $G_{k+1}$  tem  $|V(G_{k+1})| - 1$  arestas, completando-se assim a prova da implicação  $1 \Rightarrow 2$ . A prova das implicações  $2 \Rightarrow 3$  e  $3 \Rightarrow 1$ , fica como exercício.  $\square$

**Teorema 5.8.** *Um grafo  $G$  é uma floresta se e só se*

$$|E(G)| - |V(G)| + cc(G) = 0,$$

onde  $cc(G)$  denota o número de componentes de  $G$ .

**Demonstração.**  $\Rightarrow$  A prova da condição necessária vai ser feita por indução sobre o número de arestas de  $G$ , tendo em conta que o resultado se verifica trivialmente para  $|E(G)| = 0$ .

Suponha  $|E(G)| > 0$  e que o resultado se verifica para todas as florestas com menos do que  $|E(G)|$  arestas. Seja  $G'$  um subgrafo de  $G$  obtido por eliminação de uma aresta arbitrária. Logo  $G'$  é uma floresta com  $|E(G)| - 1$  arestas,  $|V(G)|$  vértices e  $cc(G) + 1$  componentes. Por hipótese de indução, aplicada a  $G'$ ,

$$0 = |E(G')| - |V(G')| + c(G') = |E(G)| - 1 - |V(G)| + cc(G) + 1 = |E(G)| - |V(G)| + cc(G).$$

$\Leftarrow$  Suponha que  $G$  tem  $p$  componentes,  $G_1, \dots, G_p$ , pelo que  $|E(G)| - |V(G)| + p = \sum_{j=1}^p (|E(G_j)| - |V(G_j)| + 1)$ . Então

$$|E(G)| - |V(G)| + p = 0 \Leftrightarrow \sum_{j=1}^p (|E(G_j)| - |V(G_j)| + 1) = 0$$

e, uma vez que  $\forall j \in \{1, \dots, p\} \quad |E(G_j)| - |V(G_j)| + 1 \geq 0$ , conclui-se que

$$\forall j \in \{1, \dots, p\} \quad |E(G_j)| - |V(G_j)| + 1 = 0.$$

Consequentemente, de acordo com o Teorema 5.7, todos os grafos  $G_j$ , com  $j \in \{1, \dots, p\}$ , são árvores.  $\square$

Deste teorema decorre que todo o grafo  $G$  tal que  $|E(G)| \geq |V(G)|$  contém pelo menos um circuito.

Dado um grafo conexo  $G$ , designa-se por árvore *abrangente* ou de *suporte* de  $G$  todo o subgrafo de  $G$  que é uma árvore e contém todos os vértices de  $G$ .

**Teorema 5.9.** *Todo o grafo conexo tem uma árvore abrangente.*

**Demonstração.** Seja  $G$  um grafo conexo. Se  $G$  não tem circuitos então, por definição, é uma árvore e o resultado verifica-se. Suponha que  $G$  tem um circuito. Então retirando uma aresta a esse circuito o grafo mantém-se conexo (porquê?). Repetindo este processo, ao fim de um número finito de arestas eliminadas, obtém-se uma árvore abrangente (uma vez que o conjunto de vértices não foi alterado).  $\square$

### 5.3.1 Árvores abrangentes de custo mínimo

No caso de grafos com custos nas arestas, o *custo de um subgrafo* corresponde ao custo determinado pela soma dos custos das respectivas arestas. Em muitas aplicações, é frequente a determinação de árvores abrangentes de custo mínimo. Por exemplo, supondo que se pretendem ligar  $n$  cidades por intermédio de uma rede de telecomunicações, e que se conhecem os custos de ligação de cada par de cidades, a determinação dos pares de cidades que se devem ligar com o menor custo, mas de modo que seja possível o estabelecimento de uma ligação entre quaisquer cidades, corresponde à determinação de uma árvore abrangente de custo mínimo. Os algoritmos mais populares para a determinação de árvores abrangentes de custo mínimo, são o algoritmo de Kruskal (que é um algoritmo guloso, *greedy* na terminologia inglesa) e o algoritmo de Prim (que escolhe o vizinho mais próximo).

#### Algoritmo de Kruskal

O passo básico do algoritmo de Kruskal, aplicado a um grafo  $G$  com custos nas arestas, consiste na escolha sucessiva de uma aresta com custo mínimo e na sua posterior eliminação, obtendo-se uma versão modificada do grafo original. Uma vez determinado um subconjunto de arestas  $S \subset E(G)$  com menor custo e que não formam um ciclo, determina-se uma aresta  $e$  de custo mínimo em  $E(G) \setminus S$  tal que  $S \cup \{e\}$  continua a não formar um ciclo (note-se que a aresta de custo mínimo que é analisada é eliminada do grafo modificado, independentemente de ser ou não escolhida para  $S$ ). Este procedimento é repetido até se obterem  $|V(G)| - 1$  arestas ou até não existirem mais arestas no grafo modificado (neste último caso, conclui-se que o grafo original não é conexo). Segue-se o algoritmo de Kruskal, cujos dados de entrada são o grafo  $G$  e a função de pesos nas arestas  $w$ .

- **Algoritmo de Kruskal( $G, w$ )**

1. Ordenar as arestas de  $G$ ,  $E(G) = \{e_1, \dots, e_m\}$  de tal forma que  $w(e_1) \leq \dots \leq w(e_m)$ ;
2. Fazer  $V = V(G)$ ,  $E^0 = \emptyset$  e  $k = 0$ ;
3. Se  $(V, E^k \cup \{e_{k+1}\})$  não tem circuitos,
  - (a) então  $E^{k+1} = E^k \cup \{e_{k+1}\}$ ;
  - (b) senão  $E^{k+1} = E^k$ ;

4. Se  $|E^{k+1}| < |V| - 1$  e  $k < m$ ,
  - (a) então  $k \leftarrow k + 1$  e voltar ao passo 3
  - (b) senão, se  $k < m$ ,
    - i. então devolver o grafo  $(V, E^{k+1})$ ;
    - ii. caso contrário,  $G$  não é conexo;

- Fim do algoritmo.

O teorema que se segue, mostra que o algoritmo de Kruskal determina uma árvore abrangente de custo mínimo de um grafo conexo com custos nas arestas.

**Exemplo 5.6.** *Vamos determinar um árvore abrangente de custo mínimo do grafo  $G$  com custos nas arestas, definido pela matriz de custos*

$$W_G = \begin{pmatrix} 0 & 1 & \infty & 10 & 8 & 3 \\ 1 & 0 & 13 & 10 & 6 & 4 \\ \infty & 13 & 0 & 15 & \infty & 4 \\ 10 & 10 & 15 & \infty & 9 & \infty \\ 8 & 6 & \infty & 9 & 0 & 7 \\ 3 & 4 & 4 & \infty & 7 & 0 \end{pmatrix},$$

utilizando o algoritmo de Kruskal.

**Solução.** O primeiro passo do algoritmo consiste na ordenação das arestas de  $G$ . Vamos assumir que, como resultado desta ordenação, se obtém:  $e_1 = v_1v_2$ ,  $e_2 = v_1v_6$ ,  $e_3 = v_2v_6$ ,  $e_4 = v_3v_6$ ,  $e_5 = v_2v_5$ ,  $e_6 = v_5v_6$ ,  $e_7 = v_1v_5$ ,  $e_8 = v_4v_5$ ,  $e_9 = v_1v_4$ ,  $e_{10} = v_2v_4$ ,  $e_{11} = v_2v_3$  e  $e_{12} = v_3v_4$ .

Segue-se uma tabela, onde se pretende descrever cada um dos passos resultantes da aplicação do algoritmo.

$k$	$e_k$	$w(e_k)$	Inserir $e_k$ ?	Árvore
1	$v_1v_2$	1	sim	$\{v_1v_2\}$
2	$v_1v_6$	3	sim	$\{v_1v_2, v_1v_6\}$
3	$v_2v_6$	4	não	
4	$v_3v_6$	4	sim	$\{v_1v_2, v_1v_6, v_3v_6\}$
5	$v_2v_5$	6	sim	$\{v_1v_2, v_1v_6, v_3v_6, v_2v_5\}$
6	$v_5v_6$	7	não	
7	$v_1v_5$	8	não	
8	$v_4v_5$	9	sim	$\{v_1v_2, v_1v_6, v_3v_6, v_2v_5, v_4v_5\}$

Note-se que o algoritmo termina, após a inserção da quinta aresta. Na Figura 5.6, representa-se o grafo  $G$  e a árvore abrangente de custo mínimo obtida.  $\square$

**Teorema 5.10.** *Se  $G$  é um grafo conexo, então o algoritmo de Kruskal determina uma árvore abrangente de custo mínimo.*

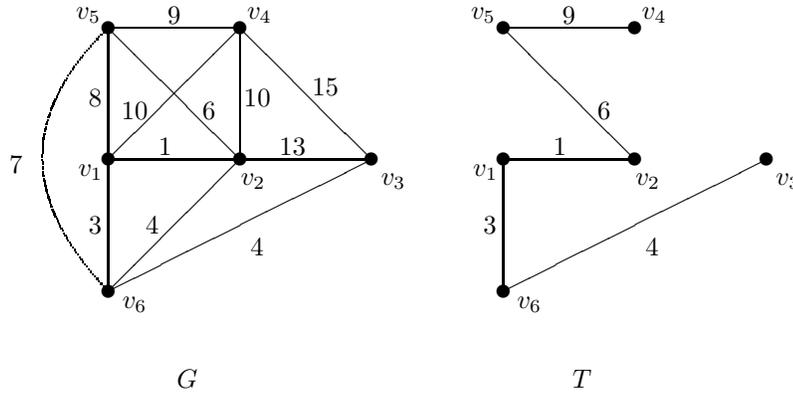


Figura 5.6: Grafo  $G$  do Exemplo 5.6 e árvore abrangente de custo mínimo  $T$  (obtida pelo algoritmo de Kruskal).

**Demonstração.** Uma vez que é imediato concluir que o algoritmo de Kruskal determina uma árvore abrangente (no caso de grafos conexos), apenas vamos mostrar, por redução ao absurdo, que esta árvore tem custo mínimo.

Sem perda de generalidade, vamos supor que após Ordenar( $E(G)$ ) (que ordena as arestas segundo os respectivos custos), se verifica  $w(e_1) \leq w(e_2) \leq \dots \leq w(e_m)$ . Suponha que a árvore  $\tilde{T}$ , determinada pelo algoritmo de Kruskal, não é ótima (ou seja, não tem custo mínimo) e seja  $T$  uma árvore abrangente ótima tal que  $T$  e  $\tilde{T}$  têm as mesmas arestas com índices não superiores a  $k - 1$  e que  $k$  é o maior índice nesta condições. Então,  $e_k$  é a próxima aresta a considerar para (eventual) inserção no conjunto de arestas que vão formar  $\tilde{T}$ . Sem perda de generalidade, suponhamos que  $e_k$  é inserida em  $\tilde{T}$  (ou seja, não forma qualquer ciclo conjuntamente com as arestas anteriormente inseridas) e é tal que  $e_k \notin E(T)$ . Logo,  $E(T) \cup \{e_k\}$  contém um ciclo  $C$  que, necessariamente, contém uma aresta  $e$  que não pertence a  $E(\tilde{T})$  e é tal que  $w(e) \geq w(e_k)$ . Consequentemente, substituindo  $E(T)$  por  $(E(T) \setminus \{e\}) \cup \{e_k\}$  obtém-se uma árvore abrangente  $T'$  de custo não superior ao custo de  $T$ .

1. Se  $w(e) > w(e_k)$ , então  $T'$  tem custo inferior ao de  $T$ , o que é absurdo, uma vez que, por hipótese,  $T$  é uma árvore ótima.
2. Se  $w(e) = w(e_k)$ , então  $T'$  tem custo igual ao de  $T$ , o que também é absurdo, uma vez que  $T'$  tem pelo menos as  $k$  primeiras arestas coincidentes com as de  $\tilde{T}$  (em contradição com a definição de  $T$ ).

Como consequência, podemos concluir que  $\tilde{T}$  é uma árvore abrangente ótima.  $\square$

### Algoritmo de Prim

O algoritmo de Prim, também conhecido por *algoritmo do vizinho mais próximo*, aplicado a um grafo  $G$ , começa com um vértice arbitrário  $u \in V(G)$ , a partir do qual se escolhe a aresta de custo mínimo  $uv$ , de entre as que lhe são incidentes, com a qual se forma o conjunto corrente de arestas candidatas a incluir na árvore abrangente de custo mínimo. Sendo *Árvore* o subconjunto corrente de arestas da árvore abrangente de custo mínimo e  $T$  o respectivo subconjunto de vértices, determina-se a aresta de custo mínimo, de entre as arestas do corte  $\partial(T)$  que, conjuntamente com as arestas em *Árvore*, não forma qualquer ciclo. Este procedimento é repetido, até não ser possível incluir em *Árvore* qualquer aresta adicional.

**Exemplo 5.7.** *Vamos determinar uma árvore abrangente de custo mínimo, para o grafo do Exemplo 5.6, por aplicação do algoritmo de Prim.*

**Solução.** Seja  $v_3$ , por exemplo, o vértice inicial, pelo que  $T = \{v_3\}$  e *Árvore* =  $\emptyset$ . Depois de iniciadas as diferentes variáveis, na tabela a seguir, para cada vértice  $v_i \in V \setminus T$  apresenta-se  $\alpha_i$  (o vértice mais próximo de  $v_i$  de entre os vértices da árvore corrente  $T$ ) e  $\beta_i$  (a distância entre  $v_i$  e a árvore  $T$ , ou seja, o custo da aresta  $\alpha_i v_i$ ).

$i$	1	2	4	5	6
$\alpha_i$		$v_3$	$v_3$		$v_3$
$\beta_i$	$\infty$	13	15	$\infty$	4

Seguem-se as tabelas de valores obtidos ao longo das cinco iterações necessárias para a determinação da árvore abrangente de custo mínimo, por aplicação do algoritmo de Prim.

1. O menor valor de  $\beta_i$  é  $\beta_6 = 4$ . Logo, inserimos o vértice  $v_6$  em  $T$  e a aresta  $v_3 v_6$  na árvore. Como consequência,  $T = \{v_3, v_6\}$ , *Árvore* =  $\{v_3 v_6\}$  e obtém-se:

$i$	1	2	4	5
$\alpha_i$	$v_6$	$v_6$	$v_3$	$v_6$
$\beta_i$	3	4	15	7

2. Neste caso, o menor valor de  $\beta_i$  é  $\beta_1 = 3$ . Logo,  $T = \{v_1, v_3, v_6\}$ , *Árvore* =  $\{v_3 v_6, v_1 v_6\}$  e obtém-se:

$i$	2	4	5
$\alpha_i$	$v_1$	$v_1$	$v_6$
$\beta_i$	1	10	7

3. O menor valor de  $\beta_i$  é  $\beta_2 = 1$ . Logo,  $T = \{v_1, v_2, v_3, v_6\}$ , *Árvore* =  $\{v_3 v_6, v_1 v_6, v_1 v_2\}$  e obtém-se:

$i$	4	5
$\alpha_i$	$v_1$	$v_2$
$\beta_i$	10	6

4. O menor valor de  $\beta_i$  é  $\beta_2 = 6$ . Logo,  $T = \{v_1, v_2, v_3, v_5, v_6\}$ ,  $\text{Árvore} = \{v_3v_6, v_1v_6, v_1v_2, v_2v_5\}$  e

$i$	4
$\alpha_i$	$v_5$
$\beta_i$	9

5. Tendo em conta a última tabela obtida, conclui-se que o vértice  $v_5$  é o vértice da árvore mais próximo de  $v_4$ . Logo, obtém-se a árvore abrangente de custo mínimo:  $T = V(G)$  e  $\text{Árvore} = \{v_3v_6, v_1v_6, v_1v_2, v_2v_5, v_4v_5\}$ .

Note-se que, embora as arestas tenham sido inseridas por uma ordem distinta, a árvore abrangente de custo mínimo obtida pelo algoritmo de Prim é a mesma que foi obtida pelo algoritmo de Kruskal. No caso geral, porém, quando existem várias árvores abrangentes ótimas, nem sempre os algoritmos de Kruskal e de Prim produzem a mesma árvore abrangente de custo mínimo.  $\square$

Em geral, é difícil dizer qual é o melhor algoritmo, de entre os algoritmos de Kruskal e de Prim. Na prática, porém, verifica-se que o algoritmo de Prim é mais rápido para grafos de ordem pequena e dimensão elevada e o algoritmo de Kruskal comporta-se melhor para grafos de ordem elevada e com poucas arestas.

## 5.4 Exercícios.

**Exercício 5.1.** *Dado um grafo  $G$  de ordem  $n \geq 2$  prove que*

1.  $\exists x, y \in V(G)$  tal que  $x \neq y \wedge d_G(x) = d_G(y)$ ;
2. o número de vértices de grau ímpar é par.

**Exercício 5.2.** *Sabendo que uma bola de futebol é fabricada unindo pentágonos e hexágonos regulares, de tal forma que cada pentágono é cosido a cinco hexágonos e cada hexágono a três pentágonos e três hexágonos, alternadamente, diga se o número de pentágonos é par ou ímpar.<sup>1</sup>*

**Exercício 5.3.** *Determine o número de arestas do grafo bipartido completo  $K_{mn}$ .*

**Exercício 5.4.** *Determine o número de arestas de um grafo conexo  $G$ , com apenas dois circuitos, sabendo que  $|V(G)| = n$ .*

<sup>1</sup>Este problema foi motivado por um outro (um pouco mais complicado, mas muito interessante) publicado em [4].

**Exercício 5.5.** Sabendo que um grafo tem dimensão  $m$  e ordem  $n$ , determine a dimensão do complementar.

**Exercício 5.6.** Dado um grafo  $G$ , prove as desigualdades:  $\text{raio}(G) \leq \text{diam}(G) \leq 2\text{raio}(G)$ .

**Exercício 5.7.** Prove as implicações  $2 \Rightarrow 3$  e  $3 \Rightarrow 1$  do Teorema 5.7.

**Exercício 5.8.** Dado um grafo  $G$ , prove que se  $|E(G)| \geq |V(G)|$ , então  $G$  tem pelo menos um circuito.

**Exercício 5.9.** Prove que a noção de distância entre vértices é uma métrica.

**Exercício 5.10.** Dada uma matriz quadrada  $A$ , denote-se o traço de  $A$  (ou seja, a soma dos seus elementos diagonais) por  $\text{tr}(A)$ , o menor e o maior valor próprio de  $A$  por  $\lambda_{\min}(A)$  e  $\lambda_{\max}(A)$ , respectivamente, a matriz quadrada com entradas todas iguais à unidade por  $J$  e a matriz identidade de ordem  $n$  por  $I_n$ . Tendo em conta esta notação e sendo  $G$  um grafo arbitrário tal que  $V(G) = \{v_1, \dots, v_n\}$ , prove cada uma das afirmações:

1.  $\text{tr}(A_G^3)$  é igual a  $6t$ , onde  $t$  denota o número de triângulos em  $G$ .
2.  $\text{diam}(G) \leq k$  se e só se  $\sum_{i=0}^k A^i \geq J$ , onde  $A^0 = I_n$  e a desigualdade é considerada componente a componente.

**Exercício 5.11.** Dado um digrafo  $\vec{G}$  tal que  $V(\vec{G}) = \{x_1, \dots, x_n\}$ , prove a igualdade  $B_{\vec{G}} B_{\vec{G}}^T = D_{\vec{G}} - A_{\vec{G}}$ , onde  $D_{\vec{G}} = \text{diag}(d_G(x_1), \dots, d_G(x_n))$ .

**Exercício 5.12.** Prove cada uma das desigualdades (6.14).

**Exercício 5.13.** Determine os caminhos mais curtos entre cada par de vértices do multidigrafo definido pela seguinte matriz de custos nas arestas:

$$W = \begin{pmatrix} 0 & 2 & \infty & 1 & \infty & \infty \\ \infty & 0 & 1 & 8 & \infty & \infty \\ 2 & \infty & 0 & \infty & \infty & 2 \\ \infty & \infty & \infty & 0 & 10 & \infty \\ \infty & 4 & \infty & \infty & 0 & 2 \\ \infty & 6 & 1 & \infty & 3 & 0 \end{pmatrix},$$

utilizando o algoritmo de Dijkstra.

**Exercício 5.14.** Determine os caminhos mais curtos entre cada par de vértices do grafo definido pela seguinte matriz de custos nas arestas:

$$W_G = \begin{pmatrix} 0 & 1 & \infty & 5 & 4 & 13 \\ 1 & 0 & 3 & 7 & 3 & 1 \\ \infty & 3 & 0 & 12 & \infty & 14 \\ 5 & 7 & 12 & 0 & 9 & \infty \\ 4 & 3 & \infty & 9 & 0 & 7 \\ 13 & 1 & 14 & \infty & 7 & 0 \end{pmatrix},$$

utilizando o algoritmo de Dijkstra.

**Exercício 5.15.** A exploração de uma determinada rede de transportes obriga a garantir um conjunto de ligações regulares (ida e volta) entre pares de cidades de modo a permitir o transporte de passageiros (com eventuais mudanças de autocarro em pontos intermédios) entre quaisquer duas cidades da tabela a seguir, onde se indicam as distâncias entre pares de cidades.

	2	3	4	5	6	7	8
1. Beja	507	348	141	372	183	454	144
2. Braga		166	621	256	357	49	388
3. Coimbra			462	155	198	114	230
4. Faro				511	297	568	257
5. Guarda					368	203	381
6. Lisboa						304	44
7. Porto							336
8. Setúbal							

Supondo que uma dada companhia pretende obter esta concessão, com um conjunto de ligações regulares cuja distância total seja mínima, indique o grafo correspondente ao conjunto de ligações a propor,

- a) utilizando o algoritmo de Kruskal;
- b) utilizando o algoritmo de Prim.

**Exercício 5.16.** Determina a árvores abrangente de custo mínimo do grafo representado pela matriz de custos do Exercício 5.14,

- a) utilizando o algoritmo de Kruskal;
- b) utilizando o algoritmo de Prim.



## Capítulo 6

# Problemas Clássicos em Optimização Combinatória

### 6.1 Poliedros combinatórios e teorema de König

Uma ideia muito simples, com consequências muito úteis em Optimização Combinatória e que tem sido alvo de intensa investigação, consiste em formular os problemas combinatórios como problemas de optimização linear inteira ou mesmo (em certos casos) como problemas de optimização linear.

Suponha que temos um problema de minimização de uma função  $c : 2^E \mapsto \mathbb{R}$ , num conjunto  $\mathcal{S} \subseteq 2^E$  de soluções admissíveis (note-se que  $2^E$  corresponde ao conjunto das partes de  $E$ ) e seja  $\mathbb{R}^E$  o espaço vectorial real  $|E|$ -dimensional, onde cada vector  $x$  tem as suas componentes indexadas pelos elementos de  $E$ . Para cada subconjunto  $F \subseteq E$ , denotando por  $x(F)$  o vector característico de  $F$  (ou seja, o vector  $x = x(F)$  tal que qualquer que seja  $e \in E$ ,  $x_e = 1$  se  $e \in F$  e  $x_e = 0$  se  $e \notin F$ ), vamos associar ao conjunto de soluções admissíveis  $\mathcal{S}$ , o invólucro convexo dos vectores característicos dos elementos de  $\mathcal{S}$ , ou seja, um politopo  $P_{\mathcal{S}} = H(\{x(F) \in \mathbb{R}^E : F \in \mathcal{S}\})$ . Assim, o problema de optimização combinatória

$$\min\{c(F) : F \in \mathcal{S}\}, \quad (6.1)$$

pode escrever-se (considerando  $c$  como vector do espaço dual de  $\mathbb{R}^E$ ) na forma

$$\min\{c^T x : x \in P_{\mathcal{S}}\}. \quad (6.2)$$

Qualquer solução admissível para o problema de optimização combinatória (6.1) corresponde a um vértice do politopo  $P_{\mathcal{S}}$  e reciprocamente. Uma vez que  $P_{\mathcal{S}}$  é um politopo, o problema (6.2) é um programa linear que tem uma solução óptima que é um ponto extremo de  $P_{\mathcal{S}}$ , qualquer que seja a função objectivo. Logo, resolvendo (6.2), obtemos um vector característico  $x^*$  que corresponde a um conjunto  $F^* \in \mathcal{S}$  que é solução óptima para (6.1).

Antes de prosseguirmos convém introduzir o conceito de matriz *totalmente unimodular*, bem como alguns resultados associados. Uma matriz  $B$  com entradas inteiras diz-se *unimodular* se  $\det(B) = \pm 1$ . Por sua vez, uma matriz arbitrária  $A$  com entradas inteiras diz-se *totalmente unimodular* se toda a submatriz quadrada de  $A$ , não singular, é unimodular. Como consequência desta definição, podemos concluir o seguinte resultado.

**Lema 6.1.** *Se no sistema  $Bx = \hat{b}$ , a matriz quadrada de ordem  $n$ ,  $B$ , é unimodular e o vector de termos independentes  $\hat{b}$  tem componentes inteiras, então este sistema tem uma solução inteira.*

**Demonstração.** Com efeito, sendo  $B = [\hat{b}_1, \dots, \hat{b}_n]$ , por aplicação da regra de Cramer, para  $i = 1, \dots, n$ , vem

$$x_i = \frac{\det([\hat{b}_1, \dots, \hat{b}_{i-1}, \hat{b}, \hat{b}_{i+1}, \dots, \hat{b}_n])}{\det(B)} \quad (6.3)$$

e, conseqüentemente, as componentes de  $x$  são inteiras (uma vez que, na expressão (6.3), o numerador é inteiro e o denominador é  $\pm 1$ ).  $\square$

**Teorema 6.2.** *Dado um politopo definido pelo conjunto  $X = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$ , onde  $A$  é uma matriz totalmente unimodular, se as componentes de  $b$  são inteiras, então os pontos extremos de  $X$  (ou seja, as soluções básicas admissíveis para  $X$ ), têm componentes inteiras.*

**Demonstração.** A prova decorre directamente das considerações anteriores.  $\square$

**Teorema 6.3.** *Uma matriz  $B$  cujas entradas  $b_{ij} \in \{0, +1, -1\}$  é totalmente unimodular se não mais do que duas componentes em cada coluna são não nulas e podemos partir o conjunto de índices linha de  $B$  em dois subconjuntos  $I_1$  e  $I_2$  de tal modo que*

1. *se uma coluna tem duas componentes não nulas com o mesmo sinal, então as linhas correspondentes têm os seus índices em subconjuntos diferentes;*
2. *se uma coluna tem duas componentes não nulas de sinal diferente, então as linhas correspondentes têm os seus índices no mesmo subconjunto.*

**Demonstração.** A prova vai ser feita por indução sobre a ordem das submatrizes quadradas de  $B$ , tendo em conta que qualquer submatriz não nula com um único elemento é unimodular. Assim, seja  $C$  uma submatriz quadrada de  $B$  de ordem  $k > 1$  e vamos supor que o resultado se verifica para submatrizes quadradas de ordem inferior a  $k$ . Se  $C$  tem uma coluna de zeros então é singular. Se  $C$  tem uma coluna com uma única componente não nula, então podemos expandir o respectivo determinante ao longo dessa coluna, pelo que o resultado decorre da hipótes de indução. Finalmente, quando  $C$  tem duas componentes

não nulas, em cada uma das suas colunas, as as condições 1. e 2. implicam que para qualquer índice coluna de  $C$ ,  $j$ , se verifique a igualdade

$$\sum_{i \in I_1} b_{ij} = \sum_{i \in I_2} b_{ij},$$

pelo que as linhas de  $C$  são linearmente dependentes e, conseqüentemente,  $C$  é singular.  $\square$

Com base neste teorema, podemos concluir que as matrizes de incidência arco vértice dos digrafos são matrizes totalmente unimodulares. Com efeito, no caso de um digrafo,  $\vec{G}$ , podemos partir os índices linha de  $B_{\vec{G}}$  nos subconjuntos  $I_1 = \emptyset$  e  $I_2 = V(\vec{G})$ , os quais verificam as condições exigidas pelo Teorema 6.3. Por outro lado, também podemos concluir que a matriz de incidência de um grafo  $G$  não orientado bipartido também é totalmente unimodular, uma vez que sendo  $G = (V_1, V_2, E)$ , se pode fazer  $I_1 = V_1$  e  $I_2 = V_2$ .

## 6.2 Conjuntos independentes e colorações

Dois arestas dizem-se adjacentes se têm um vértice em comum. Um subconjunto de vértices (arestas) dois a dois não adjacentes diz-se um *estável* (*emparelhamento*) ou *conjunto independente de vértices* (arestas). Um estável (emparelhamento) de máxima cardinalidade designa-se por *estável máximo* (*emparelhamento máximo*). O número de vértices de um estável máximo de um grafo  $G$  denota-se por  $\alpha(G)$  e designa-se *número de estabilidade* ou *número de independência* de  $G$ .

Diz-se que um emparelhamento  $M$  satura um vértice  $v$  se  $v \in V(M)$ . Um emparelhamento que satura todos os vértices do grafo, ou seja, um emparelhamento  $M \subseteq E(G)$  tal que  $V(M) = V(G)$ , designa-se por emparelhamento *perfeito*.

Para o caso dos grafos linha, conclui-se que a um estável máximo de  $L(G)$  corresponde um emparelhamento máximo em  $G$ . Logo, se  $M \subseteq E(G)$  é um emparelhamento máximo de  $G$ , então  $\alpha(L(G)) = |M|$ .

Um caminho (ou um passeio) diz-se  *$M$ -alternado* ou, simplesmente, *alternado*, se as suas sucessivas arestas estão, alternadamente, em  $M$  e em  $\bar{M}$  (onde  $\bar{M} = E(G) \setminus M$ ). Se um caminho  $M$ -alternado, com pelo menos uma aresta, começa e termina em vértices não-saturados, dizemos que é um caminho *de aumento*, relativamente a  $M$ . O teorema a seguir, caracteriza os emparelhamentos máximos em função da existência de caminhos de aumento.

**Teorema 6.4** (Berge). *Dado um grafo conexo  $G$ , um emparelhamento  $M$  é máximo se e só se não existe nenhum caminho de aumento relativamente a  $M$ .*

**Demonstração.** Seja  $M$  um emparelhamento máximo de  $G$  e suponha que existe um caminho  $P$  de aumento, relativamente a  $M$ . Para  $M' = E(P) \setminus M$ , vem que  $M'$  é um emparelhamento em  $G$  e  $|M'| = |M| + 1$ , o que constitui uma contradição (uma vez que, por hipótese,  $M$  é um emparelhamento máximo).

Reciprocamente, seja  $M$  um emparelhamento de  $G$  que não é máximo e que  $G$  não contém nenhum caminho de aumento, relativamente a  $M$ . Sendo  $M'$  um emparelhamento máximo de  $G$ , por definição,  $|M'| > |M|$ . Considerando o subgrafo  $H$ , cujas arestas pertencem à diferença simétrica dos emparelhamentos  $M$  e  $M'$  e cujos vértices são exactamente os extremos destas arestas, ou seja,  $E(H) = M \Delta M'$  e  $V(H) = \{x : xy \in M \Delta M'\}$ , podemos concluir que cada vértice de  $H$  tem grau um ou dois (uma vez que é extremo de não mais do que uma aresta de cada um dos emparelhamentos). Como consequência, cada componente de  $H$  é um ciclo de comprimento par ou um caminho com arestas alternadamente em  $M$  e em  $M'$ . Uma vez que  $|M'| > |M|$ ,  $H$  contém mais arestas de  $M'$  do que de  $M$  e, consequentemente, dado que os ciclos contêm tantas arestas de  $M$  como de  $M'$ , existe pelo menos uma componente  $P$  de  $H$  que é um caminho que começa e termina numa aresta de  $M'$ . Logo, é claro que os vértices extremos de  $P$  são  $M$ -livres, não só em  $H$ , como também em  $G$ . Assim, considerando o caminho  $P$  em  $G$ , podemos concluir que  $P$  é um caminho de aumento, relativamente a  $M$ , o que constitui uma contradição.  $\square$

Teoricamente, com base no Teorema 6.4 (teorema de Berge), é fácil descrever um algoritmo combinatório que determine um emparelhamento máximo. Com efeito, este algoritmo pode descrever-se do seguinte modo:

1.  $M \leftarrow$  um emparelhamento;
2. **enquanto** existe um caminho  $P$  de aumento relativamente a  $M$ 
  - $M \leftarrow E(P) \setminus M'$ ;
3. devolver  $M$ .

O principal problema deste algoritmo, porém, reside na determinação de um caminho de aumento. Com efeito, em geral, uma tal tarefa exige um elevado número de operações, pelo que é computacionalmente pesada. No caso de grafos bipartidos, porém, existem métodos muito eficientes para a determinação de emparelhamentos máximos.

### 6.2.1 Emparelhamentos em grafos bipartidos

Quer pela sua maior simplicidade, quer pela sua grande aplicação prática, o problema da determinação de emparelhamentos (perfeitos ou não) em grafos bipartidos, ocupa lugar de destaque na teoria dos grafos. Seguem-se dois exemplos que evidenciam, precisamente, a aplicabilidade destes emparelhamentos.

1. Suponha que temos um conjunto finito de raparigas e um conjunto finito de rapazes, cada um dos quais conhece um dado subconjunto de raparigas. Será que é possível casar todos os rapazes de tal forma que cada um se case com uma rapariga que conhece? Por exemplo, supondo que  $X = \{g_1, g_2, g_3, g_4, g_5\}$  é o conjunto das raparigas e  $Y = \{b_1, b_2, b_3, b_4\}$  é o conjunto dos rapazes, e ainda que  $b_1$  conhece  $g_1, g_2$  e  $g_3$ ,  $b_2$  conhece  $g_2$ ,

$b_3$  conhece  $g_2$ ,  $g_3$  e  $g_5$  e  $b_4$  conhece  $g_4$  e  $g_5$ , uma solução possível consiste em casar  $b_1$  com  $g_1$ ,  $b_2$  com  $g_2$ ,  $b_3$  com  $g_3$  e  $b_4$  com  $g_5$ . Na linguagem da teoria dos grafos, o grafo bipartido  $G$  tal que  $V(G) = X \cup Y$  e  $E(G) = \{xy \in X \times Y : x \text{ conhece } y\}$  (ver Figura 6.1-(A)), modela este problema que, assim, consiste na determinação de um emparelhamento que satura todos os vértices de  $Y$ .

2. Uma universidade convida os professores  $P_1, P_2, P_3$  e  $P_4$ , para leccionarem disciplinas do conjunto  $\{D_1, D_2, D_3, D_4, D_5\}$ . Supondo que o professor  $P_1$  está preparado para leccionar  $D_1$ ,  $P_2$  para leccionar  $D_1$  e  $D_3$ ,  $P_3$  para leccionar  $D_1$  e  $D_3$  e  $P_4$  para leccionar  $D_2, D_3, D_4$  e  $D_5$ , como se distribuem as disciplinas, de modo que cada professor leccione uma única disciplina de entre aquelas para as quais está preparado. Neste caso, considerando-se o grafo bipartido representado na Figura 6.1-(B), o objectivo é determinar um emparelhamento que sature todos os vértices do conjunto  $\{P_1, P_2, P_3, P_4\}$ .

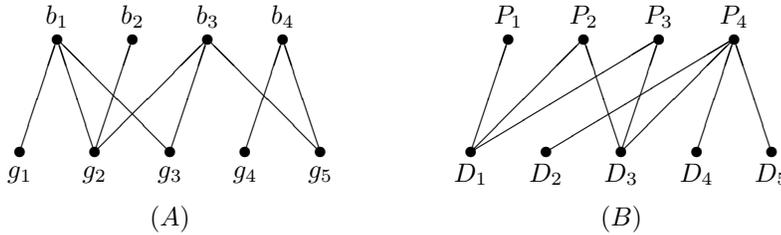


Figura 6.1: Exemplos de grafos bipartidos.

**Teorema 6.5.** *Se  $G$  é um grafo bipartido, então o invólucro convexo dos vectores característicos dos emparelhamentos  $M$  de  $G$  (matching polytope de  $M$ ) fica definido pelo seguinte conjunto de desigualdades:*

$$\sum_{e \in \partial(\{v\})} x_e \leq 1, \quad \forall v \in V(G), \quad (6.4)$$

$$x_e \geq 0, \quad \forall e \in E(G). \quad (6.5)$$

**Demonstração.** Vamos denotar por  $M(G)$  o conjunto dos vectores característicos representativos dos emparelhamentos de  $G$ , ou seja,  $M(G) = \{x(M) : M \text{ é um emparelhamento de } G\}$ , e por  $X \subseteq \mathbb{R}^E$  o conjunto de soluções do sistema de desigualdades (6.4)-(6.5).

- Dado um emparelhamento arbitrário  $M$  de  $G$ , é claro que o seu vector característico  $x(M)$  verifica as restrições (6.4)-(6.5), donde  $M(G) \subseteq X$ .
- Se às restrições (6.4) acrescentarmos as variáveis de desvio, obtém-se um sistema de  $|V(G)|$  equações e  $|E(G)| + |V(G)|$  variáveis, cuja matriz dos coeficientes  $[A, I]$  verifica as condições do Teorema 6.3. Logo, trata-se de

uma matriz totalmente unimodular, donde as suas submatrizes quadradas não singulares têm determinantes iguais a  $\pm 1$ . Como consequência, as soluções básicas admissíveis para (6.4)-(6.5), são soluções inteiras que, de acordo com as restrições, têm componentes com valores no conjunto  $\{0, 1\}$ , ou seja, são vectores característicos de subconjuntos de arestas sem vértices comuns, pelo que  $X \subseteq M(G)$ .

□

É fácil verificar que as restrições (6.4)-(6.5) não são suficientes para caracterizarem emparelhamentos em grafos não bipartidos. Por exemplo, considerando o triângulo com vértices  $v_1, v_2$  e  $v_3$ , e arestas  $a_1 = v_1v_2$ ,  $a_2 = v_2v_3$  e  $a_3 = v_3v_1$ , concluímos que  $x = \frac{1}{2}(1, 1, 1)$  é um ponto extremo do poliedro definido pelas restrições (6.4)-(6.5), ou seja,

$$\begin{array}{rcl} x_{a_1} & + & x_{a_2} & & & \leq & 1, \\ & & x_{a_2} & + & x_{a_3} & \leq & 1, \\ x_{a_1} & + & & + & x_{a_3} & \leq & 1, \\ x_{a_1} & , & x_{a_2} & , & x_{a_3} & \geq & 0. \end{array}$$

**Teorema 6.6.** *Se  $A$  é uma matriz totalmente unimodular e  $c$  um vector de componentes inteiras, então qualquer vértice do poliedro convexo  $P = \{y : y^T A \geq c^T, y \geq 0\}$ , tem componentes inteiras.*

**Demonstração.** Dado que  $A$  é totalmente unimodular,  $A^T$  é também totalmente unimodular e, consequentemente, a matriz  $[A^T, -I]$  é totalmente unimodular. Com efeito, o cálculo do determinante de qualquer submatriz quadrada  $C$  de  $[A^T, -I]$  que contenha (no seu todo ou em parte) uma ou mais colunas de  $-I$ , pode ser feito (no caso de nenhuma delas ser a coluna nula), expandindo sucessivamente cada uma dessas colunas, até se obter uma submatriz  $C'$  de  $A^T$ . Logo, tendo em conta que essas colunas são subvectores da base canónica,  $\det(C) = (-1)^p \det(C') = \pm 1$  e, consequentemente, do Teorema 6.2, decorre que os pontos extremos de  $P$  têm componentes inteiras. □

Como corolário deste teorema, podemos concluir que dada uma matriz  $A$  totalmente unimodular e dois vectores de inteiros  $b$  e  $c$ , cada um dos lados da relação de dualidade

$$\max\{c^T x : Ax \leq b, x \geq 0\} = \min\{y^T b : y^T A \geq c^T, y \geq 0\},$$

tem solução inteira,  $x^*$  e  $y^*$ , respectivamente, desde que os programas lineares tenham óptimo finito.

Dado um grafo  $G$ , designa-se por *cobertura por vértices* de  $G$ , um subconjunto de vértices  $C \subseteq V(G)$ , relativamente ao qual, cada aresta  $e \in E(G)$  tem pelo menos um extremo em  $C$ . Se  $C$  é uma cobertura por vértices e não existe uma cobertura por vértices  $C'$  tal que  $|C'| < |C|$ , diz-se que  $C$  é uma *cobertura por vértices mínima* (ou de cardinalidade mínima) e a sua cardinalidade designa-se por número de cobertura de  $G$  e denota-se por  $\beta(G)$ .

Note-se que o conceito de cobertura por vértices aparece em muitas aplicações. Por exemplo, sendo  $H$  um grafo cujos vértices são salas e dois vértices são adjacentes se existe um corredor entre as respectivas salas, o qual pode ser iluminado por uma fonte de luz colocada à entrada de uma das salas, podemos concluir que um conjunto de fontes de luz que ilumine todos os corredores constitui um conjunto de cobertura por vértices de  $H$ .

Do Teorema 6.2 decorre ainda que a determinação de um emparelhamento de cardinalidade máxima de um grafo bipartido  $G$  se pode fazer resolvendo o programa linear (P):

$$\max \sum_{e \in E(G)} x_e \quad (6.6)$$

$$s.a \sum_{e \in \partial(\{v\})} x_e \leq 1 \quad \forall v \in V(G), \quad (6.7)$$

$$x_e \geq 0, \quad \forall e \in E(G), \quad (6.8)$$

que, numa forma mais condensada, toma o aspecto:

$$\begin{aligned} \max \quad & \hat{e}^T x \\ \text{s. a} \quad & B_G x \leq \bar{e} \\ & x \geq 0, \end{aligned}$$

onde  $\hat{e}$  denota o vector com  $|E(G)|$  componentes todas unitárias e  $\bar{e}$  denota o vector com  $|V(G)|$  componentes todas unitárias. Denotando por  $V(e)$  o conjunto dos vértices extremos da aresta  $e$ , (P) tem como dual o programa linear (D):

$$\min \sum_{v \in V(G)} y_v \quad (6.9)$$

$$s.a \sum_{v \in V(e)} y_v \geq 1 \quad \forall e \in E(G), \quad (6.10)$$

$$y_v \geq 0 \quad \forall v \in V(G), \quad (6.11)$$

que se pode escrever na forma mais condensada:

$$\begin{aligned} \min \quad & y^T \bar{e} \\ \text{s. a} \quad & y^T B_G \geq \hat{e} \\ & y \geq 0. \end{aligned}$$

Com facilidade se verifica que o programa linear (D), determina o vector característico de um conjunto de cobertura por vértices do grafo bipartido  $G$ . Com efeito, qualquer vector de inteiros com componentes em  $\{0, 1\}$  que verifique as restrições (6.9)-6.11 é vector característico de uma cobertura por vértices de  $G$  e, reciprocamente, o vector característico de qualquer cobertura por vértices de  $G$  verifica as restrições (6.9)-6.11.

Na sequência do Teorema 6.6, existe uma solução óptima para (D) que corresponde ao vector característico de um conjunto de cobertura por vértices de

cardinalidade mínima. Adicionalmente, dado que os valores óptimos primal e dual coincidem, temos o seguinte teorema.

**Teorema 6.7** (König, 1931). *Num grafo bipartido, a cardinalidade de um emparelhamento máximo é igual à cardinalidade de uma cobertura mínima por vértices.*

**Demonstração.** Esta prova decorre, directamente, da análise que acabamos de fazer.  $\square$

### 6.2.2 Estáveis e conjuntos de cobertura

Um problema muito conhecido, relacionado com a determinação de um estável máximo de um grafo, é o problema das oito rainhas que consiste em colocar 8 rainhas num tabuleiro de xadrez de tal forma que as posições que ocupem não permita que nenhuma das rainhas possa atacar as outras (ou seja, de tal forma que não existam duas rainhas na mesma linha, coluna ou diagonal). Existem 92 soluções distintas para este problema. Como exercício, modele o problema das 8 rainhas como um problema de determinação de um estável máximo de um grafo.

Um conjunto de vértices adjacentes entre si designa-se por *clique*. Uma clique de máxima cardinalidade designa-se por *clique máxima* e a respectiva cardinalidade por *número de clique*. O número de clique de um grafo  $G$  denota-se por  $\omega(G)$ . É claro que  $\alpha(G) = \omega(\bar{G})$  e  $\omega(G) = \alpha(\bar{G})$ . Por outro lado, também se conclui facilmente que  $\omega(G) \leq \chi(G)$ .

O teorema a seguir estabelece uma certa correspondência entre cliques, estáveis e coberturas por vértices.

**Teorema 6.8.** *Dado um grafo  $G$  e um subconjunto de vértices  $S \subseteq V(G)$ , são equivalentes as seguintes proposições:*

1.  $S$  é uma clique de  $G$ .
2.  $S$  é um estável de  $\bar{G}$ .
3.  $V(G) \setminus S$  é uma cobertura por vértices de  $\bar{G}$ .

**Demonstração.** Vamos fazer a prova da equivalência das proposições, provando as implicações  $1 \Rightarrow 2$ ,  $2 \Rightarrow 3$  e  $3 \Rightarrow 1$ .

- (1  $\Rightarrow$  2) Se  $S$  é uma clique de  $G$ , então, tendo em conta a definição de  $\bar{G}$ , é claro que  $S$  é um estável de  $\bar{G}$ .
- (2  $\Rightarrow$  3) Se  $S$  é um estável de  $\bar{G}$ , então não existem arestas de  $\bar{G}$  a ligar dois vértices de  $S$ , donde todas as arestas de  $\bar{G}$  têm pelo menos um dos seus vértices extremos fora de  $S$ . Como consequência,  $V(G) \setminus S$  é uma cobertura por vértices de  $\bar{G}$ .
- (3  $\Rightarrow$  1) Se  $V(G) \setminus S$  é uma cobertura por vértices de  $\bar{G}$ , então nenhuma aresta de  $\bar{G}$  tem ambos os extremos em  $S$ . Consequentemente, não existe um par de vértices de  $S$  com uma aresta de  $\bar{G}$  entre eles, donde  $S$  é uma clique de  $G$ .

□

**Teorema 6.9** (Gallai, 1959). *Dado um grafo  $G$ ,  $\alpha(G) + \beta(G) = |V(G)|$ .*

**Demonstração.** Seja  $S$  um estável máximo para o grafo  $G$ . Pelo Teorema 6.8,  $V(G) \setminus S$  é um conjunto de cobertura por vértices e, conseqüentemente,

$$|V(G) \setminus S| = |V(G)| - \alpha(G) \geq \beta(G). \quad (6.12)$$

De modo semelhante se conclui que, sendo  $C$  um conjunto de cobertura mínima por vértices de  $G$ ,  $V(G) \setminus C$  é um estável de  $G$  e, conseqüentemente,

$$|V(G) \setminus C| = |V(G)| - \beta(G) \leq \alpha(G). \quad (6.13)$$

Tendo em conta as desigualdades (6.12)-(6.13), vem  $|V(G)| = \alpha(G) + \beta(G)$ . □

### 6.2.3 Colorações de vértices e arestas

Uma *coloração de vértices* de um grafo  $G$  é uma aplicação sobrejectiva

$$\begin{aligned} c : V(G) &\mapsto \{1, \dots, k\} \\ v &\rightsquigarrow c(v), \end{aligned}$$

A coloração  $c$  diz-se própria se para cada  $xy \in E(G)$  se verifica  $c(x) \neq c(y)$ . Como consequência, uma  $k$ -coloração própria dos vértices de um grafo  $G$  é equivalente à partição do conjunto dos seus vértices em  $k$  estáveis  $(V_1, V_2, \dots, V_k)$ . Se existe uma  $k$ -coloração própria de um grafo  $G$ , então  $G$  diz-se  *$k$ -colorável (para os vértices)*.

Dado um grafo arbitrário  $G$ , a determinação de uma coloração de vértices própria pode facilmente obter-se recorrendo, por exemplo, ao seguinte algoritmo:

#### 1. Algoritmo guloso (*greedy*) para coloração de vertices

- *Dados de entrada:* grafo  $G$ ;
- *Resultados de saída:* uma coloração própria dos vértices de  $G$ ;

2. Seja  $V(G) = \{v_1, v_2, \dots, v_n\}$ ;

3. Para  $i = 1, 2, \dots, n$ , colorir  $v_i$  com a menor cor possível.

Porém, para um valor fixo de  $k$ , verificar se um grafo é  $k$ -colorável pode tornar-se muito mais difícil, em particular, se pretendermos determinar o menor  $k$  para o qual tal acontece. O menor  $k$  para o qual existe uma coloração própria dos vértices de  $G$  designa-se por *número cromático* de  $G$  denota-se por  $\chi(G)$ .

**Teorema 6.10.** *Dado um grafo arbitrário  $G$ ,  $\chi(G) \leq \Delta(G) + 1$ .*

**Demonstração.** O algoritmo guloso (*greedy*) não utiliza mais do que  $\Delta(G) + 1$  cores. Logo,  $\chi(G) \leq \Delta(G) + 1$ . □

Seguem-se mais alguns resultados sobre o número cromático de um grafo.

**Teorema 6.11.** *Qualquer que seja o grafo  $G$ ,*

$$\chi(G) \leq \max_{U \subseteq V(G)} \delta(G[U]) + 1.$$

**Demonstração.** Suponha que  $G$  tem ordem  $n$  e seja

$$k = \max_{U \subseteq V(G)} \delta(G[U]).$$

Seja  $v_n$  um vértice de  $G$  tal que  $d_G(v_n) \leq k$  e  $H_{n-1} = G - \{v_n\}$ . Por hipótese,  $H_{n-1}$  tem um vértice de grau não superior a  $k$ . Seja  $v_{n-1}$  um desses vértices e seja  $H_{n-2} = H_{n-1} - \{v_{n-1}\}$ , isto é,  $H_{n-2} = G - \{v_n, v_{n-1}\}$ . Continuando este processo, obtém-se uma sequência de vértices de  $G$ ,  $v_n, \dots, v_1$ , tal que  $v_j$  é adjacente a um máximo de  $k$  vértices que o precedem. Consequentemente, para os colorir, no máximo, são necessárias  $k + 1$  cores.  $\square$

Se  $G$  é um grafo conexo não regular, então

$$\max_{U \subseteq V(G)} \delta(G[U]) \leq \Delta(G) - 1$$

e, consequentemente, a partir do Teorema 6.11, podemos concluir a desigualdade  $\chi(G) \leq \Delta(G)$ . O teorema de Brooks, que a seguir se apresenta sem demonstração, estende esta desigualdade aos grafos conexos regulares que não são completos nem ciclos de comprimento ímpar.

**Teorema 6.12** (Brooks, 1941). *Sendo  $G$  um grafo conexo, se  $G$  não é completo nem um ciclo de comprimento ímpar, então  $\chi(G) \leq \Delta(G)$ .*

Por sua vez, uma *coloração de arestas* de um grafo  $G$  é uma aplicação sobrejectiva

$$\begin{aligned} c' : E(G) &\mapsto \{1, \dots, k'\} \\ xy &\rightsquigarrow c(xy). \end{aligned}$$

Uma coloração de arestas  $c'$  diz-se própria se, para cada par de arestas  $xy, uv \in E(G)$  adjacentes, ou seja, tais que  $\{x, y\} \cap \{u, v\} \neq \emptyset$ , se verifica  $c'(xy) \neq c'(uv)$ . O menor  $k'$  para o qual existe uma coloração própria de arestas de  $G$ , designa-se por *índice cromático* de  $G$  e denota-se por  $\chi'(G)$ .

É claro que se o grafo  $G$  é não nulo, então  $\chi'(G) = \chi(L(G))$ , onde  $L(G)$  denota o grafo linha de  $G$ . O teorema de Vizing que a seguir se apresenta sem demonstração, estabelece limites muito apertados para a coloração de arestas de grafos arbitrários.

**Teorema 6.13** (Vizing, 1964). *Dado um grafo  $G$ , verificam-se as desigualdades*

$$\Delta(G) \leq \chi'(G) \leq \Delta(G) + 1.$$

Deve observar-se que, de acordo com a definição de coloração de arestas, um conjunto de arestas com a mesma cor é um emparelhamento. Como consequência, dado um grafo arbitrário  $G$  e sendo  $M \subseteq E(G)$  um emparelhamento máximo de  $G$ ,

$$|M| \geq \frac{|E(G)|}{\chi'(G)} \geq \frac{|E(G)|}{\Delta(G) + 1}. \quad (6.14)$$

Para o caso particular dos grafos bipartidos, segue-se (sem demonstração) um teorema publicado por König em 1916.

**Teorema 6.14** (König, 1916). *Se  $G$  é um grafo bipartido então  $\chi'(G) = \Delta(G)$ .*

A coloração de arestas de grafos bipartidos tem também aplicação na produção de quadrados latinos. Com efeito, denotando o conjunto dos índices das linhas de um quadrado latino  $L$  de ordem  $n$ , por  $I = \{i_1, \dots, i_n\}$  e o conjunto dos índices das colunas por  $J = \{j_1, \dots, j_n\}$ , qualquer coloração das arestas do grafo bipartido completo  $K_{nn} = (I, J, E)$ , com  $n$  cores, determina um quadrado latino  $L$ . Note-se que, de acordo com o Teorema 6.14,  $\chi'(K_{nn}) = n$ . Logo, identificando os  $n$  símbolos do quadrado latino  $L$  com as  $n$  cores utilizadas na coloração das arestas, conclui-se que não existem dois símbolos iguais na mesma linha ou coluna (uma vez que tal equivaleria à existência de duas arestas com a mesma cor incidentes num mesmo vértice).

- **Exemplo de aplicação.**

No caso do quadrado latino de ordem 3, obtém-se o grafo bipartido completo  $K_{33}$ , representado na Figura 6.2, onde as letras indicam a coloração adoptada para as arestas e se observa que cada conjunto de arestas com a mesma letra constitui um emparelhamento. Como consequência, a partir

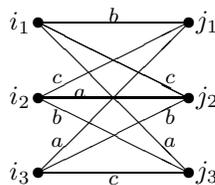


Figura 6.2: Grafo  $K_{33}$  com uma coloração de arestas

desta coloração de arestas, obtém-se o quadrado latino

$$L = \begin{matrix} & j_1 & j_2 & j_3 \\ i_1 & \left( \begin{array}{ccc} b & c & a \end{array} \right) \\ i_2 & \left( \begin{array}{ccc} c & a & b \end{array} \right) \\ i_3 & \left( \begin{array}{ccc} a & b & c \end{array} \right) \end{matrix}.$$

## 6.3 Circuitos de Euler e ciclos de Hamilton

Nesta secção, vamos analisar os circuitos de Euler e os ciclos de Hamilton, algumas implicações que a sua existência tem na estrutura de um grafo e alguns algoritmos para a sua determinação, no contextos dos problemas de optimização combinatória. No caso dos circuitos de Euler, em geral, consideramos multigrafos em vez de grafos.

### 6.3.1 Grafos de Euler

Um trajecto designa-se por *trajecto de Euler* se contém todas as arestas do multigrafo. Por sua vez, designa-se por *circuito de Euler*, todo o circuito que é um trajecto de Euler (ou seja, trajecto de Euler fechado). Também se podem definir trajectos e circuitos de Euler em multigrafos orientados (utilizando, nesse caso, as noções de trajecto e circuito orientado). Um multigrafo diz-se *euleriano* (ou *multigrafo de Euler*) se admite um circuito de Euler e diz-se *semi-euleriano* se admite um trajecto de Euler. É claro que todo o multigrafo euleriano é também semi-euleriano. Com base nestes conceitos, o problema das sete pontes de Königsberg reduz-se à questão de saber se o multigrafo que modela este problema é ou não euleriano.

**Exemplo 6.1.** *Vamos demonstrar que se um multigrafo  $G$  é euleriano, então todos os vértices têm grau par.*

**Solução.** Seja  $G$  um multigrafo euleriano e  $C$  um circuito de Euler. Escolha-se um vértice  $u$  para vértice inicial e final do circuito  $C$ . Percorrendo o circuito  $C$ , cada vez que passamos por um vértice  $v \neq u$  percorremos duas novas arestas incidentes em  $v$  donde, dado que o circuito utiliza todas as arestas, vem que o grau de  $v$  é par. Analogamente, quando passamos pelo vértice  $u$  percorremos duas arestas incidentes em  $u$ . Juntando a todas estas arestas percorridas a primeira e a última aresta incidentes em  $u$ , podemos concluir que todos os vértices de  $G$  têm grau par.  $\square$

**Exemplo 6.2.** *Seja  $G$  um multigrafo e  $C$  um circuito de  $G$ . Vamos demonstrar que para cada vértice  $v \in V(G)$ , a paridade do grau de  $v$  em  $G$  é a mesma que em  $G - E(C)$ .*

**Solução.** Sendo  $C$  um circuito de  $G$  e  $H = (V, E(C))$ , todos os vértices de  $H$  têm grau par. Uma vez que

$$\forall v \in V(G) \quad d_G(v) = d_{G-E(C)}(v) + d_H(v)$$

e  $d_H(v)$  é par, então  $d_G(v)$  e  $d_{G-E(C)}(v)$  têm a mesma paridade.  $\square$

O teorema que se segue é conhecido como teorema de Euler (dada a sua relação com a resolução do problema das sete pontes de Königsberg, obtida por Euler em 1736). A sua primeira demonstração, porém, foi publicada por Carl Hierholzer em 1873.

**Teorema 6.15** (Euler-Hierholzer). *Um multigrafo conexo não trivial é euleriano se e só se nenhum dos seus vértices tem grau ímpar.*

**Demonstração.** Utilizando o Exemplo 6.1, podemos concluir que se um multigrafo é euleriano então não tem vértices de grau ímpar. Assim, resta provar a implicação recíproca que demonstraremos por redução ao absurdo.

Suponha que  $G$  é um multigrafo conexo, não trivial, sem vértices de grau ímpar que não é euleriano e ainda que qualquer submultigrafo de  $G$  (com menos arestas) não tem esta propriedade. Uma vez que todos os vértices têm grau par, podemos concluir que  $G$  contém um circuito.

Seja  $C$  um circuito de  $G$  com comprimento máximo. Uma vez que  $G$  não é euleriano,  $C$  não é um circuito de Euler. Como consequência, o multigrafo  $G - E(C)$  tem uma componente  $G'$  com  $|E(G')| > 0$ . Do Exemplo 6.2 decorre que o multigrafo conexo  $G'$  não tem vértices de grau ímpar. Porém, uma vez que  $|E(G')| < |E(G)|$ ,  $G'$  admite um circuito de Euler  $C'$ . Adicionalmente, dado que  $G$  é conexo, existe um vértice  $v \in V(C) \cap V(C')$ , o qual podemos escolher para vértice inicial (e final) do circuito  $C \cup C'$  que, naturalmente, tem mais arestas do que as de  $C$ , o que constitui uma contradição (tendo em conta a maximalidade de  $C$ ).  $\square$

**Exemplo 6.3.** *Vamos demonstrar que um multigrafo conexo é semi-euleriano se e só se não tem mais do que dois vértices de grau ímpar.*

**Solução.** Se  $G$  contém um trajecto de Euler com vértices inicial e final coincidentes, então este trajecto é um circuito de Euler e pelo Teorema 6.15, todos os vértices têm grau par. Se os vértices inicial e final do trajecto são distintos, então ligando-os por uma nova aresta  $e$ , obtém-se um grafo  $G + e$  que é euleriano e, novamente pelo Teorema 6.15, todos os vértices de  $G + e$  têm grau par. Logo, no grafo  $G$  apenas os vértices extremos da aresta  $e$  têm grau ímpar.

Reciprocamente, seja  $G$  um multigrafo conexo com não mais do que dois vértices de grau ímpar. Uma vez que em qualquer grafo o número de vértices de grau ímpar é par, temos dois casos – o número de vértices de grau ímpar é zero ou dois. No primeiro caso (ausência de vértices de grau ímpar), o Teorema 6.15 implica que  $G$  tenha um circuito (que é também um trajecto) de Euler. No segundo caso, existindo dois vértices de grau ímpar, podemos ligar estes vértices por uma nova aresta  $e$  e o grafo obtido,  $G + e$ , não tem vértices de grau ímpar. Logo, aplicando o Teorema 6.15, o multigrafo  $G + e$  tem um circuito de Euler  $C$ , donde  $C - e$  é um trajecto de Euler para  $G$ .  $\square$

Segue-se um algoritmo para determinação de circuitos de Euler.

**Descrição de algoritmo de Hierholzer.**

*Dados de entrada:* Um grafo  $G$ .

*Resultados de saída:* Um circuito de Euler  $C$ , se um tal circuito existe.

1. Escolher um vértice  $v \in V(G)$  e determinar um circuito (ciclo)  $C$  que se inicie e termine em  $v$ . Fazer  $G \leftarrow G - E(C)$ .
2. Repetir (a) ... (c) até que  $G$  seja um grafo nulo (sem arestas).

- (a) Determinar  $v \in V(C)$  com pelo menos uma aresta de  $G$  incidente.
- (b) Determinar um circuito (ciclo)  $C'$  que se inicie e termine em  $v$ .
- (c) Juntar os circuitos  $C$  e  $C'$  e denotar o circuito resultante por  $C$ ;  
Fazer  $G \leftarrow G - E(C)$ .

3. Devolver o circuito de Euler  $C$ .

### Problema do carteiro chinês

Neste problema, um carteiro levanta a correspondência numa estação de correios, faz a respectiva distribuição e regressa ao ponto de partida. Assumimos que o carteiro deve percorrer todas as ruas da respectiva área de distribuição pelo menos uma vez. Sujeito a estas condições, pretendemos determinar o percurso que deve fazer o carteiro de modo a andar o menos possível. A designação de *problema do carteiro chinês* deve-se ao facto de ter sido primeiramente formulado pelo matemático chinês Mei-Ko Kwan em 1962. Podemos formular este problema na linguagem dos grafos com pesos (não negativos) nas arestas, da seguinte forma: considere-se o grafo correspondente à rede de ruas da área de distribuição do carteiro e associe-se a cada aresta o peso igual ao comprimento da rua a que se refere. Neste grafo, define-se *peso de um passeio fechado*  $v_0e_1, v_1e_2, \dots, e_nv_0$  como sendo a soma dos pesos de suas arestas, ou seja,  $\sum_{i=1}^n w(e_i)$ . Assim, o objectivo é encontrar um passeio fechado que contenha

todas as arestas e tenha peso mínimo, o qual será designado por *passeio óptimo para o carteiro chinês*. É claro que se o grafo considerado é de Euler, então um passeio é óptimo para o carteiro chinês se e só se é um circuito de Euler e, neste caso, o problema resolve-se por aplicação do algoritmo de Hierholzer. No caso contrário, porém, o problema é mais complicado, uma vez que o carteiro terá de passar mais do que uma vez por algumas ruas e temos de escolher quais. Para resolver este problema, vamos introduzir a operação de duplicação de arestas.

Dado um grafo (ou multigrafo)  $G$  e uma função de pesos nas arestas  $w : E(G) \rightarrow \mathbb{R}$ , a duplicação da aresta  $e \in E(G)$ , com peso  $w(e)$ , consiste em adicionar ao grafo (ou multigrafo) uma nova aresta com os mesmos vértices extremos de  $e$  e o mesmo peso (ou seja, consiste em criar uma aresta paralela a  $e$  com o mesmo peso).

Utilizando a operação de duplicação de arestas, o problema do carteiro chinês pode reformular-se como a seguir se indica.

Seja  $G$  um grafo com pesos não negativos nas arestas.

- (i) Com recurso à duplicação de arestas, determinar um supermultigrafo  $G^*$  do multigrafo  $G$  que seja euleriano e tal que o somatório

$$\sum_{e \in E(G^*) - E(G)} w(e)$$

tenha o menor valor possível.

(ii) Determinar um circuito de Euler para o multigrafo  $G^*$ .

Seja  $V^-$  o conjunto de vértices de grau ímpar num grafo  $G$  (é claro que  $|V^-|$  é par) e seja  $M$  o conjunto de arestas que definem caminhos entre vértices de  $V^-$ . Então, denotando por  $G^+(M)$  o grafo obtido a partir de  $G$  por duplicação das arestas contidos em  $M$ , temos o seguinte resultado.

**Teorema 6.16.** *Dado um multigrafo  $G$  com pesos nas arestas, existe um conjunto de caminhos  $M$  em  $G$  que emparelham os vértices de grau ímpar tal que o peso de  $M$  é mínimo e o circuito de Euler do grafo  $G^+(M)$  é uma solução ótima para o problema do carteiro chinês.*

**Demonstração.** Observe-se que as arestas duplicadas (ou seja, as arestas de  $M$ ) não contêm nenhum ciclo. Caso contrário, tendo em conta o Exemplo 6.2 e o teorema de Euler-Hierholzer, eliminadas as arestas deste ciclo, o grafo teria um circuito de Euler com menor peso.

Seja  $v_1$  um vértice de grau ímpar no grafo  $G$ . Então, qualquer passeio fechado ótimo  $\mathcal{C}$  passa por uma das arestas incidentes em  $v_1$  mais do que uma vez. Seja  $e_1 = v_1v_2$  a aresta repetida durante a construção do passeio ótimo  $\mathcal{C}$ . Se  $v_2$  tem grau ímpar em  $G$ , então  $v_1e_1v_2$  é um caminho de  $M$  que liga vértices de grau ímpar em  $G$ . Se  $v_2$  tem grau par em  $G$  então, com a duplicação de  $e_1$ ,  $v_2$  passa a ter grau ímpar e, tal como anteriormente, qualquer passeio fechado ótimo passa por uma das arestas incidentes em  $v_2$  mais do que uma vez. Este procedimento repete-se até que todos os vértices de grau ímpar estejam ligados por passeio formados pelas aresta repetidas pertencentes a  $M$ . Acrescentando estas arestas ao grafo original obtemos um novo grafo onde todos os vértices têm grau par. Assim, admitindo que  $M$  é o conjunto de arestas que definem os caminhos entre os pares de vértices de grau ímpar que no seu conjunto têm peso mínimo, obtém-se o resultado pretendido.  $\square$

No caso particular de um grafo  $G$  com exactamente dois vértices  $u$  e  $v$  de grau ímpar, a solução de (i) é equivalente à determinação de um  $(u, v)$ -caminho mais curto e à duplicação das arestas desse caminho.

**Exemplo 6.4.** *Vamos resolver o problema do carteiro chinês para o grafo definido pela seguinte matriz de pesos:*

$$W(G) = \begin{matrix} & v_1 & v_2 & v_3 & v_4 & v_5 & v_6 & v_7 & v_8 \\ \begin{matrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \\ v_8 \end{matrix} & \left( \begin{array}{cccccccc} \infty & 2 & \infty & \infty & \infty & \infty & \infty & 2 \\ 2 & \infty & 3 & \infty & \infty & \infty & 5 & 4 \\ \infty & 3 & \infty & 8 & \infty & 6 & 3 & \infty \\ \infty & \infty & 8 & \infty & 5 & 2 & \infty & \infty \\ \infty & \infty & \infty & 5 & \infty & 7 & \infty & \infty \\ \infty & \infty & 6 & 2 & 7 & \infty & 9 & \infty \\ \infty & 5 & 3 & \infty & \infty & 9 & \infty & 2 \\ 2 & 4 & \infty & \infty & \infty & \infty & 2 & \infty \end{array} \right) \end{matrix}.$$

**Solução.** Tendo em conta a Figura 6.3, onde se representa o grafo que modela este problema, é fácil verificar que apenas os vértices  $v_4$  e  $v_8$  têm grau ímpar. Determinado o caminho mais curto entre eles, com recurso (por exemplo) ao algoritmo de Dijkstra,

$$\text{DIJKSTRA}(G = (\{v_1, \dots, v_8\}, W(G)), v_4, v_8),$$

vem que o comprimento de um  $(v_4, v_8)$ -caminho mais curto é 13 e um destes caminhos é  $v_8v_7v_3v_4$  (o outro caminho com o mesmo peso é  $v_8v_7v_6v_4$ ).

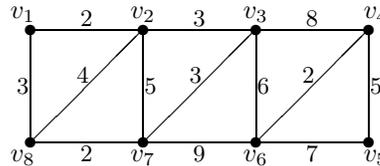


Figura 6.3: Grafo que modela o problema do carteiro chinês do Exemplo 6.4.

Duplicando as arestas deste caminho, obtemos o multigrafo euleriano representado na Figura 6.4.

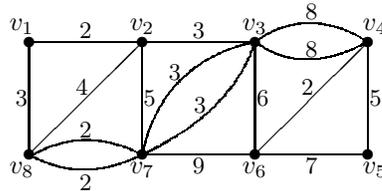


Figura 6.4: Grafo euleriano obtido a partir do grafo representado na Figura 6.3 por duplicação das arestas de um caminho mais curto entre  $v_4$  e  $v_8$ .

A soma dos pesos das arestas de um circuito de Euler deste grafo é igual a 72 e um destes circuitos é, por exemplo,

$$v_1v_2v_3v_4v_3v_6v_4v_5v_6v_7v_3v_7v_8v_2v_7v_8v_1,$$

sendo claro que este circuito é um passeio óptimo para o carteiro chinês.  $\square$

### 6.3.2 Grafos de Hamilton

Um caminho que contém todos os vértices de um grafo diz-se um caminho de Hamilton (ou hamiltoniano). Por sua vez, um ciclo que contém todos os vértices de um grafo, designa-se por ciclo de Hamilton (ou hamiltoniano). Também se podem definir caminhos e ciclos de Hamilton para grafos orientados (utilizando, nesse caso, a noção de caminho e ciclo orientado).

Como exemplo, considere-se um tabuleiro de xadrez e associe-se a cada um dos seus 64 quadrados um vértice de um grafo  $G$  cujas arestas ligam os vértices correspondentes a quadrados entre os quais é possível efectuar um movimento de

cavalo. Neste grafo existem ciclos de Hamilton que correspondem a movimentos sucessivos de um cavalo de forma a que todos os quadrados (pretos e brancos) sejam visitados uma única vez. Na Figura 6.5 representa-se um dos ciclos de Hamilton possíveis para o referido grafo.

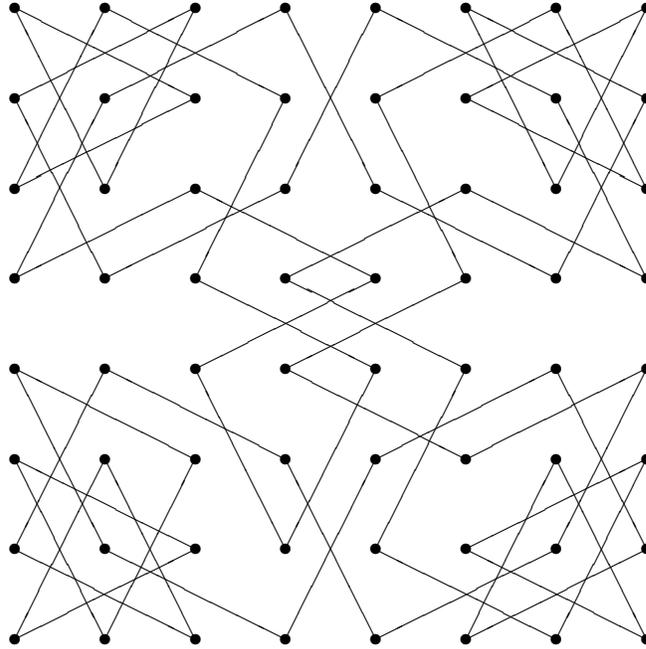


Figura 6.5: Circuito de Hamilton.

Um grafo (digrafo ou multidigrafo) que admite um ciclo de Hamilton, designa-se por grafo (digrafo ou multidigrafo) hamiltoniano ou grafo (digrafo ou multidigrafo) de Hamilton. Por sua vez, um grafo (digrafo ou multidigrafo) que admite um caminho de Hamilton diz-se um grafo (digrafo ou multidigrafo) semi-hamiltoniano.

Contrariamente ao que acontece no caso dos circuitos eulerianos, não se conhece nenhuma condição necessária e suficiente para a existência de um ciclo de Hamilton. No entanto, podem apresentar-se, separadamente, condições necessárias e condições suficientes.

Os grafos hamiltonianos apresentam propriedades muito interessantes. Por exemplo, é muito fácil colorir as arestas de um grafo cúbico hamiltoniano,  $G$ , utilizando três cores. Com efeito, sendo  $C = (x_1, x_2, \dots, x_{2k}, x_1)$  um ciclo de Hamilton em  $G$ , podemos atribuir a cor 1 às arestas  $x_{2j-1}x_{2j}$ , para  $j = 1, \dots, k$ , a cor 2 à aresta  $x_{2k}x_1$  e às arestas  $x_{2j}x_{2j+1}$ , para  $j = 1, \dots, k-1$  e a cor 3 às restantes. Logo, tendo em conta o teorema de Vizing (Teorema 6.13), podemos concluir que os grafos cúbicos hamiltonianos têm índice cromático 3.

### Problema do caixeiro viajante

O problema do caixeiro viajante que usualmente se denota por TSP (iniciais de *Traveling Salesman Problem*), consiste em determinar um percurso que permita visitar um conjunto de cidades, passando uma única vez em cada cidade e voltando à cidade de origem, com custo (ou tempo) total mínimo. Na terminologia da teoria dos grafos, dado um grafo (digrafo) completo com pesos não negativos nas arestas (arcos), este problema reduz-se à determinação de um ciclo (ciclo orientado) de Hamilton de peso mínimo. Usando apenas a notação dos grafos (uma vez que para os digrafos tudo se passa de modo equivalente), não existe perda de generalidade ao definir-se o problema para o grafo completo, uma vez que, sendo  $G$  um grafo arbitrário com pesos não negativos nas arestas definidos pela função  $w_G : E(G) \mapsto \mathbb{R}^+ \cup \{\infty\}$ , a determinação de um ciclo de Hamilton de peso mínimo em  $G$  é equivalente à resolução do TSP para o grafo completo  $K$  com função de pesos  $w_K : E(K) \mapsto \mathbb{R}^+ \cup \{\infty\}$  tal que

$$w_K(e) = \begin{cases} w_G(e), & \text{se } e \in E(G); \\ \infty, & \text{se } e \notin E(G). \end{cases}$$

O TSP é fácil de formular e, em geral, não é necessário grande talento para a obtenção de ciclos de Hamilton de peso aceitável (quando existem), mesmo para grandes problemas. No entanto, a determinação da solução do problema tem resistido à obtenção de "bons" algoritmos de resolução. Trata-se assim de um problema que contém todos os condimentos que têm atraído muitos matemáticos ao longo dos séculos – formulação simples e dificuldade de resolução. Existem também razões de carácter prático que contribuem para a importância do TSP, dado que muitos problemas da vida real se podem formular como casos particulares do TSP.

Uma abordagem imediata para a resolução do TSP consiste na pesquisa exaustiva de todos os ciclos de Hamilton do grafo (digrafo ou multidigrafo) e na escolha daquele que apresenta peso mínimo. A seguir descreve-se um algoritmo que utiliza esta pesquisa exaustiva num grafo e que, naturalmente, se pode estender a um digrafo ou multidigrafo.

#### Algoritmo de resolução do TSP por pesquisa exaustiva

*Dados de entrada:* grafo  $G$  de ordem  $n$ , com matriz de custos  $W = (w_{ij})$ ;

*Resultados de saída:* ciclo de Hamilton de peso mínimo  $C$ ;

1. Fazer  $PesoMin = \infty$ ;
2. Para todas as permutações  $\pi$  de  $[n]$  fazer
  - (a)  $Peso \leftarrow w_{\pi(n)\pi(1)} + \sum_{i=1}^{n-1} w_{\pi(i)\pi(i+1)}$ ;
  - (b) Se  $Peso < PesoMin$  então  $PesoMin \leftarrow Peso$  e  $C \leftarrow \pi$ ;
3. Devolver  $C$ .

Um digrafo (ou multidigrafo) completo de ordem  $n$  tem  $(n-1)!$  ciclos de Hamilton, os quais podem ser agrupado em pares de ciclos que diferem

apenas no sentido em que são percorridos (se a permutação do conjunto de vértices  $[1, 2, \dots, n-1, n]$  define um ciclo de Hamilton então a permutação  $[n, n-1, \dots, 2, 1]$  também define). No caso de grafos não orientados, ou seja, no caso simétrico em que  $w_{ij} = w_{ji}$ , existem "apenas"  $\frac{1}{2}(n-1)!$  ciclos de Hamilton. No entanto, mesmo para valores de  $n$  não muito grandes, este número continua a ser muito elevado. Por exemplo, para  $n = 12$ ,  $\frac{1}{2}11! = 19.958.400$  ciclos hamiltonianos candidatos a ciclos de peso mínimo. Assim, esta pesquisa exaustiva tem um custo computacional muito elevado. Alternativamente ao método de pesquisa exaustiva que exige um esforço computacional que cresce exponencialmente com  $n$ , é comum, na prática, a utilização de métodos heurísticos com os quais, embora não se tenha a garantia de optimalidade, as experiências computacionais realizadas sugerem a obtenção de soluções próximas da solução ótima de modo eficiente.

Vamos introduzir uma técnica de "branch and bound"<sup>1</sup> para a determinação da solução ótima do TSP, introduzida por Little em 1963. Esta técnica apresenta a vantagem de, em muitos casos, a solução ótima ser obtida sem se analisarem exaustivamente todos os ciclos de Hamilton. Por questões de simplicidade, utilizaremos  $W^*$  para denotar a matriz de pesos não negativos de um multidigrafo de ordem  $\nu$ , com os elementos diagonais modificados para  $w_{ii}^* = \infty, i = 1, \dots, \nu$ , a qual vamos designar por matriz de pesos modificada. Antes de prosseguirmos, convém ainda introduzir o seguinte resultado evidentes mas de grande utilidade.

**Teorema 6.17.** *Dado um digrafo definido pela matriz quadrada de ordem  $\nu$  de pesos modificada  $W^* = (w_{ij}^*)$ , para  $i = 1, \dots, \nu$ ,*

$$\begin{aligned}\hat{w}_r &= \min\{w_{rj}^*, j \in [\nu]\}, \\ \bar{w}_s &= \min\{w_{is}^*, i \in [\nu]\},\end{aligned}$$

para  $r, s = 1, \dots, \nu$ , seja  $\hat{w}$  o vector coluna cujas componentes são  $\hat{w}_r$ , para  $r = 1, \dots, \nu$ ,  $\bar{w}^T$  o vector linha cujas componentes são  $\bar{w}_s$ , para  $s = 1, \dots, \nu$  e  $\hat{e}$  o vector coluna cujas componentes são todas iguais a 1. Denote-se por  $W^* - \hat{w}\hat{e}^T$  ( $W^* - \hat{e}\bar{w}^T$ ) a matriz que se obtém de  $W^*$  subtraindo  $\hat{w}_i$  a todas as entradas da  $i$ -ésima linha, para  $i = 1, \dots, \nu$  ( $\bar{w}_j^T$  a todas as entradas da  $j$ -ésima coluna, para  $j = 1, \dots, \nu$ ). Se  $P(W^*)$  é o peso da solução ótima para o TSP do digrafo definido pela matriz  $W^*$ , então

$$P(W^*) = \sum_{i=1}^{\nu} \hat{w}_i + P(W^* - \hat{w}\hat{e}^T)$$

$$(P(W^*) = \sum_{j=1}^{\nu} \bar{w}_j + P(W^* - \hat{e}\bar{w}^T)).$$

**Demonstração.** Tendo em conta que qualquer que seja o vértice  $i$  a solução ótima utiliza necessariamente um dos arcos com cauda em  $i$  (e também um dos arcos com cabeça em  $i$ ), o menor dos pesos destes arcos está necessariamente presente na solução ótima e, conseqüentemente, o resultado verifica-se.  $\square$

<sup>1</sup> Designação inglesa que caracteriza uma família de algoritmos de optimização discreta.

O algoritmo de Little, que a seguir se descreve, em cada iteração determina um minorante para a peso óptimo (com base no Teorema 6.17) e quando termina devolve um ciclo de Hamilton de peso mínimo. Como veremos, o processo iterativo desenvolve-se ao longo de uma árvore de decisão, onde cada vértice corresponde à decisão de considerar a inclusão (ou não) de um dado arco num ciclo em construção e à determinação do minorante obtido com essa decisão. A ramificação (produção de duas novas arestas que conduzem a outros tantos vértices, um dos quais relativo à inclusão de um dado arco no ciclo e outro à situação contrária) deve ser feita a partir do vértice corrente que é o que tem minorante mais favorável de entre os vértices não ramificados. Com esta estratégia, caso se obtenha um ciclo de Hamilton com peso coincidente com o minorante, podemos concluir que esse ciclo é uma solução óptima para o TSP. Ao longo do algoritmo, vamos denotar por  $C$  e  $\bar{C}$  o par de caminhos tais que  $C$  é a parte conhecida (corrente) do ciclo de Hamilton eventualmente a determinar e  $\bar{C}$ , com início no vértice final de  $C$  e fim no vértice inicial de  $C$ , é a parte desconhecida. Quando  $|E(C)| = |V(\bar{C})|$ , vem que  $C$  define um ciclo de Hamilton.

### Algoritmo de Little

*Dados de entrada:* Matriz de pesos modificada  $W^* = (w_{ij}^*)$  de um digrafo  $\vec{G}$ ;

*Resultados de saída:* Ciclo de Hamilton  $C$  de peso mínimo;

1. Fazer  $Min = 0$  e  $E(C) = \emptyset$ .
2. (a) Determinar  $\hat{w}$ , fazer  $W^* \leftarrow W^* - \hat{w}\hat{e}^T$  e  $Min \leftarrow Min + \sum_{i=1}^n \hat{w}_i$ .  
 (b) Determinar  $\bar{w}$ , fazer  $W^* \leftarrow W^* - \bar{w}\bar{e}^T$  e  $Min \leftarrow Min + \sum_{i=1}^n \bar{w}_i$ .
3. Escolher para vértice corrente da árvore de decisão a folha da árvore com minorante ( $Min$ ) mais favorável.
4. Considerar um dos arcos  $ij$  cuja entrada em  $W^*$  tem peso corrente nulo e proceder à ramificação a partir do vértice corrente da árvore de decisão, analisando no ramo (a) a possibilidade deste arco fazer parte do ciclo de Hamilton e o caso contrário no ramo (b):
  - (a) Neste ramo  $E(C) \leftarrow E(C) \cup \{ij\}$  e a linha correspondente ao vértice  $i$  e a coluna correspondente ao vértice  $j$  são eliminadas da matriz  $W^*$ . Adicionalmente, para evitar o aparecimento de subcircuitos, os pesos dos arcos que ligam o vértice final (neste caso  $j$ ) do caminho  $C$  (designado por caminho corrente) a qualquer dos vértices já percorridos tomam o valor  $\infty$ .
  - (b) Neste ramo ( $ij \in E(G) \setminus (E(C) \cup E(\bar{C}))$ ) a entrada  $w_{ij}^*$  passa a ter o valor  $\infty$ .
5. Repetir os passos anteriores (com excepção do primeiro) até se obter um ciclo de Hamilton  $C$ .
6. Devolver o ciclo de Hamilton  $C$  e o minorante  $Min$ .

**Exemplo 6.5.** Vamos resolver o TSP para o digrafo definido pela seguinte matriz de pesos modificada:

$$W^* = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} & \begin{pmatrix} \infty & 3 & 9 & 7 \\ 3 & \infty & 6 & 5 \\ 5 & 6 & \infty & 6 \\ 9 & 7 & 4 & \infty \end{pmatrix} \end{matrix}$$

**Solução.** Depois de fazermos  $Min = 0$  e  $E(C) = \emptyset$ , vamos aplicar os restantes passos e do algoritmo de Little.

- 2 -(a) Tendo em conta que o mínimo da primeira linha de  $W^*$  é 3, o da segunda 3, o da terceira 5 e o da quarta 4, obtém-se

$$W^* \leftarrow \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} & \begin{pmatrix} \infty & 0 & 6 & 4 \\ 0 & \infty & 3 & 2 \\ 0 & 1 & \infty & 1 \\ 5 & 3 & 0 & \infty \end{pmatrix} \end{matrix} \quad (6.15)$$

$$Min \leftarrow 0 + (3 + 3 + 5 + 4) = 15. \quad (6.16)$$

- 2 -(b) Tendo em conta que o mínimo da primeira coluna de  $W^*$  é 0, o da segunda 0, o da terceira 0 e o da quarta 1, obtém-se

$$W^* \leftarrow \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} & \begin{pmatrix} \infty & 0 & 6 & 3 \\ 0 & \infty & 3 & 1 \\ 0 & 1 & \infty & 0 \\ 5 & 3 & 0 & \infty \end{pmatrix} \end{matrix} \quad (6.17)$$

$$Min \leftarrow 15 + (0 + 0 + 0 + 1) = 16. \quad (6.18)$$

3 Neste caso estamos no início da árvore de decisão.

- 4 Considerando o arco (1,2) de peso corrente  $w_{12}^* = 0$ , vamos analisar a ramificação relativa (a) à inclusão deste arco no circuito a construir ou (b) não inclusão.

- (a)  $E(C) \leftarrow \emptyset \cup \{12\} = \{12\}$  e uma vez que a presença de (1,2) impede a utilização de qualquer outro arco com cabeça em 2 e qualquer outro arco com cauda em 1, a primeira linha e a segunda coluna de  $W^*$  devem ser eliminadas. Adicionalmente o peso do arco (2,1) deve passar a ser  $\infty$ .

$$W^* \leftarrow \begin{matrix} & \begin{matrix} 1 & 3 & 4 \end{matrix} \\ \begin{matrix} 2 \\ 3 \\ 4 \end{matrix} & \begin{pmatrix} \infty & 3 & 1 \\ 0 & \infty & 0 \\ 5 & 0 & \infty \end{pmatrix} \end{matrix}$$

2 -(a) Tendo em conta que o mínimo da primeira linha de  $W^*$  corrente é 1, o da segunda 0, o da terceira 0, obtém-se

$$W^* \leftarrow \begin{array}{c} 1 \quad 3 \quad 4 \\ 2 \left( \begin{array}{ccc} \infty & 2 & 0 \\ 0 & \infty & 0 \\ 5 & 0 & \infty \end{array} \right) \end{array} \quad (6.19)$$

$$Min \leftarrow 16 + (1 + 0 + 0) = 17. \quad (6.20)$$

2 -(b) Tendo em conta que o mínimo de qualquer das colunas de  $W^*$  corrente é 0,  $W^*$  mantém-se, bem como o valor de  $Min = 17$ .

(b) Neste caso,  $E(C)$  mantém-se e a entrada relativa ao arco (1,2) passa a  $w_{12}^* = \infty$ , obtendo-se

$$W^* = \begin{array}{c} 1 \quad 2 \quad 3 \quad 4 \\ 1 \left( \begin{array}{cccc} \infty & \infty & 6 & 3 \\ 0 & \infty & 3 & 1 \\ 0 & 1 & \infty & 0 \\ 5 & 3 & 0 & \infty \end{array} \right) \end{array}$$

2 -(a) Tendo em conta que o mínimo da primeira linha de  $W^*$  é 3 e o das restantes é 0, vem

$$W^* \leftarrow \begin{array}{c} 1 \quad 2 \quad 3 \quad 4 \\ 1 \left( \begin{array}{cccc} \infty & \infty & 3 & 0 \\ 0 & \infty & 3 & 1 \\ 0 & 1 & \infty & 0 \\ 5 & 3 & 0 & \infty \end{array} \right) \end{array} \quad (6.21)$$

$$Min \leftarrow 16 + (3 + 0 + 0 + 0) = 19. \quad (6.22)$$

2 -(b) Tendo em conta que o mínimo da segunda coluna de  $W^*$  é 1 e o das restantes é 0, vem

$$W^* \leftarrow \begin{array}{c} 1 \quad 2 \quad 3 \quad 4 \\ 1 \left( \begin{array}{cccc} \infty & \infty & 3 & 0 \\ 0 & \infty & 3 & 1 \\ 0 & 0 & \infty & 0 \\ 5 & 2 & 0 & \infty \end{array} \right) \end{array} \quad (6.23)$$

$$Min \leftarrow 19 + (0 + 1 + 0 + 0) = 20. \quad (6.24)$$

3 Vamos considerar o ponto da árvore de decisão com  $Min$  mais favorável, no qual se tem  $C = \{12\}$  e  $Min = 17$ .

4 Considerando o arco (2,4) de peso corrente  $w_{24}^* = 0$ , vamos analisar a ramificação relativa à inclusão deste arco no circuito a construir.

- (a)  $E(C) \leftarrow \{12\} \cup \{24\} = \{12, 24\}$  e uma vez que a presença de (2,4) impede a utilização de qualquer outro arco com cabeça em 4 e qualquer outro arco com cauda em 2, a linha relativa ao vértice 2 e coluna relativa ao vértice 4 de  $W^*$  devem ser eliminadas. Adicionalmente, o peso do arco (4,1) deve passar a ser  $\infty$ .

$$W^* \leftarrow \begin{array}{cc} & \begin{array}{cc} 1 & 3 \end{array} \\ \begin{array}{c} 3 \\ 4 \end{array} & \begin{pmatrix} 0 & \infty \\ \infty & 0 \end{pmatrix} \end{array}$$

- 2 -(a) Não há qualquer actualização a fazer, uma vez que o mínimo em cada uma das linhas é 0.
- 2 -(b) Não há qualquer actualização a fazer, uma vez que o mínimo em cada uma das colunas é 0.
- 3 Vamos considerar o ponto da árvore de decisão com *Min* mais favorável, no qual se tem  $C = \{12, 24\}$  e  $Min = 17$ .
- 4 Resta-nos escolher os arcos (4,3) e (3,1), os quais não trazem qualquer acréscimo a *Min*, obtendo-se um ciclo de Hamilton com peso igual ao minorante  $Min = 17$  até ao momento obtido, pelo que se trata de um ciclo hamiltoniano de peso mínimo.

□



# Bibliografia

- [1] Antunes, C. H., Tavares, L. V. (ed.). *Casos de Aplicação da Investigação Operacional*, McGraw-Hill, Lisboa (2000).
- [2] Avriel, M. S., Golany, B. (ed.). *Mathematical Programming for Industrial Engineers*. Marcel Dekker, Inc. New York (1996).
- [3] Bazaraa, M. S., Jarvis, J. J., Sherali, H. D. *Linear Programming and Network Flows*. John Wiley & Sons, Inc. New York (1990).
- [4] Bastos, M. e A. G. de Oliveira. *A Matemática nos Fullerenos*, Boletim da SPM, 32 (1995): 37-43.
- [5] Beale, E. M. L. *Cycling in the dual simplex algorithm*, Naval Research Logistics Quaterly, 2 (1955): 269-276.
- [6] Bland, R. G. *New Finite Pivoting Rules for the Simplex Method*, Mathematics of Operations Research, 2 (1977): 103-107.
- [7] Cardoso, D. M. *O Método Simplex Generalizado: Uma Nova Aproximação à Programação Linear*. Tese de Doutoramento, Universidade de Aveiro, Aveiro (1992).
- [8] Charnes, A. *Optimality and degeneracy in linear programming*, Econometrica 20 (1952): 160-177.
- [9] Cunningham, W. H. *A Class of Linear Programs and Convertible to Network Problems*, Operations Research, 31, 2 (1983): 387-390.
- [10] Dantzig, G. B. *Applications of the Simplex Method to a Transportation Problem*, in T. C. Koopmans (ed.), *Activity of Production and Allocation*, John Wiley & Sons, Inc. New York (1951): 359-373.
- [11] Dantzig, G. B. *Computational algorithm of the revised simplex method*, Report RM 1266, The Rand Corporation, Santa Monica, CA (1953).
- [12] Dantzig, G. B., Orden, A., Wolfe, P. *The Generalized Simplex Method for Minimizing a Linear Form Under Linear Inequality Restraints*, Pacific Journal of Mathematics, 5, 2 (1955): 183-195.

- 
- [13] Dantzig, G. B. *Linear Programming and Extensions*, Princeton University Press, Princeton NJ (1963).
- [14] Diestel, R. *Graph Theory*. Springer (Graduate texts in mathematics; 173), New York, 1997.
- [15] Hoffman, A. *Cycling in the simplex algorithm*, Report 2974, National Bureau of Standards, Gaithersburg MD (1953).
- [16] Jogen, H. Th., Meer, K., Triesch, E. *Optimization Theory*, Kluwer Academic Publishers, Boston (2004).
- [17] Lay, S. R. *Convex Sets and their Applications*, John Wiley & Sons, Inc. (1982).
- [18] Marshal, K. T., Suurballe, J. W. *A Note on Cycling in the Simplex Method*, Naval Research Logistics Quarterly, 16 (1969): 121-137.
- [19] Nash, S. G., Sofer, A. *Linear and Nonlinear Programming*, McGraw-Hill, New York (1996).
- [20] Orlin, J. B. *On the Simplex Algorithm for Network and Generalized Networks*, Mathematical Programming, 24 (1985): 166-178.
- [21] Saigal, R. *Linear Programming - A Modern Integrated Analysis*, Kluwer Academic Publishers, Boston (1995).
- [22] Schrijver, A. *Theory of Linear and Integer Programming*, John Wiley & Sons, Inc., Chichester (1986).
- [23] A. Schrijver, *Combinatorial Optimization*, vol. A, B, and C, Springer, Berlin, 2003.
- [24] A. Schrijver, *A Course in Combinatorial Optimization*, John Wiley & Sons (2nd ed.), New York, 2004.
- [25] Sierksma, G. *Linear and Integer Programming: Theory and Practice*, Marcel Dekker, Inc. (1996).
- [26] Tavares, L. V., Oliveira, R. C., Themido, I. H., Correia, F. N. *Investigação Operacional*, McGraw-Hill, Lisboa (1996).
- [27] A. Tucker, *Applied Combinatorics*, CWI, Amsterdam, 1984.
- [28] Urruty, J. B. H., Lemaréchal, C. *Convex Analysis and Minimization Algorithms*, vol. I, Springer, Berlin (1993).

# Índice

- algoritmo
  - de Dijkstra, 124, 125
  - de Hierholzer, 149
  - de Kruskal, 129
  - de Little, 156
  - de Prim, 132
  - para o método dual simplex, 98
- análise de sensibilidade, 101
- análise pós-optimal, 101
- aplicação
  - aditiva, 21
  - homogénea, 21
  - linear, 21
- base de trabalho, 71
- cabeça de um arco, 116
- caminho, 118
- cauda de um arco, 116
- ciclo, 118, 123
- circuito, 118
  - de Euler, 148
- coeficientes tecnológicos, 9
- coloração
  - de arestas, 146
  - de vértices, 145
  - própria de arestas, 146
  - própria de vértices, 145
- coluna pivotal, 47
- combinação afim, 17
- combinação convexa, 17
- componentes
  - básicas, 31
  - não básicas, 31
- comprimento
  - de um caminho, 123
  - de um ciclo, 123
- condições de não negatividade, 9
- cone, 22
  - convexo, 22
  - polar negativo, 23
- conjunto
  - convexo, 17
  - dependente afim, 18
  - independente afim, 18
- conjunto independente
  - de arestas, 139
  - de vértices, 139
- corte, 122
- cortes de Gomory, 110
- digrafo, 116
- dimensão
  - de um conjunto, 19
  - de uma variedade linear, 19
- directão
  - distinta, 29
  - extrema, 29
- dualidade, 89
- emparelhamento, 139
  - máximo, 139
  - perfeito, 139
- estável, 139
  - máximo, 139
- exemplo
  - Beale de entrada em ciclo, 65
  - de aplicação do método dual simplex, 99
- floresta, 127
- funcional linear, 21
- função objectivo, 8, 9

- grafo, 115  
   bipartido, 127  
   bipartido completo, 127  
   complementar, 116  
   completo, 116  
   conexo, 118  
   de Hamilton, 152  
   linha, 120  
   nulo, 116  
   orientado, 116  
   regular, 117  
   simples, 115
- grafos  
  cúbicos, 117
- grau de um vértice, 117
- hiperplano, 21
- interpretação económica  
  da variável de desvio ou folga, 10
- invólucro  
  afim, 19  
  convexo, 19
- Lema  
  de Farkas (1<sup>a</sup> versão), 26  
  de Farkas (2<sup>a</sup> versão), 28
- Leonhard Euler, 115
- linha  
  de custos reduzidos, 47  
  pivotal, 47
- maior  
  grau dos vértices, 123
- matriz  
  de adjacência de um digrafo, 118  
  de adjacência de um grafo, 118  
  de incidência de um digrafo, 120  
  de incidência de um grafo, 120  
  laplaciana, 121  
  laplaciana sem sinal, 122  
  totalmente unimodular, 138  
  unimodular, 138
- meio espaço, 21
- menor  
  grau dos vértices, 123
- Método  
  das duas fases, 55  
  do big<sup>m</sup>-M, 58
- método  
  das duas fases, 55  
  de decomposição de Dantzig-Wolfe,  
    70, 77  
  do big-M, 55  
  dual simplex, 70, 95  
  primal-dual, 70  
  simplex, 52  
  simplex para variáveis limitadas,  
    70  
  simplex revisto, 61
- numero  
  de estabilidade, 139
- número  
  de independência, 139
- número cromático, 145
- operações de pivotação, 49, 51
- partições básicas admissíveis adjacente,  
  71
- passeio, 118
- pivot, 47
- poliedro convexo, 22
- politopo, 21
- ponto extremo, 28
- problema  
  da dieta óptima, 6  
  das sete pontes de Königsberg, 115  
  de gestão de “staff”, 7  
  de inventário, 8  
  de optimização da produção, 7  
  de transportes, 6  
  do caixeiro viajante (TSP), 154  
  do carteiro chinês, 150  
  dual, 89  
  primal, 89
- problema auxiliar  
  do método de Dantzig-Wolfe, 77,  
    78
- problema principal  
  do método de Dantzig-Wolfe, 77

- programa linear, 8
  - na forma canónica, 9
  - na forma padrão, 8
- programação paramétrica, 102
- propriedade
  - de monotonicidade da regra de Bland, 68
- quadro simplex
  - completo, 49
  - reduzido, 49
- região admissível, 10
- regra
  - de Bland, 67
  - em estrela, 51, 62, 63
- restrições, 9
- restrições de caixa, 70
- semi-espaço, 21
- simplex, 21
- solução
  - básica admissível degenerada, 33
  - básica degenerada, 33
  - admissível, 10
  - básica, 30
  - básica admissível, 31
  - ótima, 10
- solução básica
  - em problemas com variáveis limitadas, 70
- solução básica admissível
  - em problemas com variáveis limitadas, 71
- solução básica degenerada
  - em problemas com variáveis limitadas, 71
- solução básica não degenerada
  - em problemas com variáveis limitadas, 71
- subespaço afim, 19
- subgrafo, 116
  - induzido, 116
- supergrafo, 116
- Teorema
  - da complementaridade, 92
  - da complementaridade estrita, 93
  - da dualidade forte, 91
  - da dualidade fraca, 90
  - da representação de poliedros convexos, 36
  - da separação estrita, 24
  - de Bolzano-Weierstrass, 25
  - de Caratheodory, 20
  - de Helly, 39
  - de Radon, 39
- trajecto, 118
  - de Euler, 148
- variedade linear, 19
- variável
  - de decisão, 10
  - de desvio, 10
  - de folga, 10
  - não restringida, 11
- vector
  - característico, 122
- vizinhança, 118
- vizinho, 123
- vértice, 21
- William R. Hamilton, 115
- árvore, 127
  - abrangente, 128
  - de suporte, 128
- índice cromático, 146