

Apontamentos das aulas teóricas

Programação Matemática

1º Semestre de 2008/2009

1. NOTAÇÕES E PRELIMINARES

\mathbb{R} designa o conjunto dos números reais.

\mathbb{R}^n designa o espaço linear real canónico de dimensão n .

Um elemento x de \mathbb{R}^n será visto como um vector coluna (i.e. uma matriz $n \times 1$):

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

$\mathbb{R}^{m,n}$ designa o espaço das matrizes $m \times n$.

$A^T \in \mathbb{R}^{n,m}$ designa a transposta da matriz $A \in \mathbb{R}^{m,n}$.

Dados dois vectores x e y em \mathbb{R}^n , $x^T y$ designa o produto interno usual em \mathbb{R}^n . $\|x\| := \sqrt{x^T x}$ define a normal euclideana do vector x .

Dado dois conjuntos A e B em \mathbb{R}^n define-se

$$A + B := \{a + b : a \in A, b \in B\} \text{ (soma de Minkowski);}$$

$$\lambda A := \{\lambda a : a \in A\}.$$

Ao conjunto $A + \{x\}$ (muitas vezes denotado $A + x$) chama-se *conjunto A trasladado por x* .

Propriedades:

Para quaisquer $A, B, C \in \mathbb{R}^n$ e $\lambda, \mu \in \mathbb{R}$ temos:

(i) $A + B = B + A$;

(ii) $(A + B) + C = A + (B + C)$;

(iii) $\lambda(\mu A) = (\lambda\mu)A$;

(iv) $\lambda(A + B) = \lambda A + \lambda B$.

Nota: Em geral $(\lambda + \mu)A \neq \lambda A + \mu A$.

2. CONJUNTOS AFINS, CONVEXOS E CONES EM \mathbb{R}^n

Definições básicas.

Definição 2.1. Um conjunto S em \mathbb{R}^n diz-se:

um *conjunto afim* se $x_1, x_2 \in S$ e $\lambda \in \mathbb{R} \Rightarrow \lambda x_1 + (1 - \lambda)x_2 \in S$;

um *conjunto convexo* se $x_1, x_2 \in S$ e $\lambda \in [0, 1] \Rightarrow \lambda x_1 + (1 - \lambda)x_2 \in S$;

um *cone* se $x \in S$ e $\lambda \geq 0 \Rightarrow \lambda x \in S$;

um *cone convexo* se $x_1, x_2 \in S$ e $\lambda_1, \lambda_2 \in \mathbb{R}_0^+ \Rightarrow \lambda_1 x_1 + \lambda_2 x_2 \in S$.

Definição 2.2. Uma combinação linear $x = \lambda_1 x_1 + \dots + \lambda_m x_m$ de elementos x_1, \dots, x_m de \mathbb{R}^n diz-se:

uma *combinação afim* de x_1, \dots, x_m se $\lambda_1 + \dots + \lambda_m = 1$;

uma *combinação convexa* de x_1, \dots, x_m se $\lambda_1 + \dots + \lambda_m = 1$ e $\lambda_i \geq 0$ para $i = 1, \dots, m$;

uma *combinação cónica* de x_1, \dots, x_m se $\lambda_i \geq 0$ para $i = 1, \dots, m$.

Proposição 2.3. Um conjunto é afim se e só se for fechado para combinações afins;

Um conjunto é convexo se e só se for fechado para combinações convexas;

Um conjunto é um cone convexo se e só se for fechado para combinações cónicas.

Dem. Exercício. □

O próximo resultado permite caracterizar os conjuntos afins em \mathbb{R}^n .

Proposição 2.4. $A \subseteq \mathbb{R}^n$ é afim se e só se existem um ponto $x_0 \in \mathbb{R}^n$ e um único subespaço linear $L \subseteq \mathbb{R}^n$ tal que $A = L + x_0$.

Dem. Exercício. □

Deste modo podemos definir a dimensão de um espaço afim da seguinte maneira.

Definição 2.5. Define-se a *dimensão* de um espaço afim A como sendo a dimensão do único espaço linear L tal que $A = L + x_0$ para algum $x_0 \in \mathbb{R}^n$.

Exemplo 2.6. (definição) Um *hiperplano* em \mathbb{R}^n é conjunto afim de dimensão $n - 1$.

O próximo resultado dá-nos outra caracterização dos conjuntos afins em \mathbb{R}^n .

Proposição 2.7. Um conjunto $X \subseteq \mathbb{R}^n$ é afim se e só se existem uma matriz $A \in \mathbb{R}^{m,n}$ e um vector $b \in \mathbb{R}^m$ tal que $X = \{x \in \mathbb{R}^n : Ax = b\}$.

Além disso $\text{car}A + \dim X = n$.

Dem. Deixamos como exercício verificar que um conjunto da forma $X = \{x \in \mathbb{R}^n : Ax = b\}$ (onde $A \in \mathbb{R}^{m,n}$ é uma matriz e $b \in \mathbb{R}^m$ é um vector) é um espaço afim.

Seja X um espaço afim. Pela proposição 2.4, $X = L + x_0$ onde L é um subespaço linear de \mathbb{R}^n e x_0 é um ponto de \mathbb{R}^n . Seja $L^\perp = \{y \in \mathbb{R}^n : y^T x = 0, \forall x \in L\}$ o complemento ortogonal de L . Seja $\{b_1, \dots, b_k\}$ uma base de L^\perp , então

$$L = \left\{ x \in \mathbb{R}^n : \begin{bmatrix} b_1^T \\ \vdots \\ b_k^T \end{bmatrix} x = 0 \right\} \quad (\text{pois } (L^\perp)^\perp = L).$$

Definindo

$$A = \begin{bmatrix} b_1^T \\ \vdots \\ b_k^T \end{bmatrix}$$

temos que

$$x \in X \Leftrightarrow x - x_0 \in L \Leftrightarrow A(x - x_0) = 0 \Leftrightarrow Ax = Ax_0$$

seja $b = Ax_0$ então $X = \{x \in \mathbb{R}^n : Ax = b\}$.

Observando que $\text{car}A = \dim L^\perp$, a identidade $\text{car}A + \dim X = n$ resulta dos factos conhecidos de Álgebra Linear de que $L \cap L^\perp = \{0\}$ e $\mathbb{R}^n = L \oplus L^\perp$. □

Note-se que no caso particular de um hiperplano este pode ser definido como um conjunto da forma

$$\{x \in \mathbb{R}^n : a^T x = b\} \quad \text{para algum } a \in \mathbb{R}^n \text{ não-nulo e } b \in \mathbb{R}$$

Invólucros afins, convexos e cónicos.

Proposição 2.8. Qualquer intersecção de uma famílias de $\left\{ \begin{array}{l} \text{conjuntos afins} \\ \text{conj. convexos} \\ \text{cones convexos} \end{array} \right\}$ é um $\left\{ \begin{array}{l} \text{conjunto afim} \\ \text{conj. convexo} \\ \text{cone convexo} \end{array} \right\}$.

Dem. Exercício. □

Definição 2.9. Dado um conjunto $S \subseteq \mathbb{R}^n$, chama-se *invólucro* $\left\{ \begin{array}{l} \text{afim} \\ \text{convexo} \\ \text{cónico} \end{array} \right\}$ ao menor $\left\{ \begin{array}{l} \text{conjunto afim} \\ \text{conj. convexo} \\ \text{cone convexo} \end{array} \right\}$

que contém S e é denotado por $\left\{ \begin{array}{l} \text{aff}(S) \\ \text{conv}(S) \\ \text{cone}(S) \end{array} \right\}$.

A proposição 2.8 mostra que faz sentido falar em menor conjunto (afim, convexo, cone) que contém S . De facto, conduz à seguinte observação.

Observação 2.10. $\left\{ \begin{array}{l} \text{aff}(S) \\ \text{conv}(S) \\ \text{cone}(S) \end{array} \right\}$ é igual à intersecção de todos os $\left\{ \begin{array}{l} \text{conjuntos afins} \\ \text{conj. convexos} \\ \text{cones convexos} \end{array} \right\}$ que contém S .

O próximo resultado dá-nos uma caracterização dos invólucros afins, convexos ou cónicos.

Proposição 2.11. *Seja S em conjunto em \mathbb{R}^n temos que:*

$\text{aff}(S)$ é o conjunto de todas as combinações afins de elementos de S .

$\text{conv}(S)$ é o conjunto de todas as combinações convexas de elementos de S .

$\text{cone}(S)$ é o conjunto de todas as combinações cónicas de elementos de S .

Dem. Exercício. □

Definição 2.12. Define-se a *dimensão* de um conjunto S como sendo a dimensão do seu invólucro afim (i.e. $\dim(S) := \dim(\text{aff}(S))$).

Definição 2.13. Um conjunto (finito) de pontos em \mathbb{R}^n , $\{x_1, \dots, x_m\}$, diz-se *afimemente independente* se a sua dimensão coincide com a sua cardinalidade (i.e. $\dim\{x_1, \dots, x_m\} = m$).

A independência afim pode-se deduzir da independência dos pontos transladados como mostra o seguinte resultado.

Proposição 2.14. *Um conjunto (finito) de pontos em \mathbb{R}^n , $\{x_1, \dots, x_m\}$ é afimemente independente se e só se $\{x_2 - x_1, \dots, x_m - x_1\}$ é linearmente independente.*

Dem. Exercício. □

Os próximos dois resultados dão-nos uma forma alternativa de construir conjuntos afins ou convexos.

Proposição 2.15. (1) *Se $A \subseteq \mathbb{R}^n$ e $B \subseteq \mathbb{R}^m$ são $\left\{ \begin{array}{l} \text{conjuntos afins} \\ \text{conj. convexos} \\ \text{cones convexos} \end{array} \right\}$ então o conjunto $A \times B \subseteq \mathbb{R}^{n+m}$ é um $\left\{ \begin{array}{l} \text{conjunto afim} \\ \text{conj. convexo} \\ \text{cone convexo} \end{array} \right\}$.*

(2) *Qualquer transformação linear $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ transforma $\left\{ \begin{array}{l} \text{conjuntos afins} \\ \text{conj. convexos} \\ \text{cones convexos} \end{array} \right\}$ em $\left\{ \begin{array}{l} \text{conjuntos afins} \\ \text{conj. convexos} \\ \text{cones convexos} \end{array} \right\}$.*

(3) *Qualquer transformação afim (i.e. a composta de uma transformação linear por um translação) $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ transforma $\left\{ \begin{array}{l} \text{conjuntos afins} \\ \text{conj. convexos} \end{array} \right\}$ em $\left\{ \begin{array}{l} \text{conjuntos afins} \\ \text{conj. convexos} \end{array} \right\}$.*

Dem. Exercício. □

Corolário 2.16. *Se $A, B \subseteq \mathbb{R}^n$ são $\left\{ \begin{array}{l} \text{conjuntos afins} \\ \text{conj. convexos} \\ \text{cones convexos} \end{array} \right\}$ então o conjunto $\lambda_1 A + \lambda_2 B \subseteq \mathbb{R}^n$ é*

um $\left\{ \begin{array}{l} \text{conjunto afim} \\ \text{conj. convexo} \\ \text{cone convexo} \end{array} \right\}$ para quaisquer $\lambda_1, \lambda_2 \in \mathbb{R}$.

Dem. Exercício. □

Definição 2.17. Chama-se *polígono* ao invólucro convexo de um número finito de pontos em \mathbb{R}^n .

Definição 2.18. Chama-se *cone finitamente gerado* ao invólucro cónico de um número finito de pontos em \mathbb{R}^n .

Exemplo 2.19. (definição) Uma *semi-recta* em \mathbb{R}^n com extremo na origem é um cone gerado por único vector não-nulo.

A próxima definição é de certo modo o análogo do conceito de ortogonalidade para cones.

Definição 2.20. Para um cone convexo K em \mathbb{R}^n define-se o seu *cone polar* K° como sendo o conjunto:

$$K^\circ := \{y \in \mathbb{R}^n : y^T x \leq 0, \forall x \in K\}.$$

Proposição 2.21. Para qualquer conjunto K (cone convexo ou não), K° é um cone convexo e fechado. Se K é um subespaço linear de \mathbb{R}^n então $K^\circ = K^\perp$.

Dem. Exercício. □

Observemos que se um cone convexo K é invólucro cônico de um conjunto S então $K^\circ = \{y \in \mathbb{R}^n : y^T x \leq 0, \forall x \in S\}$. Desta observação resulta a seguinte proposição.

Proposição 2.22. Se K é um cone finitamente gerado então o seu cone polar é o conjunto solução de um número finito de desigualdades:

$$K^\circ = \{y \in \mathbb{R}^n : c_i^T y \leq 0, \forall i=1, \dots, m\}.$$

Teorema de Carathéodory. O próximo resultado, conhecido como **teorema de Carathéodory**, diz-nos que qualquer combinação convexa pode ser reduzida a uma combinação convexa de pontos afimemente independentes (e algo análogo para combinações cónicas).

Teorema 2.23. (1) Seja $C = \text{cone}(G)$ para algum $G \subseteq \mathbb{R}^n$ e seja $x \in C$. Então x pode ser escrito com uma combinação cónica de $m \leq n$ vectores linearmente independentes de G .

(2) Seja $C = \text{conv}(G)$ para algum $G \subseteq \mathbb{R}^n$ e seja $x \in C$. Então x pode ser escrito com uma combinação convexa de $m \leq n + 1$ vectores afimemente independentes de G .

Dem. (1) Se $x \in \text{cone}(G)$ então $x = \sum_{j=1}^m \lambda_j a_j$, com $a_1, \dots, a_m \in G$ e $\lambda_1, \dots, \lambda_m > 0$.

Se os a_i 's são linearmente independentes então $m \leq n$ e o resultado fica provado.

Se a_1, \dots, a_m são linearmente dependentes então existe $(\mu_1, \dots, \mu_m) \neq (0, \dots, 0)$ tal que $\sum_{j=1}^m \mu_j a_j = 0$. Seja

$$\Delta^* = \max\{\Delta \geq 0 : \lambda_j - \Delta \mu_j \geq 0, \forall j=1, \dots, m\} = \min\left\{\frac{\lambda_j}{\mu_j} : \mu_j > 0, j = 1, \dots, m\right\}$$

(podemos assumir a existência de pelo menos um $\mu_i > 0$ logo $0 < \Delta^* < \infty$). Seja $\lambda'_j := \lambda_j - \Delta^* \mu_j \geq 0$, então $x = \sum_{j=1}^m \lambda'_j a_j$ e existe um índice i tal que $\lambda'_i = 0$, logo podemos reduzir o número de a_j 's.

Aplicando este método recursivamente chega-se a subconjunto de $\{a_1, \dots, a_m\}$ linearmente independente que gera x .

(2) Se $x \in \text{conv}(G)$ então $x = \sum_{j=1}^m \lambda_j a_j$, com $a_1, \dots, a_m \in G$, $\sum_{j=1}^m \lambda_j = 1$ e $\lambda_1, \dots, \lambda_m > 0$. Considerando o ponto $(x, 1) \in \mathbb{R}^{n+1}$ temos que $(x, 1) = \sum_{j=1}^m \lambda_j (a_j, 1)$, com $(a_1, 1), \dots, (a_m, 1) \in G \times \{1\}$ e $\lambda_1, \dots, \lambda_m > 0$. Ou seja, $(x, 1) \in \text{cone}(G \times \{1\})$ logo, por (1), temos $(x, 1) = \sum_{k=1}^{m'} \lambda'_{j_k} (a_{j_k}, 1)$, com $(a_{j_1}, 1), \dots, (a_{j_{m'}}, 1) \in G \times \{1\}$ linearmente independentes, $\lambda'_{j_1}, \dots, \lambda'_{j_{m'}} > 0$ e $m' \leq n + 1$. Tal implica que x é uma combinação convexa de $a_{j_1}, \dots, a_{j_{m'}}$ e estes são afimemente independentes. □

Definição 2.24. Um *simplex* é o invólucro convexo de um conjunto de vectores de \mathbb{R}^n afimemente independentes.

Um *simplex generalizado* é o invólucro cônico de um conjunto de vectores de \mathbb{R}^n linearmente independentes.

Proposição 2.25. (1) Qualquer polítopo é a união de um número finito de simplices.

(2) Qualquer cone finitamente gerado é a união de um número finito de simplices generalizados.

Dem. (1) Consideremos o polítopo $P = \text{conv}(\{a_1, \dots, a_m\}) \subset \mathbb{R}^n$. Definimos $M = \{1, \dots, m\}$ e seja \mathcal{F} a família de todos subconjuntos J para os quais $\{a_j\}_{j \in J}$ é afimemente independente. É óbvio que \mathcal{F} é finito. Para cada $J \in \mathcal{F}$ seja $P^J = \text{conv}(\{a_j : j \in J\})$ que, note-se, é um simplex.

Vejamus que $P = \cup_{J \in \mathcal{F}} P^J$ (portanto P é uma união finita de simplices como queremos provar). A inclusão \supseteq resulta trivialmente do facto de $P^J \subseteq P$ para qualquer $J \in \mathcal{F}$. E a inclusão recíproca resulta, como corolário, do teorema de Carathéodory 2.23.

(2) A demonstração da segunda parte da proposição é *mutatis mutandis* igual à da primeira. □

Corolário 2.26. Qualquer cone finitamente gerado é fechado e qualquer polítopo é compacto.

Dem. Pelo resultado anterior basta mostrar que qualquer simplex generalizado é fechado e qualquer simplex é compacto. O que é deixado como exercício. □

Separação de conjuntos convexos. Vimos atrás que um hiperplano é um conjunto da forma $H_=(a, b) = \{x \in \mathbb{R}^n : a^T x = b\}$. Este divide o espaço \mathbb{R}^n em dois semi-espacos abertos $H_<(a, b) = \{x \in \mathbb{R}^n : a^T x < b\}$ e $H_>(a, b) = \{x \in \mathbb{R}^n : a^T x > b\}$, sendo os seus respectivos fechos topológicos, $H_{\leq}(a, b) = \{x \in \mathbb{R}^n : a^T x \leq b\}$ e $H_{\geq}(a, b) = \{x \in \mathbb{R}^n : a^T x \geq b\}$, designados de semi-espacos fechados.

Teorema 2.27. *Seja C um conjunto convexo fechado e não-vazio e seja $p \notin C$. Então existem $a \in \mathbb{R}^n \setminus \{0\}$ e $\varepsilon > 0$ tais que*

$$a^T x \leq a^T p - \varepsilon \quad \forall x \in C$$

Além disso, existe um único ponto $x_0 \in C$ tal que $\|p - x_0\| = \text{dist}(C, p) := \inf\{\|x - p\| : x \in C\}$.

Dem. A existência de um ponto $x_0 \in C$ tal que $\|p - x_0\| = \text{dist}(C, p)$ resulta do teorema de Weierstrass. De facto, $\text{dist}(C, p) = \text{dist}(C', p)$ onde $C' = \{x \in C : \|x - p\| \leq \|c - p\|\}$ para qualquer $c \in C$ (existe pois C é não-vazio), e sendo C' compacto e a função distância a p , $d: C \rightarrow \mathbb{R}$, $x \mapsto \|x - p\|$, contínua temos existência de x_0 resulta do teorema de Weierstrass.

Seja $x \in C$ e $t \in]0, 1[$. Visto que C é convexo, $(1-t)x_0 + tx \in C$ e portanto $\|(1-t)x_0 + tx - p\| \geq \|x_0 - p\|$. Elevando ao quadrado ambos os termos da desigualdade obtemos:

$$\begin{aligned} & \|(1-t)x_0 + tx - p\|^2 \geq \|x_0 - p\|^2 \\ \Leftrightarrow & ((1-t)x_0 + tx - p)^T((1-t)x_0 + tx - p) \geq (x_0 - p)^T(x_0 - p) \\ \Leftrightarrow & ((1-t)x_0 + tx - p)^T((1-t)x_0 + tx - p) \geq (x_0 - p)^T(x_0 - p) \\ \Leftrightarrow & \|x_0 - p\|^2 + 2t(x_0 - p)^T(x - x_0) + t^2\|x - x_0\|^2 \geq \|x_0 - p\|^2 \\ \Leftrightarrow & 2(x_0 - p)^T(x - x_0) + t\|x - x_0\|^2 \geq 0 \\ \Rightarrow & (x_0 - p)^T(x - x_0) \geq 0 \text{ fazendo } t \rightarrow 0^+ \\ \Leftrightarrow & (x_0 - p)^T x \geq (x_0 - p)^T(p + x_0 - p) \\ \Leftrightarrow & (x_0 - p)^T x \geq (x_0 - p)^T p + \|x_0 - p\|^2 \\ \Leftrightarrow & (p - x_0)^T x \leq (p - x_0)^T p - \|x_0 - p\|^2 \end{aligned}$$

Definindo $a := p - x_0$ e $\varepsilon := \|x_0 - p\|^2$, $a \neq 0$ e $\varepsilon > 0$ pois $p \notin C$ (logo $p \neq x_0$), temos a desigualdade enunciada no teorema.

A unicidade do ponto x_0 deriva do facto de o interior do segmento de recta que une dois pontos situados na superfície de uma esfera estar contido no interior desta. \square

O que este resultado basicamente diz é que para qualquer ponto exterior a um conjunto convexo existe um semi-espaco fechado contendo o conjunto convexo e excluindo o ponto em questão. Desta observação segue-se facilmente o seguinte corolário.

Corolário 2.28. *Qualquer conjunto convexo e fechado C é uma intersecção de semi-espacos fechados:*

$$C = \bigcap_{i \in I} H_{\leq}(a_i, b_i) \text{ com } H_{\leq}(a_i, b_i) := \{x \in \mathbb{R}^n : a_i^T x \leq b_i\}.$$

Dem. Exercício. \square

Corolário 2.29. *Seja C um cone convexo, fechado e não-vazio e seja $p \notin C$. Então existe $a \in \mathbb{R}^n \setminus \{0\}$ tal que $a^T x \leq 0 < a^T p$ para qualquer $x \in C$:*

Dem. Pelo teorema 2.27 sabemos que existem $a \in \mathbb{R}^n \setminus \{0\}$ e $\varepsilon > 0$ tais que

$$a^T x \leq a^T p - \varepsilon \quad \forall x \in C$$

logo, para qualquer $\lambda > 0$, temos

$$\lambda a^T x = a^T(\lambda x) \leq a^T p - \varepsilon \quad \forall x \in C$$

Como $a^T p - \varepsilon$ é constante e λ pode ser tão grande quanto se queira, esta última desigualdade só é válida se

$$a^T x \leq 0 \quad \forall x \in C$$

Além disso, $a^T p > a^T x \quad \forall x \in C$ o que implica que $a^T p > 0$ pois qualquer cone contem o vector nulo ($0 \in C$). \square

A próxima proposição resulta como corolário deste último resultado.

Proposição 2.30. *Seja K um cone convexo, então $K^{oo} = cl(K)$ (i.e. o cone polar do cone polar de K é o seu fecho topológico de K).*

Dem. É fácil ver que $K \subseteq K^{oo}$ e como K^{oo} é fechado (pela proposição 2.21). basta então ver que $K^{oo} \setminus cl(K) = \emptyset$.

Suponhamos por absurdo que existe $p \in K^{oo} \setminus cl(K)$. Então, como $p \notin cl(K)$, temos, pelo corolário 2.29 (note-se que $cl(K)$ é um cone convexo fechado), que existe $a \in \mathbb{R}^n \setminus \{0\}$ tal que $a^T x \leq 0 < a^T p$ para qualquer $x \in cl(K)$ (em particular para qualquer $x \in K$) o que implica que $a \in K^o$. Ora $a \in K^o$ contradiz o facto de $a^T p > 0$ e $p \in K^{oo}$ ($a \in K^o$ e $p \in K^{oo} \Rightarrow a^T p \leq 0$). \square

Definição 2.31. Sejam C_1 e C_2 conjuntos não-vazios, dizemos que um hiperplano $H_-(a, b) := \{x \in \mathbb{R}^n : a^T x = b\}$:

- (1) *separa fracamente* C_1 e C_2 se $C_1 \subseteq H_{\leq}(a, b)$ e $C_2 \subseteq H_{\geq}(a, b)$ (ou vice-versa);
- (2) *separa fortemente* C_1 e C_2 se separa fracamente C_1 e C_2 e $dist(C_i, H_-(a, b)) > 0$ para $i = 1, 2$;
- (3) *separa propriamente* C_1 e C_2 se separa fracamente C_1 e C_2 e $C_1 \cup C_2 \not\subseteq H_-(a, b)$.

Observação 2.32. Separação fraca \Rightarrow separação própria \Rightarrow separação forte.

Vimos no teorema 2.27 que um conjunto convexo C é separado fortemente de qualquer ponto exterior a C , $p \in ext(C)$. Para pontos fronteiros não existe separação forte mas existe separação fraca como é enunciado em seguida.

Proposição 2.33. *Seja C um conjunto convexo e $p \in \mathbb{R}^n$ um ponto fronteiro de C (i.e. $p \in frontC$). Então existe $a \in \mathbb{R}^n$ não-nulo tal que $a^T x \leq a^T p$ para todo o $x \in C$.*

Dem. Exercício (1ª ficha). \square

O próximo resultado dá condições equivalentes á separação forte entre dois conjuntos convexos não-vazios.

Proposição 2.34. *Seja C_1 e C_2 dois conjuntos convexos não-vazios. Então as seguintes afirmações são equivalentes:*

- (i) C_1 e C_2 são separados fortemente por um hiperplano;
- (ii) Existe um vector não-nulo $a \in \mathbb{R}^n$ tal que $\sup_{x \in C_1} a^T x < \inf_{y \in C_2} a^T y$;
- (iii) $dist(C_1, C_2) := \inf\{\|x - y\| : x \in C_1 \text{ e } y \in C_2\} > 0$.

Dem. Exercício. \square

3. POLIÉDROS, SISTEMAS LINEARES E PROGRAMAÇÃO LINEAR

Sistemas lineares e poliédros.

Definição 3.1. Um *sistema linear* é um conjunto finito de desigualdades lineares e pode-se escrever da forma $Ax \leq b$ onde $A \in \mathbb{R}^{m,n}$ e $b \in \mathbb{R}^m$.

Observação 3.2. Uma equação linear pode ser decomposta em duas desigualdade ($a^T x = \beta \Leftrightarrow a^T x \leq \beta$ e $a^T x \geq \beta$). Deste modo, um conjunto finito de equações lineares é também um sistema linear.

Definição 3.3. Um sistema linear $Ax \leq b$ diz-se *consistente* se o conjunto $\{x \in \mathbb{R}^n : Ax \leq b\}$ é não-vazio.

Definição 3.4. A um conjunto do tipo $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ (para um sistema linear $Ax \leq b$) chama-se um *poliédro*.

Proposição 3.5. *Um poliédro é um conjunto convexo fechado e é a intersecção de um número finito de semi-espacos fechados.*

Definição 3.6. Dois sistemas lineares $Ax \leq b$ e $A'x \leq b'$ dizem-se *equivalentes* se tiverem o mesmo conjunto de soluções:

$$\{x \in \mathbb{R}^n : Ax \leq b\} = \{x \in \mathbb{R}^n : A'x \leq b'\}$$

Por exemplo se o conjunto $\{x \in \mathbb{R}^n : Ax \leq b\}$ está contido num semi-espaço $H_{\leq}(c, d) = \{x \in \mathbb{R}^n : c^T x \leq d\}$ podemos juntar a desigualdade $c^T x \leq d$ sem alterar o conjunto de soluções de $Ax \leq b$. Um caso particular em que isto acontece é quando temos $y \in \mathbb{R}^m$, $y \geq 0$ e tomamos $c^T = y^T A$ e $d = y^T b$. Mais geralmente, dado um sistema $Ax \leq b$ com $A \in \mathbb{R}^{m,n}$ e $b \in \mathbb{R}^m$, se tivermos uma matriz $M \in \mathbb{R}^{k,m}$ com entradas não negativas que possui uma inversa à esquerda $M_{esq}^{-1} \in \mathbb{R}^{m,k}$ com entradas não negativas ($M_{esq}^{-1} M = I$) então os sistemas $Ax \leq b$ e $MAx \leq Mb$ são equivalentes.

Alternativa de Fredholm e Lema de Farkas.

Teorema 3.7. *Seja $A \in \mathbb{R}^{m,n}$ (com colunas $a_1, \dots, a_n \in \mathbb{R}^m$) e $b \in \mathbb{R}^m$. O sistema $Ax = b$ tem solução se e só se $y^T b = 0$ para todo o $y \in \mathbb{R}^m$ que satisfaça $y^T a_1 = \dots = y^T a_n = 0$.*

Basicamente o que este resultado diz é que temos sempre uma e uma só de duas coisas: ou $Ax = b$ tem solução x ou $A^T y = 0$ tem solução y com $y^T b \neq 0$. Daí o nome **Alternativa de Fredholm**.

Outra forma de o escrever é

$$\exists_{x \in \mathbb{R}^n} : Ax = b \Leftrightarrow (y^T A = 0 \Rightarrow y^T b = 0)$$

Segue-se a demonstração do teorema.

Dem. Seja $L := \{Ax \in \mathbb{R}^m : x \in \mathbb{R}^n\}$, temos que

$$(y^T A = 0 \Rightarrow y^T b = 0) \Leftrightarrow (y \in L^\perp \Rightarrow y^T b = 0) \Leftrightarrow b \in (L^\perp)^\perp \Leftrightarrow b \in L \Leftrightarrow \exists_{x \in \mathbb{R}^n} : Ax = b$$

□

O próximo resultado, conhecido por **Lema de Farkas**, generaliza a alternativa de Fredholm para sistema de equações lineares restritas a soluções não-negativas.

Lema 3.8. *Seja $A \in \mathbb{R}^{m,n}$ (com colunas $a_1, \dots, a_n \in \mathbb{R}^m$) e $b \in \mathbb{R}^m$. O sistema $Ax = b, x \geq 0$ tem solução se e só se $y^T b \geq 0$ para todo o $y \in \mathbb{R}^m$ que satisfaça $y^T A \geq 0$.*

Dem. Dizer que $Ax = b$ tem solução $x \geq 0$ é o mesmo que dizer que $b \in K := \text{cone}(\{a_1, \dots, a_n\})$. Como, pelo corolário 2.26, K é fechado temos, pela proposição 2.30, que $K = K^{\circ\circ}$. Portanto a existência de solução do sistema $Ax = b, x \geq 0$ equivale a dizer que $b \in K^{\circ\circ}$, ou seja, $b^T y \leq 0$ sempre que $y \in K^\circ$, que é como quem diz $y^T b \leq 0$ para qualquer $y \in \mathbb{R}^m$ que satisfaça $y^T A \leq 0$. Substituindo y pelo seu simétrico, chegamos á segunda condição do lema, $y^T b \geq 0$ para todo o $y \in \mathbb{R}^m$ que satisfaça $y^T A \geq 0$. □

Seque-se, como corolário, uma versão alternativa do lema Farkas.

Corolário 3.9. *Seja $A \in \mathbb{R}^{m,n}$ e $b \in \mathbb{R}^m$. O sistema $Ax \leq b$ tem solução se e só se $y^T b \geq 0$ para todo o $y \in \mathbb{R}^m$ que satisfaça $y^T A = 0$ e $y \geq 0$.*

Dem. A existência de um $x \in \mathbb{R}^n$ que satisfaça $Ax \leq b$ é equivalente à existência de $x_1, x_2, x_3 \in \mathbb{R}^n$ que satisfaçam $Ax_1 - Ax_2 + x_3 = b$ com $x_1 \geq 0$, $x_2 \geq 0$ e $x_3 \geq 0$. Logo, o sistema $Ax \leq b$ tem solução se e só se o sistema $\tilde{A}\tilde{x} = b, \tilde{x} \geq 0$ tem solução (onde $\tilde{A} \in \mathbb{R}^{m,3n}$ é a matriz com três blocos $\tilde{A} = [A | -A | I]$ e $\tilde{x} \in \mathbb{R}^{3n}$ é o vector com três blocos $\tilde{x} = [x_1^T | x_2^T | x_3^T]^T$). Este último sistema tem, pelo lema de Farkas (lema 3.8), solução se e só se $y^T b \geq 0$ para todo o $y \in \mathbb{R}^m$ que satisfaça $y^T \tilde{A} \geq 0$. Ora satisfazer $y^T \tilde{A} \geq 0$ é equivalente a satisfazer as três condições $y^T A \geq 0$, $-y^T A \geq 0$ e $y \geq 0$, o que é o mesmo que satisfazer $y^T A = 0$ e $y \geq 0$. □

Observação 3.10. Se $\ker A^T = \{0\}$ então $Ax \leq b$ tem solução para qualquer b .

Problemas de programação linear e dualidade.

Definição 3.11. Um *problema de programação linear* (ou *problema PL*) consiste em maximizar (ou minimizar) um funcional linear $f(x)$ num dado poliedro P :

$$(P) \quad \max\{f(x) : x \in P\}.$$

Designamos o valor $v(P) = \max\{f(x) : x \in P\}$ por *valor optimal* e a um ponto x^* tal que $f(x^*) = v(P)$ chamamos de *solução optimal*.

Como vimos um poliedro é definido por um sistema linear $Ax \leq b$, e é sabido da álgebra linear que um funcional linear em \mathbb{R}^n , $f : \mathbb{R}^n \rightarrow \mathbb{R}$, é representado por um vector $c \in \mathbb{R}^n$, $f(x) = c^T x$. Assim podemos apresentar um problema de programação linear do seguinte modo:

$$(P) \quad \max\{c^T x : Ax \leq b\} \text{ com } c \in \mathbb{R}^n, A \in \mathbb{R}^{m,n} \text{ e } b \in \mathbb{R}^m$$

designando este por *problema primal* e chamamos *problema dual* ao problema PL:

$$(D) \quad \min\{y^T b : y^T A = c^T, y \geq 0\}$$

Estes dois problemas estão relacionados pelo simples facto de os seus valores optimais (quando existem) serem iguais. É isso que diz o seguinte resultado conhecido por **teorema da dualidade**.

Teorema 3.12. *Seja $A \in \mathbb{R}^{m,n}$, $b \in \mathbb{R}^m$ e $c \in \mathbb{R}^n$. Assumindo que cada um dos dois sistema ($Ax \leq b$) e ($y^T A = c^T, y \geq 0$) é consistente, temos que*

$$\max\{c^T x : Ax \leq b\} = \min\{y^T b : y^T A = c^T, y \geq 0\}.$$

Dem. Seja $x \in \mathbb{R}^n$ tal que $Ax \leq b$ e $y \in \mathbb{R}^m$ tal que $y^T A = c^T, y \geq 0$. Então $c^T x = y^T Ax \leq y^T b$. Logo se os sistemas $Ax \leq b$ e $y^T A = c^T, y \geq 0$ são consistentes temos

$$\max\{c^T x : Ax \leq b\} \leq \min\{y^T b : y^T A = c^T, y \geq 0\}.$$

Esta desigualdade é conhecida por *dualidade fraca*.

Vejam agora a desigualdade contrária. Seja $v = \max\{c^T x : Ax \leq b\}$ então o sistema $Ax \leq b, c^T x \geq l$ não é consistente para qualquer $l > v$. Colocando este último sistema na forma

$$\begin{bmatrix} A \\ -c^T \end{bmatrix} \leq \begin{bmatrix} b \\ -l \end{bmatrix}$$

podemos usar o corolário 3.9 para concluir que existem $y \in \mathbb{R}^m$ e $\lambda \in \mathbb{R}$ com $y \geq 0$ e $\lambda \geq 0$ tais que $y^T A = \lambda c^T$ e $y^T b < \lambda l$. Não podemos ter $\lambda = 0$ pois nesse caso teríamos $y^T A = 0$ e $y^T b < 0$, mas $y^T A = 0, y \geq 0$ e $b \geq Ax$ implicam $y^T b \geq y^T Ax = 0$ o que contradiz $y^T b < 0$. Logo $\lambda > 0$ e tomando $\tilde{y} := \lambda^{-1} y$ temos $\tilde{y} \geq 0, \tilde{y}^T A = c^T$ e $\tilde{y}^T b < l$ pelo que $\min\{y^T b : y^T A = c^T, y \geq 0\} < l$ para qualquer $l > v$. Logo

$$\min\{y^T b : y^T A = c^T, y \geq 0\} \leq \max\{c^T x : Ax \leq b\}.$$

□

A segunda desigualdade pode ser vista com alguma intuição geométrica de seguinte modo. Seja x_0 uma solução optimal do sistema $Ax \leq b$ e seja $I := \{1 \leq i \leq m : a_i^T x_0 = b_i\}$ onde a_i^T são as linhas da matriz A . Intuitivamente vê-se que se x_0 é solução optimal então $c \in \text{cone}(\{a_i : i \in I\})$. Portanto

$$c = \sum_{i \in I} \lambda_i a_i \text{ com } \lambda_i \geq 0.$$

Tomando $y \in \mathbb{R}^m$ tal que $y_i = \lambda_i$ se $i \in I$ e $y_i = 0$ se $i \notin I$ temos que $y \geq 0, y^T A = c^T$ e

$$y^T b = \sum_{i \in I} y_i b_i = \sum_{i \in I} y_i a_i^T x_0 = y^T A x_0 = c^T x_0.$$

Logo

$$\min\{y^T b : y^T A = c^T, y \geq 0\} \leq c^T x_0 = \max\{c^T x : Ax \leq b\}.$$

Note-se que isto não é uma demonstração pois ficou por demonstrar que se x_0 é solução optimal então $c \in \text{cone}(\{a_i : i \in I\})$.

As condições do teorema anterior podem ser enfraquecidas ligeiramente de modo a obter um resultado mais geral.

Teorema 3.13. *Seja $A \in \mathbb{R}^{m,n}$, $b \in \mathbb{R}^m$ e $c \in \mathbb{R}^n$. Consideremos os problemas PL duais (P) $\max\{c^T x : Ax \leq b\}$ e (D) $\min\{y^T b : y^T A = c^T, y \geq 0\}$. Se um dos dois problemas (P) e (D) é viável (i.e. o sistema linear associado é consistente), então*

$$\sup\{c^T x : Ax \leq b\} = \inf\{y^T b : y^T A = c^T, y \geq 0\}.$$

Demonstração. Exercício. □

Pontos interiores e equações implícitas.

Definição 3.14. Uma desigualdade $a_i^T x \leq b_i$ dentro de um sistema linear $Ax \leq b$ diz-se *igualdade implícita* se qualquer solução x de $Ax \leq b$ satisfaz $a_i^T x = b_i$.

Exemplo 3.15. O ponto $(x, y) = (0, 0)$ é a única solução do sistema

$$\begin{cases} x + y \leq 0 \\ x \geq 0 \\ y \geq 0 \end{cases}$$

portanto cada uma das desigualdades do sistema é uma igualdade implícita.

Definição 3.16. Dado um sistema linear $Ax \leq b$ designamos por $A^-x \leq b^-$ o subsistema formado pelas igualdades implícitas e por $A^+x \leq b^+$ o subsistema formado pelas restantes desigualdades.

Observação 3.17. Os sistemas $Ax \leq b$ e $A^-x = b^-$, $A^+x \leq b^+$ são equivalentes.

Definição 3.18. Um dado ponto x_0 diz-se *ponto interno* de um poliedro $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ se x_0 satisfaz $A^-x_0 = b^-$ e $A^+x_0 < b^+$ (o que é equivalente dizer que $x_0 \in P$ e $A^+x_0 < b^+$).

Observação 3.19. Note-se que se o sistema $A^+x < b^+$ for vazio (i.e. $A^- = A$) qualquer ponto de P é um ponto interno.

Lema 3.20. *Se um poliedro $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ é não-vazio então tem pelo menos um ponto interno.*

Demonstração. Seja I^+ o conjunto dos índices das desigualdades do subsistema $A^+x \leq b^+$, ou seja

$$I^+ := \{i : \exists x \in P a_i^T x < b_i\}.$$

Logo, pela definição de I^+ , para qualquer índice $i \in I^+$ existe pelo menos um ponto $x^i \in P$ tal que $a_i^T x^i < b_i$. Assim sendo, podemos construir o ponto

$$x_0 = \frac{1}{|I^+|} \sum_{i \in I^+} x^i$$

que é fácil ver que se trata de um ponto interno de P . □

Proposição 3.21. *Seja $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ um poliedro não-vazio, então*

$$\text{aff}(P) = \{x \in \mathbb{R}^n : A^-x = b^-\} = \{x \in \mathbb{R}^n : A^-x \leq b^-\}$$

e

$$\dim(P) = n - \text{car}(A^-).$$

Dem. Seja $A_1 = \{x \in \mathbb{R}^n : A^-x = b^-\}$ e $A_2 = \{x \in \mathbb{R}^n : A^-x \leq b^-\}$. Como $Ax \leq b \Rightarrow A^-x = b^- \Rightarrow A^-x \leq b^-$ temos $P \subseteq A_1 \subseteq A_2$. Ora sendo A_1 um conjunto afim que contém P temos que $\text{aff}(P) \subseteq A_1$.

Vejam agora que $A_2 \subseteq \text{aff}(P)$. Pelo lema anterior, existe um ponto interno x_0 de P , tal ponto satisfaz, por definição, $A^+x_0 < b^+$ e $A^-x_0 = b^-$. Seja $x \in A_2$, se $x = x_0$ então $x \in P \subseteq \text{aff}(P)$. Se $x \neq x_0$ então existe $\lambda > 0$ suficientemente pequeno de modo a que o ponto $z := \lambda x + (1 - \lambda)x_0$ esteja contido em P (a condição $A^-z \leq b^-$ é satisfeita pois ambos x e x_0 estão em A_2 e a condição $A^+z \leq b^+$ é satisfeita para $\lambda > 0$ suficientemente pequeno). Deste modo temos que $x = \lambda^{-1}z + (1 - \lambda^{-1})x_0$ é uma combinação afim de dois pontos de P , logo $x \in \text{aff}(P)$.

$$\dim(P) = \dim A_1 = \ker(A^-) = n - \text{car}(A^-). \quad \square$$

Decomposições de poliedros.

Definição 3.22. O *espaço linearizado* de um poliedro $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ é o conjunto $\text{lin. space}(P) = \{x \in \mathbb{R}^n : Ax = 0\}$.

Observação 3.23. $y \in \text{lin. space}(P)$ se e só se $\{x + \lambda y : \lambda \in \mathbb{R}\} \subseteq P$ para todo o $x \in P$.

Observação 3.24. $x + \text{lin. space}(P) \subseteq P$ para todo o $x \in P$.

Definição 3.25. Um poliedro P diz-se *pontuado* se $\text{lin. space}(P) = \{0\}$.

Observação 3.26. Um poliedro $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ é *pontuado* se e só se $\text{car}(A) = n$.

Observação 3.27. Se um poliedro P é limitado então P é *pontuado*. No entanto um poliedro *pontuado* não é necessariamente limitado.

Para poliedros *pontuados* podemos definir o conceito de vértice do seguinte modo.

Definição 3.28. Um ponto x_0 de um poliedro $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ *pontuado* diz-se um *vértice* se for a única solução de n equações linearmente independentes do sistema $Ax \leq b$.

Exemplo 3.29. Os vértices (no sentido da definição anterior) do poliedro $P = \{(x, y) \in \mathbb{R}^2 : x \leq 2, y \leq 2, x + y \geq 1\}$ correspondem aos vértices (no conceito usual da geometria elementar) do triângulo P .

Definição 3.30. Um ponto x_0 de um conjunto convexo C diz-se *ponto extremo* de C se não existirem pontos distintos x_1 e x_2 de C tais que $x_0 = \frac{x_1 + x_2}{2}$.

Note-se que isto é equivalente a dizer que x_0 não pertence ao interior de qualquer segmento de recta contido em C .

Este último conceito é, como mostra o seguinte resultado, uma generalização do conceito de vértice para conjuntos convexos não necessariamente poliedros.

Proposição 3.31. Um ponto x_0 de um poliedro $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ *pontuado* é um vértice se e só se é um ponto extremo de P .

Dem. Seja x_0 um vértice de P , então existe uma submatriz $A' \in \mathbb{R}^{n,n}$ de $A \in \mathbb{R}^{m,n}$ com linhas linearmente independentes tal que x_0 é a única solução do sistema $A'x = b'$ onde $b' \in \mathbb{R}^n$ é o correspondente subvector de $b \in \mathbb{R}^m$. Sejam x_1 e x_2 dois quaisquer pontos de P que satisfaçam $x_0 = \frac{x_1 + x_2}{2}$. Como $x_1, x_2 \in P$ temos

$$\begin{cases} A'x_1 \leq b' \\ A'x_2 \leq b' \end{cases}$$

e por outro lado

$$\begin{cases} A'x_1 = A'(2x - x_2) = 2b' - A'x_2 \geq 2b' - b' = b' \\ A'x_2 = A'(2x - x_1) = 2b' - A'x_1 \geq 2b' - b' = b' \end{cases}$$

logo $A'x_1 = A'x_2 = b'$ o que implica $x_1 = x_2$.

Vejamus que se x_0 não for vértice então também não será ponto extremo. Se x_0 não é vértice então o maior subsistema $A'x \leq b'$ de $Ax \leq b$ tal que $A'x_0 = b'$ é tal que a característica de A' não é total ($\text{car}(A') < n$). Logo existe $z \in \mathbb{R}^n$ não-nulo tal que $A'z = 0$. Como $a_i^T x_0 < b_i$ para qualquer desigualdade que não esteja no subsistema $A'x \leq b'$, temos que, para um $\varepsilon > 0$ suficientemente pequeno, $x_0 + \varepsilon z$ e $x_0 - \varepsilon z$ pertencem a P , o que implica que x_0 não é um ponto extremo. \square

Definição 3.32. Uma semi-recta (ou raio) $R(x; r) = \{x + \lambda r : \lambda \geq 0\} = x + \text{cone}(\{r\})$ (com $x, r \in \mathbb{R}^n$ e $r \neq 0$) contida num conjunto convexo C diz-se *semi-recta extrema* de C se x for um ponto extremo¹ de C e não existirem semi-rectas $R(x; r_1)$ e $R(x; r_2)$ não-paralelas (i.e. $\text{cone}(r_1) \neq \text{cone}(r_2)$), contidas em C com $r = \frac{r_1 + r_2}{2}$.

Observação 3.33. Se $x \in P = \{x \in \mathbb{R}^n : Ax \leq b\}$ então $R(x; r) = \{x + \lambda r : \lambda \geq 0\} = x + \text{cone}(\{r\})$ é uma semi-recta contida no poliedro P se e só se $Ar \leq 0$.

¹Esta condição não é requerida no texto de Geir Dahl.

Proposição 3.34. *Seja x um ponto extremo de um poliedro $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ e $r \in \mathbb{R}^n \setminus \{0\}$. Então $R(x; r) = \{x + \lambda r : \lambda \geq 0\} \subseteq P$ é uma semi-recta extrema se e só se existe um subsistema $A'x = 0$ do sistema $Ax = 0$ consistido em $n - 1$ equações linearmente independentes tal que $A'r = 0$.*

Dem. Exercício. □

Corolário 3.35. *Qualquer poliedro pontuado possui um número finito de vértices e semi-rectas extremas.*

Dem. Exercício. □

De facto temos que, se P for um poliedro pontuado definido pelo sistema $Ax \leq b$ onde $A \in \mathbb{R}^{m,n}$, o número de vértices de P não poderá exceder $\binom{m}{n}$ e o número de semi-rectas extremas de P não poderá exceder $\binom{m}{n-1}$ vezes o número de vértices de P .

Lema 3.36. *Qualquer poliedro pontuado não-vazio possui pelo menos um vértice.*

Dem. Exercício. □

Proposição 3.37. *Seja $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ um poliedro não-vazio e pontuado, e consideremos o problema PL*

$$\max\{c^T x : x \in P\}$$

(i) *Se $v(P) = \max\{c^T x : x \in P\}$ é finito, então existe um vértice x_0 que é solução optimal ($c^T x_0 = v(P)$);*

(ii) *Se $v(P) = +\infty$, então existe uma semi-recta extrema $R(x, r) = x + \text{cone}(\{r\})$ de P com $c^T r > 0$ (logo $c^T(x + \lambda r) \rightarrow +\infty$ quando $\lambda \rightarrow +\infty$).*

Dem. (i) Consideremos o conjunto

$$F = \{x \in P : c^T x = v(P)\} = \left\{ x \in \mathbb{R}^n : \begin{bmatrix} A \\ c^T \\ -c^T \end{bmatrix} x \leq \begin{bmatrix} b \\ v(P) \\ -v(P) \end{bmatrix} \right\}$$

É claro que F é um poliedro pontuado (pois P também o é) e não-vazio (se P é fechado e $v(P) < \infty$ então existe $x_0 \in P$ tal que $c^T x_0 = v(P)$). Então, pelo lema 3.36, F tem vértices.

Seja x_0 um vértice de F , vejamos que também é um vértice de P . Pela proposição 3.31 tal é equivalente a mostrar que x_0 é ponto extremo de P . Sejam $x_1, x_2 \in P$ tais que $x_0 = \frac{x_1 + x_2}{2}$ temos então que

$$v(P) = c^T x_0 = \frac{c^T x_1 + c^T x_2}{2} \leq \frac{v(P) + v(P)}{2} = v(P)$$

logo $x_1, x_2 \in F$ e sendo x_0 um ponto extremo de F (pois é um vértice de F) temos que $x_1 = x_2 = x_0$. Ou seja, x_0 é um ponto extremo (logo vértice) de P .

(ii) Se $v(P) = +\infty$ então, pelo teorema 3.13, o poliedro dual $D = \{y \in \mathbb{R}^m : y^T A = c^T, y \geq 0\}$ é vazio. Assim, pelo lema de Farkas (lema 3.8), existe $x \in \mathbb{R}^n$ tal que $Ax \geq 0$ e $c^T x < 0$. Ora tomando $r_0 = -x$ temos que, para qualquer $p \in P$ a semi-recta $R(p, r_0)$ está contida em P e $c^T r_0 > 0$. Falta agora encontrar uma semi-recta como esta mas que seja extrema.

Consideremos o problema PL

$$(Q) \quad \max\{c^T z : Az \leq 0, c^T z \leq 1\}.$$

Este problema é viável (λr_0 é solução para $\lambda > 0$ suficientemente pequeno) e é limitado ($\max\{c^T z : Az \leq 0, c^T z \leq 1\} \leq 1 < \infty$). Deste modo, pela alínea (i) da proposição, o problema possui uma solução optimal r que é um vértice de Q . É fácil ver que tomando um vértice x_0 de P a semi-recta $R(x_0, r)$ é extrema e $c^T r > 0$ (de facto $c^T r = 1$). □

Este resultado mostra que a resolução de um problema PL reduz-se a problema finita, basta verificar se existe uma semi-recta extrema $R(x, r)$ com $c^T r > 0$ (caso em que o problema não tem solução) e, caso não exista, determinar entre os vértices qual o que tem maior valor $c^T x$.

No entanto este procedimento torna-se em geral pouco eficiente há medida que a dimensão do problema cresce.

Outra consequencia deste resultado é que qualquer poliedro pontuado não-vazio pode ser decomposto do modo que é enunciado a seguir.

Teorema 3.38. *Seja $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ um poliedro pontuado não-vazio com conjunto de vértices V . Seja R um conjunto formado por um vector de direcção para cada semi-recta extrema de P . Então P pode ser decomposto do seguinte modo*

$$P = \text{conv}(V) + \text{cone}(R).$$

Dem. Seja $Q = \text{conv}(V) + \text{cone}(R)$. Se $x \in Q$ então $x = x_1 + x_2$ com $x_1 \in \text{conv}(V)$ e $x_2 \in \text{cone}(R)$. Como $V \subseteq P$ e P é convexo temos $\text{conv}(V) \subseteq P$, logo $Ax_1 \leq b$. Por outro lado, para qualquer r que seja vector de direcção de uma semi-recta extrema de P , temos, pela observação 3.33, $Ar \leq 0$. Logo se x_2 é uma combinação cónica de elementos de R também irá satisfazer a condição $Ax_2 \leq 0$. Deste modo, $Ax = Ax_1 + Ax_2 \leq b$ (i.e. $x \in P$), portanto $Q \subseteq P$.

Vejam agora que se $x \notin Q$ então $x \notin P$. Seja $V = \{v_1, v_2, \dots, v_k\}$ e $R = \{r_1, r_2, \dots, r_l\}$. Se $x_0 \notin Q$, então o sistema

$$\begin{bmatrix} v_1 & v_2 & \dots & v_k & r_1 & r_2 & \dots & r_l \\ 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_k \\ \mu_1 \\ \mu_2 \\ \vdots \\ \mu_l \end{bmatrix} = \begin{bmatrix} x_0 \\ 1 \end{bmatrix}, \lambda_1 \geq 0, \lambda_2 \geq 0, \dots, \lambda_k \geq 0, \mu_1 \geq 0, \mu_2 \geq 0, \dots, \mu_l \geq 0$$

não tem solução. Então, pelo lema de Farkas 3.8, existem $y_0 \in \mathbb{R}^n$ e $d \in \mathbb{R}$ tais que

$$[y_0^T d] \begin{bmatrix} v_1 & v_2 & \dots & v_k & r_1 & r_2 & \dots & r_l \\ 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 \end{bmatrix} \geq 0 \text{ e } [y_0^T d] \begin{bmatrix} x_0 \\ 1 \end{bmatrix} < 0.$$

Ou seja, $y_0^T v_i \geq -d$ para todo o $1 \leq i \leq k$, $y_0^T r_j \geq 0$ para todo o $1 \leq j \leq l$ e $y_0^T x_0 < -d$. Tomando $c = -y_0$ temos $c^T v_i \leq d$ para todo o $1 \leq i \leq k$, $c^T r_j \leq 0$ para todo o $1 \leq j \leq l$ e $c^T x_0 > d$. Ora $c^T r_j \leq 0$ para todo o $1 \leq j \leq l$ implica, pela proposição 3.37 (ii), que o problema $\max\{c^T x : x \in P\}$ é limitado ($\max\{c^T x : x \in P\} < \infty$). E pela alínea (i) da mesma proposição temos $\max\{c^T x : x \in P\} = \max\{c^T v_1, \dots, c^T v_k\} \leq d$. Como $c^T x_0 > d$ temos que $x_0 \notin P$. \square

Lema 3.39. *Seja $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ um poliedro não-vazio. Então P pode ser decomposto como $P = Q + L$ onde $L = \text{lin. space}(P)$ e $Q = P \cap L^\perp$ é um poliedro pontuado.*

Dem. Seja $x \in P$ e $x = x_1 + x_2$ a sua decomposição ortogonal em L (i.e. $x_1 \in L$ e $x_2 \in L^\perp$), então $Ax_2 = Ax \leq b$ logo $x_2 \in P \cap L^\perp$. Logo $x \in L + Q$.

Por outro lado, se $x \in L + Q$ ($x = x_1 + x_2$ com $x_1 \in L$ e $x_2 \in Q$) então $Ax = Ax_1 + Ax_2 = Ax_2 \leq b$, logo $x \in P$.

Falta ver que Q é um poliedro pontuado, ou seja que $\text{lin. space}(Q) = \{0\}$. Seja $y \in \text{lin. space}(Q)$, sabemos pela observação 3.23 que $\{x + \lambda y : \lambda \in \mathbb{R}\} \subseteq Q$ para qualquer $x \in Q$. Como $Q \subseteq P$, temos que $A(x + \lambda y) \leq b$ para qualquer $\lambda \in \mathbb{R}$, logo $Ay = 0$. Portanto $y \in L$, e como $Q \subseteq L^\perp$, $y^T(x + \lambda y) = y^T x + \lambda \|y\|^2 = 0$ para todo o $\lambda \in \mathbb{R}$. Donde se tira que $y = 0$. Portanto $\text{lin. space}(Q) = \{0\}$ (Q é pontuado). \square

Resulta imediatamente dos dois últimos resultados a seguinte proposição.

Proposição 3.40. *Qualquer poliedro P não-vazio pode ser decomposto da forma*

$$P = \text{conv}(V) + \text{cone}(R) + L$$

onde $L = \text{lin. space}(P)$, V é o conjunto dos vértices de $P \cap L^\perp$ e R é um conjunto que consiste em um vector direcção para cada semi-recta extrema de $P \cap L^\perp$.

Observação 3.41. Esta decomposição é mínima no sentido que se $P = \text{conv}(V') + \text{cone}(R') + L'$ de modo que $\text{cone}(R')$ não contenha nenhum subespaço próprio, então $L' = L$, $V \subseteq V'$ e $R \subseteq R'$ (sendo que nesta última inclusão devemos assumir que os vectores direcção são unitários).

Outra observação que podemos fazer é a seguinte.

Observação 3.42. Qualquer subespaço linear é um cone finitamente gerado.

De facto, se b_1, b_2, \dots, b_s formam uma base do subespaço L então $L = \text{cone}(b_1, \dots, b_s, u)$ onde $u = -b_1 - \dots - b_s$.

Uma vez que a soma de dois cones é um cone ($\text{cone}(R_1) + \text{cone}(R_2) = \text{cone}(R_1 \cup R_2)$), temos que

Proposição 3.43. *Qualquer poliedro não-vazio P é a soma de um polítopo $\text{conv}(V)$ com um cone finitamente gerado $\text{cone}(R)$. Sendo que podemos tomar por V o conjunto dos vértices de $P \cap L^\perp$ onde $L = \text{lin. space}(P)$.*

Iremos mais à frente demonstrar que o recíproco também é válido. Para tal vamos recorrer ao seguinte resultado, conhecido por **teorema de Farkas-Minkowski-Weyl**, não sem antes introduzir a seguinte definição.

Definição 3.44. Um cone poliedral é um cone da forma $\{x \in \mathbb{R}^n : Ax \leq 0\}$.

Teorema 3.45. *Um cone convexo K é poliedral se e só se é finitamente gerado.*

Dem. Primeiro vamos mostrar que um cone poliedral é um cone finitamente gerado.

Seja $K = \{x \in \mathbb{R}^n : Ax \leq 0\}$ um cone poliedral não-vazio (se for vazio é claramente finitamente gerado). Então, pela proposição 3.43 (pois K é um poliedro), $K = \text{conv}(V) + \text{cone}(R)$ onde V é o conjunto dos vértices de $K \cap L^\perp$ (onde $L = \text{lin. space}(K)$) e R é um conjunto vazio. Ora $V = \{0\}$ logo $K = \text{cone}(R)$ é um cone finitamente gerado.

Por outro lado, para qualquer cone finitamente gerado $K = \text{cone}(\{a_1, \dots, a_k\})$, o seu cone polar é poliedral $K^\circ = \{x \in \mathbb{R}^n : a_i^T x \leq 0, 1 \leq i \leq k\}$. Logo pelo que acabamos de demonstrar, K° é finitamente gerado $K^\circ = \text{cone}(\{b_1, \dots, b_s\})$. Portanto, sendo K pelo corolário 2.26 fechado, temos que $K = K^{\circ\circ} = \{x \in \mathbb{R}^n : b_i^T x \leq 0, 1 \leq i \leq s\}$, ou K é poliedral. \square

Estamos agora em condições de enunciar e demonstrar o próximo resultado conhecido como **teorema da representação de Motzkin** ou **teorema da decomposição para poliedros**.

Teorema 3.46. *Um conjunto convexo $P \subseteq \mathbb{R}^n$ é um poliedro se e só se existem conjuntos finitos V e R tais que*

$$P = \text{conv}(V) + \text{cone}(R).$$

Dem. Vimos pela proposição 3.43 que se P é um poliedro então $P = \text{conv}(V) + \text{cone}(R)$ com V e R conjuntos finitos.

Consideremos então um conjunto da forma $Q = \text{conv}(V) + \text{cone}(R)$ com V e R finitos, e vejamos que Q é um poliedro. $x \in Q$ se e só se

$$x = \sum_i \lambda_i v_i + \sum_j \mu_j r_j$$

com $\sum_i \lambda_i = 1$, λ_i 's e μ_j 's não-negativos, $v_i \in V$ e $r_j \in R$. O que é equivalente a dizer que

$$(x, 1) = \sum_i \lambda_i (v_i, 1) + \sum_j \mu_j (r_j, 0)$$

com λ_i 's e μ_j 's não-negativos, $v_i \in V$ e $r_j \in R$ (ou seja $(x, 1)$ é uma combinação cónica dos vectores $(v_i, 1)$ e $(r_j, 0)$ em \mathbb{R}^{n+1}). Portanto $x \in Q$ se e só se $(x, 1) \in Q'$ onde $Q' = \text{cone}(\{(v, 1) \in \mathbb{R}^{n+1} : v \in V\} \cup \{(r, 0) \in \mathbb{R}^{n+1} : r \in R\})$. Como Q' é um cone finitamente gerado existe, pelo teorema de Farkas-Minkowski-Weyl, uma matriz $A \in \mathbb{R}^{m, n+1}$ tal que $Q' = \{x \in \mathbb{R}^{n+1} : Ax \leq 0\}$. Temos então que $Q = \{x \in \mathbb{R}^n : (x, 1) \in Q'\} = \{x \in \mathbb{R}^n : A'x \leq b\}$ onde $A' \in \mathbb{R}^{m, n}$ é a submatriz de A formada pelas primeiras n colunas de A e $b \in \mathbb{R}^m$ é o vector simétrico da última coluna de A ($b_i = -a_{i, n+1}$). Portanto Q é um poliedro. \square

Deste teorema conclui-se facilmente o seguinte corolário.

Corolário 3.47. *Um conjunto é um polítopo se e só se é um poliedro limitado.*

Dem. Exercício. □

Faces de um poliedro.

Definição 3.48. Seja $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ um poliedro. Uma *desigualdade válida* para P é uma desigualdade linear $c^T x \leq \alpha$ que é satisfeita por todo o $x \in P$. Ou seja $P \subseteq H_{\leq}(c, \alpha)$.

Definição 3.49. Uma desigualdade válida $c^T x \leq \alpha$ diz-se *suportiva* se existe um ponto $x_0 \in P$ tal que $c^T x_0 = \alpha$. Ou seja $P \cap H_{=}(c, \alpha) \neq \emptyset$.

Definição 3.50. Uma *face* de P é um conjunto F da forma

$$F = P \cap H_{=}(c, \alpha)$$

para uma desigualdade válida $c^T x \leq \alpha$, dizemos então que F é a *face induzida* por $c^T x \leq \alpha$.

Um caso particular de uma face de P é o próprio P como se vê na seguinte observação.

Observação 3.51. P é face de P induzida pela desigualdade $0^T x \leq 0$.

Definição 3.52. P e \emptyset são chamadas de *faces triviais* de P . As restantes faces são chamadas de *faces próprias*.

Teorema 3.53. *Seja $P = \{x \in \mathbb{R}^n : Ax \leq b\}$. Um conjunto não-vazio F é uma face de P se e só se $F = \{x \in P : A'x = b'\}$ para algum subsistema $A'x \leq b'$ de $Ax \leq b$.*

Dem. Se F é uma face não-vazia então existe um $c \in \mathbb{R}^n \setminus \{0\}$ tal que $F = \{x \in P : c^T x = \alpha\}$ com $\alpha = \max\{c^T x : x \in P\}$. Pelo teorema da dualidade $\alpha = \min\{y^T b : y^T A = c^T, y \geq 0\}$. Seja $y^* \in \mathbb{R}^m$ uma solução optimal de (D) e seja $I' = \{i = 1, \dots, m : y_i^* > 0\}$. Para qualquer $x \in P$ temos que

$$c^T x = (y^{*T} A)x = \sum_{i=1}^m y_i^* (Ax)_i = \sum_{i \in I'} y_i^* (Ax)_i \leq \sum_{i \in I'} y_i^* b_i = \alpha$$

Temos portanto que $c^T x = \alpha$ se e só se $(Ax)_i = b_i$ para qualquer $i \in I'$. Ou seja $F = \{x \in P : a_i^T x = b_i, i \in I'\}$.

Por outro lado, seja F um conjunto da forma $F = \{x \in P : A'x = b'\}$ para um subsistema $A'x \leq b'$ de $Ax \leq b$. Ou seja, $F = \{x \in P : a_i^T x = b_i, i \in I'\}$ com $I' \subseteq \{1, \dots, m\}$. Vejamos que F é a face induzida pela desigualdade $c^T x \leq \alpha$ onde $c = \sum_{i \in I'} a_i$ e $\alpha = \sum_{i \in I'} b_i$. $\alpha = \sum_{i \in I'} b_i$ é uma desigualdade válida pois $x \in P \Rightarrow Ax \leq b \Rightarrow A'x \leq b' \Rightarrow c^T x \leq \alpha$. Se $x \in F$ então $x \in P$ e $A'x = b'$ o que implica que $x \in P$ e $c^T x = \alpha$, logo $F \subseteq P \cap H_{=}(c, \alpha)$. E se $x \in P \cap H_{=}(c, \alpha)$ então $\alpha = c^T x = 1^T A'x \leq 1^T b' = \alpha$ o que implica que $a_i^T x = b_i \forall i \in I'$, logo $P \cap H_{=}(c, \alpha) \subseteq F$. □

Este resultado diz-nos que uma face de P pode ser caracterizada como $F = \{x \in P : A'x = b', A''x \leq b''\}$ onde A' e A'' formam uma decomposição da matriz A , e b' e b'' formam a correspondente decomposição do vector b .

Corolário 3.54. *Seja $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ um poliedro. Então*

- (i) P tem um número finito de faces;
- (ii) Cada face de P é poliedro;
- (iii) Seja F uma face de P (logo F é um poliedro por (ii)). Então um subconjunto G de F é uma face de F se e só se é uma face de P .

Dem. Exercício. □

Definição 3.55. (1) Uma *face minimal* de P é uma face não-vazio de P que não contém estritamente uma outra face não-vazia de P ;

- (2) Uma *face maximal* de P é uma face própria de P que não está estritamente contida numa outra face própria de P .

Teorema 3.56. *Seja F um subconjunto não-vazio de um poliedro $P = \{x \in \mathbb{R}^n : Ax \leq b\}$. Então F é uma face minimal de P se e só se F é um conjunto afim da forma:*

$$F = \{x \in \mathbb{R}^n : A'x = b'\}$$

para algum subsistema $A'x \leq b'$ de $Ax \leq b$.

Dem. Seja $F = \{x \in \mathbb{R}^n : A'x = b', A''x \leq b''\}$ uma face minimal de P com $A'x \leq b'$ e $A''x \leq b''$ subsistemas de $Ax \leq b$, e seja I'' o conjunto dos índices que formam o subsistema $A''x \leq b''$. Podemos assumir que, para cada $i \in I''$, $A'x = b' \wedge a_j^T x \leq b_j \forall j \in I'' \setminus \{i\} \not\Rightarrow a_i^T x \leq b_i$ (i.e. $a_i^T x \leq b_i$ é não-redundante)² pois caso contrário poderíamos retirar a desigualdade $a_i^T x \leq b_i$ ao sistema $A''x \leq b''$. Assim existe, para cada $i \in I''$, pelo menos um $y_i \in \mathbb{R}^n$ tal que $A'y_i = b'$, $a_j^T y_i \leq b_j \forall j \in I'' \setminus \{i\}$ e $a_i^T y_i > b_i$. Como o conjunto F é não-vazio, para cada $i \in I''$ o conjunto $F_i = \{x \in \mathbb{R}^n : A'x = b', A''x \leq b'', a_i^T x = b_i\} \subseteq F$ é não-vazio (se $x \in F$ então existe um $\lambda_i \in [0, 1[$ tal que $\lambda_i y_i + (1 - \lambda_i)x \in F_i$). Como F é uma face minimal temos que $F_i = F$ (ou seja, $a_i^T x \leq b_i$ é uma igualdade implícita) e portanto podemos transferir a igualdade $a_i^T x = b_i$ do sistema $A''x \leq b''$ para o sistema $A'x \leq b'$ de modo que no final deste processo ficamos com $F = \{x \in \mathbb{R}^n : A'x = b'\}$.

Por outro lado se $\emptyset \neq F \subseteq P$ é um conjunto da forma $F = \{x \in \mathbb{R}^n : A'x = b'\}$, então qualquer face F' de F é, pelo teorema 3.53, da forma $F' = \{x \in F : A''x = b''\} = F \cap \{x \in \mathbb{R}^n : A''x = b''\}$ para algum subsistema $A''x \leq b''$ do sistema $(A'x \leq b' \wedge -A'x \leq -b')$. Mas nesse caso temos $F \subseteq \{x \in F : A''x = b''\}$ logo F é uma face minimal. \square

Corolário 3.57. *Cada face minimal F de $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ é um transladado de $\text{lin. space}(P)$, e portanto $\dim(F) = n - \text{car}(A)$. Se P é pontuado então as faces minimais de P são os conjuntos singulares formados pelos vértices de P .*

Dem. Exercício. \square

Definição 3.58. Dado um sistema do tipo

$$a_i^T x = b_i \text{ com } i \in I_1$$

$$a_i^T x \leq b_i \text{ com } i \in I_2$$

Dizemos que o sistema é *minimal* se:

- (1) nenhuma desigualdade pode ser transformada em igualdade sem alterar o conjunto de soluções;
- (2) nenhum constrangimento (desigualdade ou igualdade) é *redundante* (ou seja, pode ser removido sem alterar o conjunto de soluções).

Neste caso dizemos que o sistema é uma *representação minimal* do poliedro que é conjunto solução do sistema.

Teorema 3.59. *O sistema*

$$a_i^T x = b_i \text{ com } i \in I_1$$

$$a_i^T x \leq b_i \text{ com } i \in I_2$$

é uma representação minimal de um poliedro P não-vazio se e só se:

- (1) $\{a_i\}_{i \in I_1}$ é um conjunto linearmente independente;
- (2) A aplicação
$$I_2 \longrightarrow \mathcal{F}$$

 $i \longmapsto \{x \in P : a_i^T x = b_i\}$, onde \mathcal{F} é o conjunto das facetos (faces maximais) de P , é uma bijecção.

Dem. Ver teorema 2.25 do texto de Geir Dahl que está na bibliografia da página web da cadeira. \square

²Neste caso dir-se-ia que a desigualdade $a_i^T x \leq b_i$ seria *redundante*.

4. ALGORITMO SIMPLEX

Consideremos o problema primal de programação linear:

$$(P) \quad \max\{c^T x : Ax = b, x \geq 0\}$$

Podemos assumir que $A \in \mathbb{R}^{m,n}$ com $\text{car}(A) = n$ (logo $m \leq n$) pois se as linhas de A não forem linearmente independentes temos que ou o sistema $Ax = b$ é inconsistente ou podemos retirar-lhe equações, sem alterar o conjunto de soluções, de modo a que todas as equações sejam linearmente independentes.

Observação 4.1. O poliedro $P = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$ é pontuado.

Observemos também que qualquer problema PL pode ser reduzido a problema de tipo (P). Por exemplo as soluções ótimas do problema $\min\{c^T x : x \in P\}$ coincidem com as soluções ótimas do problema $\max\{(-c)^T x : x \in P\}$ sendo que $\min\{c^T x : x \in P\} = -\max\{(-c)^T x : x \in P\}$. Por outro lado se tivermos um problema do tipo

$$(A) \quad \max\{c^T x : Ax \leq b\}$$

este pode ser transformado no problema

$$(B) \quad \max\{\tilde{c}^T \tilde{x} : \tilde{A}\tilde{x} = b, \tilde{x} \geq 0\}$$

onde $\tilde{c}^T = [c^T \mid -c^T \mid 0] \in \mathbb{R}^{2n+m}$, $\tilde{A} = [A \mid -A \mid I] \in \mathbb{R}^{m,2n+m}$ e $\tilde{x}^T \in \mathbb{R}^{2n+m}$. Neste caso $x \in \mathbb{R}^n$ é solução (optimal) de (A) se e só se $x = x_1 - x_2$ para algum $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ que seja solução (optimal) de (B).

Notações: Dado um vector $z \in \mathbb{R}^J$ (respectivamente uma matriz $M \in \mathbb{R}^{I \times J}$), com $I = \{1, \dots, m\}$ e $J = \{1, \dots, n\}$, e um subconjunto de índices $J' = \{j_1, \dots, j_k\} \subseteq J$ denotamos por $z_{J'}$ (resp. $M_{J'}$) o subvector de z (resp. a submatriz de M) definido por

$$z_{J'} = [z_j]_{j \in J'} \quad \left(\text{resp. } M_{J'} = [m_{i,j}] \begin{matrix} i \in I \\ j \in J' \end{matrix} \right)$$

Exemplo 4.2. Considerando $z = \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{bmatrix}$, $M = \begin{bmatrix} a & b & c & d \\ e & f & g & h \end{bmatrix}$ e $J = \{2, 4\}$ temos que $z_J =$

$$\begin{bmatrix} \beta \\ \delta \end{bmatrix} \text{ e } M_J = \begin{bmatrix} b & d \\ f & h \end{bmatrix}$$

Definição 4.3. Dada uma matriz $A \in \mathbb{R}^{m,n}$ com característica m e um subconjunto de índices $B = (b_1, \dots, b_m) \subseteq \{1, \dots, n\}$, dizemos que A_B é uma *base* e que B é um *conjunto de índices de base* se A_B for não-singular.

Observação 4.4. Como $\text{car}(A) = m$, A tem pelo menos um conjunto de índices de base.

Definição 4.5. A um conjunto de índices de base $B \subseteq J$ para $A \in \mathbb{R}^{I \times J}$ associamos a *solução básica primal* $x^B = (x_B, x_N)$ onde $x_N = 0$ ($N = J \setminus B$) e $x_B = A_B^{-1}b$.

As variáveis de x_B (i.e. x_j com $j \in B$) são chamadas de *variáveis básicas* e as variáveis de x_N são chamadas de *variáveis não-básicas*.

Observação 4.6. Qualquer solução básica primal $x^B = (x_B, x_N)$ é solução do sistema $Ax = b$.

Definição 4.7. Se uma solução básica primal x^B pertence a P (ou seja $x_B \geq 0$) então x^B é chamada de *solução básica primal viável*.

Lema 4.8. O conjunto dos vértices de $P = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$ coincide com o conjunto das soluções básicas viáveis de $Ax = b, x \geq 0$.

Dem. Seja v um vértice de P , v terá de ser solução de n equações linearmente independentes retiradas do sistema $Ax = b, x \geq 0$. Como $\text{car}(A) = m$, pelo menos $n - m$ dessas equações terão de vir do subsistema $x \geq 0$. Por outro lado, como $v \in P$, v obedece a todas as m equações do subsistema $Ax = b$. Assim o subsistema que define v é formado (ou pode ser formado) pelas m equações do sistema $Ax = b$ mais $n - m$ equações do sistema $x \geq 0$. Seja $N \subseteq \{1, \dots, n\}$ o conjunto dos $n - m$ índices j tais que $v_j = 0$. A matriz associada ao subsistema que define v é da forma

$$\begin{bmatrix} A_B & A_N \\ 0 & I \end{bmatrix}$$

onde $B = \{1, \dots, n\} \setminus N$. Como, por definição de vértice, esta matriz tem de ser não-singular e por conseguinte A_B é também não-singular. Logo $v = x^B$ é uma solução básica (viável pois $v \in P$).

Por outro lado se x^B é uma solução básica viável então x^B pertence a P e é solução do subsistema

$$\begin{bmatrix} A_B & A_N \\ 0 & I \end{bmatrix} \begin{bmatrix} x_B \\ x_N \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix}$$

Ora, sendo A_B não-singular, as n equações deste sistema são linearmente independentes logo a sua única solução x^B é um vértice de P . \square

Consideremos o problema dual de (P) :

$$(D) \quad \min\{y^T b : y^T A \geq c^T\}$$

Definição 4.9. Dado um conjunto de índices de base B , seja A_B a base associada a B e x^B a correspondente solução básica primal.

Definimos a *solução básica dual* $y^B \in \mathbb{R}^m$ como sendo $(y^B)^T = c_B^T A_B^{-1}$.

Temos então que $(y^B)^T A = (y^B)^T [A_B A_N] = [c_B^T (y^B)^T A_N]$. Logo uma solução básica dual y^B é *solução básica dual viável* (i.e. $(y^B)^T A \geq c^T$) se $(y^B)^T A_N \geq c_N^T$.

As soluções básicas x^B e y^B são chamadas de *complementares* por satisfazerem a *condição de folga complementar*:

$$((y^B)^T A - c^T) x^B = 0$$

Proposição 4.10. *Assumindo que B é um conjunto de índices de base tal que as soluções básicas x^B (primal) e y^B (dual), respectivamente, são viáveis temos que x^B é optimal em (P) e y^B é optimal em (D) .*

Dem. Como x^B é uma solução básica viável primal, $x^B \in P$ logo $c^T x^B \leq v(P)$. Por outro lado, como y^B é uma solução básica viável dual, $y^B \in D$ logo $(y^B)^T b \geq v(D)$. Pelo teorema da dualidade temos que $v(P) = v(D)$, o que junto com o facto de

$$c^T x^B = c_B^T x_B = c_B^T A_B^{-1} A_B x_B = (y^B)^T A_B x_B = (y^B)^T b$$

nos leva a concluir que $c^T x^B = v(P)$ e $(y^B)^T b = v(D)$. \square

Observação 4.11. Seja $A \in \mathbb{R}^{m,n}$ ($m \leq n$), $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$ e um problema PL (P) $\max\{c^T x : Ax = b, x \geq 0\}$.

Seja $B = \{B_1, \dots, B_m\} \subseteq \{1, \dots, n\}$ com $B_1 < \dots < B_m$ um conjunto de índices de base e A_B a base associada a B e A . Então o problema (P) é equivalente ao problema $P(B)$ relativo à corrente base B :

$$(P(B)) \quad c_B^T \bar{b} + \max\{\bar{c}_N^T x_N : x_B + \bar{A}_N x_N = \bar{b}, x_B \geq 0, x_N \geq 0\}$$

onde $\bar{A}_N = A_B^{-1} A_N$, $\bar{b} = A_B^{-1} b$ e $\bar{c}_N^T = c_N^T - c_B^T A_B^{-1} A_N$.

De facto, temos

$$\begin{aligned} Ax = b, x \geq 0 &\Leftrightarrow A_B x_B + A_N x_N = b, x_B \geq 0, x_N \geq 0 \Leftrightarrow x_B + A_B^{-1} A_N x_N = A_B^{-1} b, x_B \geq 0, x_N \geq 0 \Leftrightarrow \\ &\Leftrightarrow x_B + \bar{A}_N x_N = \bar{b}, x_B \geq 0, x_N \geq 0 \end{aligned}$$

e

$$c^T x = c_B^T x_B + c_N^T x_N = c_B^T (A_B^{-1} b - A_B^{-1} A_N x_N) + c_N^T x_N = c_B^T \bar{b} + (c_N^T - c_B^T A_B^{-1} A_N) x_N = c_B^T \bar{b} + \bar{c}_N^T x_N$$

Definição 4.12. Ao vector $\bar{c}_N^T = c_N^T - c_B^T A_B^{-1} A_N$ chamamos *vector de custo reduzido*.

Observação 4.13. Se $\bar{c}_N^T \leq 0$ então a solução básica primal associada a B , $x^B = (x_B, x_N) = (A_B^{-1}b, 0) = (\bar{b}, 0)$, é solução optimal de $(P(B))$ (logo também o é de (P)). Além disso a solução básica dual, $y^B = c_B^T A_B^{-1}$, é solução viável de (D) ($y^B A = c_B^T A_B^{-1} [A_N | A_N] = [c_B^T | c_B^T A_B^{-1} A_N] \geq [c_B^T | c_N^T] = c^T$).

Observação 4.14. Se $\bar{c}_N^T \not\leq 0$ então existe um índice $r \in N$ tal que $\bar{c}_r > 0$. Neste caso solução básica primal x^B não é solução optimal pois podemos incrementar a variável x_r .

De facto, para $\lambda_r \geq 0$ podemos construir uma solução de $Ax = b$ do seguinte modo:

$$\begin{cases} (x(\lambda_r))_r &= \lambda_r \\ x(\lambda_r)_j &= 0 & \text{para } j \in N \setminus \{r\} \\ x(\lambda_r)_B &= \bar{b} - \bar{a}_r \lambda_r & \text{onde } \bar{a}_r \text{ é a coluna de índice } r \text{ de } \bar{A}_N \end{cases}$$

É fácil ver que $x(\lambda_r)$ satisfaz $x(\lambda_r)_B + \bar{A}_N x(\lambda_r)_N = \bar{b}$ e que $x(\lambda_r)_N \geq 0$ para qualquer $\lambda_r \geq 0$. No entanto, para que $x(\lambda_r)$ seja uma solução viável (em particular $x(\lambda_r)_B \geq 0$), é necessário que λ_r satisfaça a condição $\bar{a}_{i,r} \lambda_r \leq \bar{b}_i$ para todo o $i \in B$. Uma vez que $\bar{b} \geq 0$ (pois partimos de uma solução básica primal viável) temos que se $\bar{a}_{i,r} \leq 0$ então $\bar{a}_{i,r} \lambda_r \leq \bar{b}_i$ para todo o $\lambda_r \geq 0$, e se $\bar{a}_{i,r} > 0$ então $\bar{a}_{i,r} \lambda_r \leq \bar{b}_i$ se e só se $\lambda_r \leq \frac{\bar{b}_i}{\bar{a}_{i,r}}$.

Logo, temos que o valor máximo de λ_r para o qual $x(\lambda_r)$ é solução viável é

$$\lambda_r^* = \begin{cases} \min\{\frac{\bar{b}_i}{\bar{a}_{i,r}} : 1 \leq i \leq m, \bar{a}_{i,r} > 0\} & \text{se } \bar{a}_r \not\leq 0 \\ \infty & \text{se } \bar{a}_r \leq 0 \end{cases}$$

No caso em que λ_r^* é finito (ou seja, quando $\bar{a}_r \not\leq 0$) consideremos o conjunto

$$B^*(r) = \{B_i : 1 \leq i \leq m, \bar{a}_{i,r} > 0, \frac{\bar{b}_i}{\bar{a}_{i,r}} = \lambda_r^*\}$$

temos que $x(\lambda_r^*)_i = 0$ para qualquer $i \in B^*(r)$.

Observação 4.15. $x(\lambda_r) \geq 0$ para qualquer $0 \leq \lambda_r \leq \lambda_r^*$.

Definição 4.16. Dizemos que uma solução básica primal x^B é *degenerada* se $x_{B_i} = 0$ para algum $i \leq m$ (ou dito de outra forma, $x_i = 0$ para algum $i \in B$). Caso contrário (i.e. $x_B > 0$) dizemos que x^B é não-degenerada.

Observação 4.17. $x(\lambda_r) \neq x(0) = x^B$ se e só se $\lambda_r > 0$.

Observação 4.18. Se x^B é não-degenerada então $\lambda_r > 0$.

Proposição 4.19. Considere-se um conjunto de índices de base B com $x_B \geq 0$ e um $r \in J \setminus B$. Defina-se λ_r^* e $x(\lambda_r^*)$ como foi feito atrás. Seja

$$F = \{x(\lambda_r) : 0 \leq \lambda_r \leq \lambda_r^*\}.$$

Então temos sempre que:

- (i) Se $\lambda_r = 0$, o vértice $x_0 = x^B$ é degenerado e para cada $s \in B^*(r)$ o conjunto $(B \setminus \{s\}) \cup \{r\}$ é um conjunto de índices de base com solução básica viável x_0 .
- (ii) Se $0 < \lambda_r^* < \infty$, cada um dos conjuntos $(B \setminus \{s\}) \cup \{r\}$ (com $s \in B^*(r)$) é um conjunto de índices de base com solução básica viável $x(\lambda_r^*)$. Além disso, $F = \{x(\lambda_r) : 0 \leq \lambda_r \leq \lambda_r^*\} = \text{conv}(\{x^B, x(\lambda_r^*)\})$ é aresta (i.e. uma face de dimensão 1) de P que junta os dois vértices x^B e $x(\lambda_r^*)$.
- (iii) Se $\lambda_r^* = \infty$, então o conjunto $F = \{x(\lambda) : \lambda \geq 0\}$ é uma semi-recta extrema de P (i.e. uma face de dimensão 1 ilimitada).

Dem. Vejamos para os casos (i) e (ii) que, se B é um conjunto de índices de base com solução básica x^B viável e $s \in B^*$, $B' = (B \setminus \{s\}) \cup \{r\}$ é um conjunto de índices de base com solução básica viável $x(\lambda_r^*)$ ($= x^B$ se $\lambda_r^* = 0$).

Sendo A_B não-singular, temos que a coordenada s da solução única x do sistema $A_B x = a_r$ (onde a_r é a coluna de índice r de A) é dada, pela regra de Cramer, por

$$x_r = \frac{\det(A_{B'})}{\det(A_B)}$$

Por outro lado pode ser dada por

$$x_r = e_r^T x = e_s^T A_B^{-1} a_r = e_s^T A_B^{-1} A_N e_r = e_s^T \bar{A}_N e_r = \bar{a}_{s,r}$$

Como $s \in B^*$, temos que $\bar{a}_{s,r} > 0$ logo $\det(A_{B'}) \neq 0$. Ou seja, B' é um conjunto de índices de base.

Do modo como $x(\lambda_r^*)$ foi construído temos que $x(\lambda_r^*)_i = 0$ para qualquer $i \in N' = J \setminus B' = (N \setminus \{r\}) \cup \{s\}$ e $Ax(\lambda_r^*) = b$:

$$\begin{aligned} Ax(\lambda_r^*) &= \sum_{i \in J} a_i x(\lambda_r^*)_i \\ &= \sum_{i \in B} a_i (\bar{b} - \bar{a}_r \lambda_r^*)_i + a_r \lambda_r^* \\ &= A_B (\bar{b} - \bar{a}_r \lambda_r^*) + a_r \lambda_r^* \\ &= A_B \bar{b} - A_B \bar{a}_r \lambda_r^* + a_r \lambda_r^* \\ &= A_B \bar{b} && \text{pois } \bar{a}_r = A_B^{-1} a_r \\ &= b && \text{pois } \bar{b} = A_B^{-1} b \end{aligned}$$

logo $x(\lambda_r^*)$ é solução básica de B' (i.e. $x(\lambda_r^*)_{B'} = A_B^{-1} b$ e $x(\lambda_r^*)_{N'} = 0$), e é viável pois foi construído de modo a garantir tal ($x(\lambda_r^*)_{B'} \geq 0$).

Fica como exercício ver que F , para (ii) e para (iii), é uma face de dimensão 1. \square

Algoritmo simplex:

- Passo 0 (inicialização) Seja B um conjunto de índices de base tal que a solução básica primal x^B é viável. Calcula-se $x_B = A_B^{-1} b$ e $(y^B)^T = c_B^T A_B^{-1}$.
- Passo 1 (verificação optimal) Calcula-se o custo reduzido $\bar{c}_N^T = c_N^T - (y^B)^T A_N$.
- Se $\bar{c}_N \leq 0$, então termina-se; x^B é optimal em (P) e y^B é optimal em (D) .
 - Caso contrário, escolhe-se $r \notin B$ com $\bar{c}_r > 0$ e faz-se passo 2.
- Passo 2 (pivoting) Determina-se λ_r^* e $B^*(r)$.
- Se $B^*(r)$ é vazio (caso em que $\bar{a}_r \leq 0$ logo $\lambda_r^* = +\infty$), então (P) é ilimitado; $x(\lambda_r) \in P$ para todo o $\lambda_r \geq 0$ e $c^T x(\lambda_r) \rightarrow +\infty$ quando $\lambda_r \rightarrow +\infty$.
 - Caso contrário, escolhe-se um $s \in B^*(r)$ e redefine-se o conjunto de índices de base como sendo $B := (B \setminus \{s\}) \cup \{r\}$. Determina-se as novas soluções básicas primal e dual x^B e y^B , e regressa-se ao passo 1.

Note-se que para inilizar o algoritmo temos de encontrar uma solução básica viável. Isso pode ser feito do seguinte modo:

Dado um problema de programação linear

$$(P) \quad \max\{c^T x : Ax = b, x \geq 0\}$$

podemos assumir primeiro que $b \geq 0$ (se algum b_i for menor que zero podemos substituir a equação $a_i^T x = b_i$ do sistema pela sua simétrica $-a_i^T x = -b_i$ sem alterar a resolução do problema). Consideremos então o problema de programação linear

$$(PI) \quad \min\{\mathbf{1}^T y : Ax + y = b, x \geq 0, y \geq 0\}$$

onde $\mathbf{1} \in \mathbb{R}^m$ é o vector em que todas as entradas são 1 e $y \in \mathbb{R}^m$ é um vector de variáveis de folga. É fácil ver que se (P) é consistente então (PI) tem valor optimal $v(PI) = 0$ com solução optimal (x_0, y_0) em que $y_0 = 0$ e x_0 é um vértice de P (logo é uma solução básica viável de (P)). Além disso temos que (PI) tem uma solução básica viável imediata que é dada por $(x, y) = (0, b)$.

Exemplo 4.20. Consideremos o problema de programação linear (P) :

$$\begin{aligned}
\text{Maximizar: } & x_1 - x_2 + 2x_3 - x_4 \\
\text{Sujeito a: } & x_1 + x_2 - x_3 + x_4 = 2 \\
& x_1 + x_2 + x_3 = 3 \\
& x_1 - x_2 + x_3 = 1 \\
\text{Com: } & x_1, x_2, x_3 \geq 0
\end{aligned}$$

Para determinar uma solução básica viável para este problema procedemos do seguinte modo (sobretudo para problemas mais complexos que este). Tomamos o problema (PI):

$$\begin{aligned}
\text{Minimizar: } & x_5 + x_6 + x_7 \\
\text{Sujeito a: } & x_1 + x_2 - x_3 + x_4 + x_5 = 2 \\
& x_1 + x_2 + x_3 + x_6 = 3 \\
& x_1 - x_2 + x_3 + x_7 = 1 \\
\text{Com: } & x_1, x_2, x_3, x_5, x_6, x_7 \geq 0
\end{aligned}$$

e resolvemo-lo com o algoritmo simplex inicializando o problema com o conjunto de índices de base $B = \{5, 6, 7\}$, $A_B = I$ e $(x_1, x_2, x_3, x_4, x_5, x_6, x_7)^B = (0, 0, 0, 0, 2, 3, 1)$. A solução optimal obtida $(x_1^*, x_2^*, x_3^*, x_4^*, x_5^*, x_6^*, x_7^*)$ será tal que $x_5^* = x_6^* = x_7^* = 0$ sendo que $(x_1^*, x_2^*, x_3^*, x_4^*)$ é uma solução básica viável do problema (P).

Em alguns casos porém, é imediato encontrar uma solução básica viável sem ter que recorrer ao procedimento anterior. É por exemplo o caso de um problema do tipo $\max\{c^T x : Ax \leq b, x \geq 0\}$ com $b \geq 0$. Este problema é equivalente a $\max\{c^T x : Ax + y = b, x \geq 0, y \geq 0\}$ sendo $(0, b)$ uma solução básica viável associada ao conjunto de índices de base correspondente às variáveis de folga (variáveis associadas a y).

Exemplo 4.21. Considerando o problema de programação linear:

$$\begin{aligned}
\text{Maximizar: } & x_1 - x_2 - 2x_3 \\
\text{Sujeito a: } & x_1 + x_2 - x_3 \leq 2 \\
& x_1 + 2x_2 + x_3 \leq 3 \\
\text{Com: } & x_1, x_2, x_3 \geq 0
\end{aligned}$$

Este problema é equivalente ao problema

$$\begin{aligned}
\text{Maximizar: } & x_1 - x_2 - 2x_3 \\
\text{Sujeito a: } & x_1 + x_2 - x_3 + x_4 = 2 \\
& x_1 + 2x_2 + x_3 + x_5 = 3 \\
\text{Com: } & x_1, x_2, x_3 \geq 0
\end{aligned}$$

Neste caso inicializamos o problema com o conjunto de índices de base $B = \{4, 5\}$, $A_B = I$ e $(x_1, x_2, x_3, x_4, x_5)^B = (0, 0, 0, 2, 3)$.

Em seguida vamos mostrar que para aplicar o algoritmo simplex não é necessário calcular a matriz A_B^{-1} e os vectores $x_B = A_B^{-1}b$ e $\bar{c}_N^T = c_N^T - c_B^T A_B^{-1} A_N$ à parte. Estes podem ser calculados em conjunto num processo algo semelhante ao método de eliminação de Gauss.

Consideremos para um problema de programação linear

$$\max\{c^T x : Ax = b, x \geq 0\}$$

o seguinte *tableau*:

A	b
c^T	0

tendo já um conjunto de índices de base B , podemos tableau da forma

A_B	A_N	b
c_B^T	c_N^T	0

Note-se que as colunas de A_B podem estar misturadas com as colunas de A_N (o mesmo para c_B^T e c_N^T). Assim, multiplicando a parte superior do tableau por A_B^{-1} (o que pode ser feito usando o método de Gauss-Jordan) obtemos:

I	\bar{A}_N	\bar{b}
c_B^T	c_N^T	0

onde $\bar{A}_N = A_B^{-1} A_N$ e $\bar{b} = A_B^{-1} b$.

Para obter $\bar{c}_N^T = c_N^T - c_B^T \bar{A}_N$ basta aplicar o método de eliminação de Gauss à última linha de modo a anular todas a componente c_B^T , ficamos portanto com:

I	\bar{A}_N	\bar{b}
0	\bar{c}_N^T	$-c_B^T \bar{b}$

Note-se que na última coluna temos a parte não-nula da solução básica $x_B = \bar{b}$ e o simétrico do valor associado à solução básica $c^T x^B = c_B^T \bar{b}$.

Estamos agora no **passo 1** do algoritmo.

Se $\bar{c}_N \leq 0$ o algoritmo termina, a solução optimal corresponde a tomar $x_i = 0$ para as variáveis não-básicas ($i \in N$) e $x_i = \bar{b}_j$ se a coluna i do tableau for igual a e_j para as variáveis básicas ($i \in B$). O valor optimal será o simétrico do valor obtido no canto inferior direito do tableau.

Se $\bar{c}_N \not\leq 0$ então escolhe-se uma variável não-básica $\bar{c}_r > 0$ e vamos para o **passo 2**.

Se a coluna por cima de \bar{c}_r (coluna \bar{a}_r) é menor ou igual a zero ($\bar{a}_r \leq 0$) então o problema é ilimitado ($v(P) = +\infty$). Senão escolhe-se de entre as entradas positivas $\bar{a}_{i,r} > 0$ de \bar{a}_r uma $\bar{a}_{s,r}$ que tenha o menor rácio $\frac{\bar{b}_i}{\bar{a}_{i,r}}$. Essa entrada será um novo pivot do tableau, através do método de eliminação de Gauss coloca-se a zero todas as restantes entradas da coluna r do tableau (incluindo a entrada \bar{c}_r).

Obtemos um novo tableau com novo conjunto de índices de base $(B \setminus \{s\}) \cup \{r\}$. Voltamos ao passo 1.

Vejamos um exemplo:

Exemplo 4.22. Consideremos o problema de programação linear:

$$\begin{array}{ll} \text{Maximizar:} & x_1 \\ \text{Sujeito a:} & x_1 + 2x_2 \leq 11 \\ & 2x_1 + x_2 \leq 8 \\ & x_1 - x_2 \leq 1 \\ \text{Com:} & x_1, x_2 \geq 0 \end{array}$$

este problema é equivalente (juntando as variáveis de folga x_3, x_4 e x_5) ao problema:

$$\begin{array}{ll} \text{Maximizar:} & x_1 \\ \text{Sujeito a:} & x_1 + 2x_2 + x_3 = 11 \\ & 2x_1 + x_2 + x_4 = 8 \\ & x_1 - x_2 + x_5 = 1 \\ \text{Com:} & x_1, x_2, x_3, x_4, x_5 \geq 0 \end{array}$$

Temos neste caso, uma solução básica imediata dada pelos índices $B = \{3, 4, 5\}$ (corresponde às variáveis de folga). O tableau fica com o passo 0 automaticamente feito:

		↓	↓	↓	
1	2	1	0	0	11
2	1	0	1	0	8
1	-1	0	0	1	1
1	0	0	0	0	0

O \bar{c}_N^T é dado por $(1, 0)$ que é a parte da linha de baixo que não corresponde às variáveis básicas (assinaladas com setas). Como $\bar{c}_N \not\leq 0$, tomamos uma coluna r com $\bar{c}_r > 0$ (neste caso é a primeira) e procuramos de entre as entradas positivas (neste caso são todas) da coluna a que tem menor rácio $\frac{\bar{b}_i}{\bar{a}_{i,r}}$ (neste caso é a terceira pois $\frac{1}{1} < \frac{8}{2} < \frac{11}{1}$):

1	2	1	0	0	11
2	1	0	1	0	8
(1)	-1	0	0	1	1
<u>1</u>	0	0	0	0	0

tornamos, usando eliminação de Gauss, essa entrada num pivot:

0	3	1	0	-1	10
0	3	0	1	-2	6
1	-1	0	0	1	1
0	1	0	0	-1	-1

Mudamos então de conjunto de índices de base:

↓		↓	↓		
0	3	1	0	-1	10
0	3	0	1	-2	6
1	-1	0	0	1	1
0	1	0	0	-1	-1

Repetimos o processo até obtermos $\bar{c}_N \leq 0$:

0	3	1	0	-1	10
0	(3)	0	1	-2	6
1	-1	0	0	1	1
0	<u>1</u>	0	0	-1	-1

pois $\frac{6}{3} < \frac{10}{3}$.

↓	↓	↓			
0	0	1	-1	1	4
0	1	0	$\frac{1}{3}$	$-\frac{2}{3}$	2
1	0	0	$\frac{1}{3}$	$\frac{1}{3}$	3
0	0	0	$-\frac{1}{3}$	$-\frac{1}{3}$	-3

Temos portanto que a solução optimal é $(x_1, x_2, x_3, x_4, x_5) = (3, 2, 4, 0, 0)$ (no problema original é simplesmente $(x_1, x_2) = (3, 2)$ pois as restantes variáveis são de folga) e o valor optimal é 3.

Note-se que, se não ocorreram erros de cálculo, devem-se verificar as seguintes situações após o passo 0:

A coluna superior direita do tableau (correspondente a \bar{b}) é maior ou igual a zero.

O valor no canto inferior direito do tableau nunca aumenta à medida que o algoritmo é iterado.

Assim, se alguma destas situações não se verificar é porque houve algum erro de cálculo.

Correcção do algoritmo simplex.

Observação 4.23. Se num dado momento tivermos que a solução básica x^B é não-degenerada (logo $\lambda_r^* > 0$ pela observação 4.18) então no passo seguinte o valor alvo $c^T x$ é incrementado.

De facto, se $B' = (B \setminus \{s\}) \cup \{r\}$ for o novo conjunto de índices de base, a nova solução básica é dada por $x(\lambda_r^*)$ e o novo valor alvo será:

$$\begin{aligned} c^T x^{B'} &= c^T x(\lambda_r^*) = c_B^T x(\lambda_r^*)_B + c_N^T x(\lambda_r^*)_N = c_B^T (\bar{b} - \bar{a}_r \lambda_r^*)_B + c_N^T e_r \lambda_r^* = \\ &= c_B^T \bar{b} + (c_N^T - c_B^T \bar{A}_N) e_r \lambda_r^* = c_B^T x_B + \bar{c}_r \lambda_r^* > c_B^T x_B = c^T x^B \end{aligned}$$

Observação 4.24. Em cada iteração do algoritmo o vector x^B é um dos vértice de $P = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$ e estes são em número finito.

Resulta da observação 4.24 que se o algoritmo não termina é porque entra em ciclo fechado. Pela observação 4.23 tal só poderá acontecer se as iterações que formam o ciclo dão-se todas em soluções básicas degeneradas.

Embora hipoteticamente o algoritmo simplex possa entrar em ciclo, tal é muito raro de acontecer. Há, no entanto, regras que permitem evitar de todo qualquer possibilidade do algoritmo entrar em ciclo. Uma delas é a **regra de Bland** que se pode definir da seguinte maneira:

- (a) Escolhe-se para a coluna r com $\bar{c}_r > 0$ a primeira nessa condição. Ou seja, $r = \min\{i : \bar{c}_i > 0\}$.
- (b) Escolhe-se para novo pivot a entrada $\bar{a}_{s,r} > 0$, de entre as $\bar{a}_{i,r} > 0$ que satisfazem a condição de rácio $\bar{b}_i / \bar{a}_{i,r}$ mínimo, cujo o pivot correspondente à sua linha tem índice de coluna mais baixo³.

Esta regra simples garante que o algoritmo simplex não entra em ciclo.

Teorema 4.25. *Usando a regra de Bland, o algoritmo simplex termina em tempo finito.*

Dem. Suponhamos por absurdo que o algoritmo entra em ciclo (já vimos que é o único caso que pode impedir que o algoritmo termina em tempo finito). Para simplificar, podemos retirar todas as linhas e colunas do tableau que não contenham pivots que entram no ciclo sem que isso condicione a existência do ciclo. Tomamos então o tableau com tais linhas e colunas removidas:

A	b
c^T	0

Como foi observado a existência de um ciclo implica que todas as soluções básicas sejam degeneradas, resulta então, da remoção das linhas que não contenham os pivots do ciclo, que $b = 0$. Consideremos agora os dois seguintes tableau:

$$T_1 : \begin{array}{|c|c|} \hline A' & 0 \\ \hline c'^T & 0 \\ \hline \end{array} \quad \text{e} \quad T_2 : \begin{array}{|c|c|} \hline A'' & 0 \\ \hline c''^T & 0 \\ \hline \end{array}$$

onde T_1 corresponde ao tableau antes da última coluna entrar na base e T_2 corresponde ao tableau antes da última coluna sair da base sendo substituída por uma coluna p .

Como estamos a usar a regra de Bland temos que no tableau T_1 :

$$\forall_{i < n} c'_i \leq 0 \text{ e } c'_n > 0$$

e no tableau T_2 temos:

$$c''_p > 0, c''_n = 0, a''_{m,p} > 0$$

e

$$\forall_{i < m} a''_{i,p} \leq 0$$

³Se tivermos o cuidado de ao fim de cada passo do algoritmo permutarmos as linhas do tableau de modo a que os pivots apareçam em ordem crescente, então esta regra resume-se a escolher a entrada $\bar{a}_{s,r} > 0$ com menor índice s .

Note-se que estamos a assumir que temos n variáveis e m equações (i.e. A , A' e A'' são matrizes m por n). Consideremos então o seguinte vector $y \in \mathbb{R}^n$ definido por:

$$y_j = \begin{cases} -a''_{i,p} & \text{se } B''(i) = j \\ 1 & \text{se } j = p \\ 0 & \text{caso contrário} \end{cases}$$

Onde $B'' = \{B''(1), \dots, B''(m)\}$ é o conjunto de índices de base associado ao tableau T_2 . Do mesmo modo, $B' = \{B'(1), \dots, B'(m)\}$ é o conjunto de índices de base associado ao tableau T_2 . É fácil ver que y é solução de $A''x = 0$. Como, pelo algoritmo simplex, $A' = A_{B'}^{-1}A$ e $A'' = A_{B''}^{-1}A$, temos também $A'y = 0$. E sendo $c'^T = c^T - c_{B'}^T A'$ e $c''^T = c^T - c_{B''}^T A''$, temos que $c'^T y = c''^T y$.

Ora $c''^T y = c''_p > 0$ e por outro lado

$$c'^T y = \sum_{j=1}^n c'_j y_j = c'_p + \sum_{i=1}^m c'_{B''(i)} (-a''_{i,p}) = c'_p - c'_n a''_{m,p} - \sum_{i=1}^m c'_{B''(i)} a''_{i,p} < 0$$

Chegamos assim a uma contradição. Donde concluímos que o algoritmo não pode entrar em ciclo. \square

5. GRAFOS

Definições básicas.

Definição 5.1. Um *grafo não-orientado* (respectivamente *grafo não-orientado* ou *digrafo*) é um par ordenado (V, E) , onde V é um conjunto (finito) de elementos chamados *vértices* e E é um conjunto (finito) de pares não-ordenados (resp. orientados) de vértices.

Uma aresta $e = \{x, y\} \in E$ (usaremos também as notações $[x, y]$ ou simplesmente xy) diz-se *incidente* aos vértices X e y , e estes dizem-se *incidentes* à aresta e . Neste caso, também se diz que os vértices x e y são *adjacentes* entre si e que os *extremos* da aresta e . Duas arestas dizem-se *adjacentes* se partilham um mesmo vértice como um dos seus dois extremos.

Se os extremos de uma aresta e forma iguais (i.e. $e = xx$), dizemos que a aresta e é um *loop* (ou *laço fechado*).

Em alguns casos, pode-se considerar grafos com arestas paralelas também chamados de *multigrafos*. Em tais grafos, para cada par de vértices x e y pode haver múltiplas arestas (ditas *arestas paralelas*) que tenham x e y como extremos. Para definir rigorosamente um multigrafo ou consideramos que o conjunto das arestas E é um multiconjunto de pares não-ordenados de V , ou consideramos que o conjunto das arestas E não é formado por pares não-ordenados de V mas juntamos à definição de multigrafo uma aplicação (dita de incidência) que a cada elemento de E faz corresponder um par não-ordenado de V .

Sempre que neste texto for escrito a palavra grafo sem mais adjectivos está-se a referir apenas a um grafo não-orientado sem loops nem arestas paralelas (aquilo que muitos autores chamam de *grafo simples*).

Chama-se *ordem* de um grafo $G = (V, E)$ ao número de vértices $|V|$, e *tamanho* do grafo ao número de arestas $|E|$.

Definição 5.2. Dado um grafo $G = (V, E)$ com ordem $|V| = n$ e tamanho $|E| = m \geq 1$ definimos a *matriz de incidência* como sendo uma matriz $A \in \mathbb{R}^{n,m}$ em que as linhas são numeradas pelos vértices, as colunas são numeradas pelas arestas e a entrada genérica $a_{i,j}$ da matriz é 1 se o vértice i é incidente à aresta j , 0 caso contrário.

Observação 5.3. A matriz de incidência de um grafo é única a menos de permutações de linhas ou colunas.

Definição 5.4. Dado um grafo $G = (V, E)$ com ordem $|V| = n$ e definimos a *matriz de adjacência* como sendo uma matriz quadrada $A \in \mathbb{R}^{n,n}$ cujas as linhas e as colunas são numeradas pelos vértices e a entrada genérica $a_{i,j}$ da matriz é 1 se os vértices i e j são adjacentes, 0 caso contrário.

Observação 5.5. A matriz de adjacência de um grafo é simétrica e única a menos de conjugação por matrizes de permutação.

Definição 5.6. Chamamos *grau* de um vértice v ao número de aresta incidentes a v e geralmente denota-se tal número por $d(v)$.

Observação 5.7. O número de entradas iguais a 1 numa dada linha, quer da matriz de incidência quer da matriz de adjacência de um dado grafo, é igual ao grau do vértice associado a essa linha. O número de entradas iguais a 1 em cada coluna da matriz de incidência é sempre igual a 2.

Proposição 5.8. Se J é uma matriz de incidência de dado grafo G , então $JJ^T = A + D$ onde A é uma matriz de adjacência de G e D é uma matriz diagonal cuja entrada (i, i) é o grau do vértice associado à linha i da matriz de incidência J .

Dem. Exercício. □

Proposição 5.9.

$$\sum_{v \in V} d(v) = 2|E|$$

Dem. Exercício. □

Corolário 5.10. O número de vértices de um grafo com grau ímpar é par.

Dem. Exercício. □

Definição 5.11. Um *homomorfismo* entre dois grafos $G_1 = (V_1, E_1)$ e $G_2 = (V_2, E_2)$ é uma aplicação $\phi : V_1 \rightarrow V_2$ tal que para cada $e = [u, v]$ de G_1 o par $[\phi(u), \phi(v)]$ é uma aresta de G_2 . Ou seja, $[u, v] \in E_1 \Rightarrow [\phi(u), \phi(v)] \in E_2$.

Se aplicação ϕ for bijetiva, quer nos vértices quer nas arestas (i.e. $[u, v] \in E_1 \Leftrightarrow [\phi(u), \phi(v)] \in E_2$), dizemos que ϕ é um *isomorfismo* de grafos e que G_1 e G_2 são *isomorfos*.

Observação 5.12. Se dois grafos são isomorfos então possuem as mesmas matrizes de incidência e de adjacência.

Sai imediatamente desta observação que podemos recuperar o grafo original a partir das suas matrizes de incidência e de adjacência. Na verdade basta a matriz de incidência ou a matriz de adjacência (a menos que estejamos a considerar multigrafos em vez de grafos).

Definição 5.13. Um grafo $H = (U, F)$ diz-se *subgrafo* de $G = (V, E)$ se $U \subseteq V$ e $F \subseteq E$.

Seja $G = (V, E)$ um grafo e sejam $U \subseteq V$ e $F \subseteq E$. O *subgrafo induzido* por U , denotado por $G[U]$, é o grafo $(U, E[U])$ onde $E[U] := \{[u, v] \in E : u, v \in U\}$, dizemos então que $G[U]$ é um *subgrafo vértice-induzido*. O *subgrafo induzido* por F , denotado por $G[F]$, é o grafo (V, F) , dizemos então que $G[F]$ é um *subgrafo aresta-induzido*.

Observação 5.14. O subgrafo vértice-induzido $G[U]$ é o maior subgrafo de G que tem U como conjunto de vértices. O subgrafo aresta-induzido $G[F]$ é o maior subgrafo de G que tem F como conjunto de aretas.

Definição 5.15. O *complemento* de um grafo $G = (V, E)$ é o grafo $\overline{G} = (V, \overline{E})$ onde $\overline{E} := \{[u, v] : u, v \in V \text{ e } [u, v] \notin E\}$ (i.e. \overline{E} é o conjunto de pares não-ordenados de vértices que pertencem a E).

Um grafo $G = (V, E)$ diz-se *nulo* se $V = \emptyset$ (logo $E = \emptyset$), *vazio* se $E = \emptyset$ (sem que V seja necessariamente vazio) e *completo* se o seu complemento for vazio (i.e. $\overline{E} = \emptyset$).

Problema do caminho mais curto.

Definição 5.16. Um *percurso* ou *trilho* num grafo $G = (V, E)$ é uma sequência finita alternada de vértices e arestas $W : v_0, e_0, v_1, \dots, e_{n-1}, v_n$ com $v_i \in V$ e $e_i = [v_i, v_{i+1}] \in E$ para todo o i . Dizemos que W é um percurso- v_0v_n e que v_0 e v_n são os vértices finais do percurso.

Definição 5.17. O *comprimento* (usual) de um percurso é o número de arestas do percurso.

Proposição 5.18. Seja A a matriz de adjacência de um grafo G e seja $n \geq 1$. Então a entrada (i, j) de A^n é igual à quantidade de diferentes percursos- $v_i v_j$ com comprimento n .

Dem. Exercício. □

Definição 5.19. Um *caminho* é um percurso onde os vértices não se repetem (logo as arestas também não).

Definição 5.20. Dois vértices u e v de um dado grafo dizem-se *conectáveis* se existe um percurso- uv . Se qualquer par de vértices de dado grafo forem conectáveis então o grafo diz-se *conexo*.

Observação 5.21. A existência de um percurso- uv garante a existência de um caminho- uv . Pelo que na definição anterior podemos trocar percurso- uv por caminho- uv sem alterar o seu significado.

Definição 5.22. A noção de conectável forma uma relação de equivalência no conjunto dos vértices de um grafo G . Aos subgrafos induzidos por cada uma suas classe de equivalência chama-se *componentes conexas* de G .

Definição 5.23. Um *grafo com pesos* (ou *comprimentos*) é um grafo $G = (V, E)$ com uma *função peso* (ou *comprimento*) nas arestas $w : E \rightarrow \mathbb{R}_0^+$. O *peso* ou *comprimento* $w(P)$ de um caminho (ou percurso) $P : v_0, e_0, \dots, e_{n-1}, v_n$ é definido como sendo a soma do pesos das suas arestas:

$$w(P) = \sum_{e \in P \cap E} w(e).$$

Definição 5.24. Dado um grafo $G = (V, E)$ e uma função comprimento $w : E \rightarrow \mathbb{R}_0^+$, define-se a *distância* entre dois vértices u e v de G como sendo o menor comprimento entre os caminho- uv (caso u e v sejam conectáveis):

$$d_w(u, v) = \min\{w(P) : P \text{ é caminho-}uv\}$$

Caso u e v não sejam conectáveis, define-se $d_w(u, v) = +\infty$.

Observação 5.25. Para qualquer percurso- uv existe um caminho- uv mais curto (ou com peso menor). Logo na definição de distância entre dois vértices também podemos considerar percursos em vez de apenas caminhos.

Observação 5.26. Se $w \geq 0$ então d_w é uma pseudométrica ($d_w(x, y) = 0 \not\Rightarrow x = y$), se $w > 0$ então d_w é uma métrica e se $w = 1$ temos a métrica usual no grafo: $w(P) =$ número de aretas em P .

Definição 5.27. Um *problema do caminho mais curto* consiste em dado um grafo $G = (V, E)$ com uma função comprimento nas arestas $w : E \rightarrow \mathbb{R}_0^+$ e dois vértices u e v de G , encontrar um caminho- uv , P , que minimize o comprimento (ou peso). Ou seja, $w(P) = d_w(u, v)$.

Lema 5.28. Seja u_0 um vértice fixo de um dado grafo conexo $G = (V, E)$ e S um conjunto de vértices que contém u_0 . Seja $\bar{S} = V \setminus S$ o complemento de S .

Se $\bar{S} \neq \emptyset$, existe sempre um vértice $v_0 \in \bar{S}$ e um caminho- u_0v_0 , $P : u_0, u_1, \dots, u_k, v_0$ ⁴, tal que:

- (i) $u_j \in S$ para $j \leq k$;
- (ii) O caminho u_0, \dots, u_k é o caminho- u_0u_k mais curto em G e portanto também o é em $G[S]$ (o subgrafo de G induzido por S);
- (iii) $w(P) = d(u_0, v_0) = d(u_0, \bar{S}) := \min\{d(u_0, v) : v \in \bar{S}\}$

Donde resulta que

$$d(u_0, v_0) = d(u_0, u_k) + w([u_k, v_0]) = \min\{d(u_0, u) + w([u, v]) : u \in S, v \in \bar{S}\}.$$

Dem. Sendo \bar{S} um conjunto finito não-vazio é óbvio que existe um vértice $v' \in \bar{S}$ tal que $d(u_0, v') = d(u_0, \bar{S})$. Seja $P' : u_0, u_1, \dots, u_{k'}, v'$ um caminho- u_0v' com $w(P') = d(u_0, v')$ (i.e. P' é um caminho- u_0v' mais curto), e seja v_0 o primeiro vértice do caminho P' tal que $v_0 \in \bar{S}$. Se $v_0 = v'$ então temos o pretendido $P = P'$. Senão, $v_0 = u_{k+1}$ com $k+1 \leq k'$. Temos então que $P : u_0, u_1, \dots, u_k, v_0$ é um caminho que satisfaz (i) (claro do modo como P é definido) e (ii) (se u_0, u_1, \dots, u_k não fosse um caminho- u_0u_k minimal então $P' : u_0, u_1, \dots, u_k, \dots, u_{k'}, v'$ não

⁴Note-se que um caminho pode ser dado como uma sequência de apenas vértices (estando as arestas implícitas) ou de apenas arestas (estando os vértices implícitas).

seria um caminho- u_0v' minimal). Além disso v_0 e P satisfazem também (iii) pois o facto de P estar contido em P' implica que $d(u_0, v_0) = w(P) \leq w(P') = d(u_0, \bar{S})$ e por outro lado $v_0 \in \bar{S}$ implica $d(u_0, v_0) \geq d(u_0, \bar{S})$. É fácil verificar (exercício) que as igualdades $d(u_0, v_0) = d(u_0, u_k) + w([u_k, v_0]) = \min\{d(u_0, u) + w([u, v]) : u \in S, v \in \bar{S}\}$ resultam das condições (i), (ii) e (iii). \square

Este último resultado fundamenta o seguinte algoritmo para determinar o caminho mais curto.

Algoritmo de Dijkstra:

Dado um grafo conexo $G = (V, E)$ e um vértice inicial u_0 procedemos do seguinte modo:

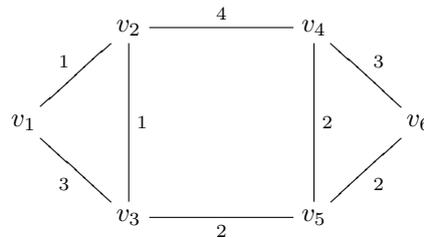
- Passo 0 Definimos $S := \{u_0\}$, $L(u_0) := 0$ e $L(v) := +\infty$ para cada $v \in V \setminus \{u_0\}$. Se $|V| = 1$, o algoritmo termina.
- Passo 1 Para cada $v \notin S$ procede-se do seguinte modo: Se o mínimo de $L(u) + w([u, v])$ para $u \in S$ e $[u, v] \in E$ é menor que $L(v)$, então redefine-se $L(v) := \min\{L(u) + w([u, v]) : u \in S, [u, v] \in E\}$ e o vértice v recebe uma nova etiqueta $(L(v), u)$.
- Passo 2 Determina-se $\min_{v \in \bar{S}} L(v)$ e seja v um vértice onde o mínimo é atingido.
- Passo 3 Redefine-se $S := S \cup \{v\}$. Se $S = V$, então termina-se o algoritmo, a distância de u_0 a v é dada por $L(v)$ e um⁵ caminho mais curto pode ser deduzido através das etiquetas $(L(v), u)$ (se v recebe a etiqueta $(L(v), u)$ então u é o vértice que precede v num caminho- u_0v mais curto). Senão retorna-se ao passo 1.

Definição 5.29. Um conjunto do tipo $\{[u, v] : u \in S, v \in \bar{S}\}$ chama-se um *corte* de S e denota-se por $\delta(S)$.

Observação 5.30. Este algoritmo pode ser aplicado a grafos orientados (digrafos) para determinar o caminho dirigido mais curto. Note-se que nesse caso $[u, v] \neq [v, u]$.

Observação 5.31. O algoritmo também pode ser ligeiramente modificado de modo a poder ser aplicado a grafos não conexos. Nesse caso, se após o passo 1 tivermos $\min_{v \in \bar{S}} L(v) = +\infty$ o algoritmo deve terminar pois os vértices em \bar{S} não conectáveis a u_0 . O S final será o conjunto dos vértices da componente conexa de G que contém u_0 e o algoritmo determina as distâncias $d(u_0, v) = L(v)$ bem como caminhos- u_0v mais curtos para todos os vértices de S .

Exemplo 5.32. Considere-se o seguinte grafo:



Para aplicar o algoritmo vamos construir a seguinte tabela, na primeira coluna colocamos os vértices por ordem de inclusão no conjunto S (as linhas da tabela vão sendo construídas à medida que o algoritmo é aplicado), na segunda coluna colocamos as respectivas etiquetas. Quanto às restantes colunas, estas serão indexadas pelos vértices do grafo e servem para o cálculo dos $L(v_i)$'s. Na linha em que o vértice v_i surge na primeira coluna, a entrada correspondente ao vértice v_j será um X se v_j já estiver em S (i.e. v_j aparece antes de v_i na primeira coluna), $L(v_i) + w([v_i, v_j])$ se $v_j \notin S$ e $[v_i, v_j]$ for uma aresta de G ou ∞ se $v_j \notin S$ e v_i for adjacente a v_j . Por exemplo a primeira linha da tabela será:

S	etiqueta	v_2	v_3	v_4	v_5	v_6
v_1	0	1	3	∞	∞	∞

⁵Note-se que no passo 2 o vértice a escolher pode não ser único, pelo que o caminho mais curto não será necessariamente único.

Em seguida determina-se de entre as entradas correspondente às linhas de vértices de S e às colunas de vértices de \bar{S} (coloridas de verde) aquela que tem menor valor. Se tal entrada corresponder aos vértices $v_i \in S$ e $v_j \in \bar{S}$ e tiver valor α construímos a nova linha correspondente ao vértice v_j com etiqueta (α, v_i) . Por exemplo o próximo passo será:

S	etiqueta	v_2	v_3	v_4	v_5	v_6
v_1	0	(1)	3	∞	∞	∞
v_2	$(1, v_1)$	X	$1 + 1$	$1 + 4$	∞	∞

seguinto-se:

S	etiqueta	v_2	v_3	v_4	v_5	v_6
v_1	0	1	3	∞	∞	∞
v_2	$(1, v_1)$	X	(2)	5	∞	∞
v_3	$(2, v_2)$	X	X	∞	$2 + 2$	∞

e

S	etiqueta	v_2	v_3	v_4	v_5	v_6
v_1	0	1	3	∞	∞	∞
v_2	$(1, v_1)$	X	2	5	∞	∞
v_3	$(2, v_2)$	X	X	∞	(4)	∞
v_5	$(4, v_3)$	X	X	$4 + 2$	X	$4 + 2$

e

S	etiqueta	v_2	v_3	v_4	v_5	v_6
v_1	0	1	3	∞	∞	∞
v_2	$(1, v_1)$	X	2	(5)	∞	∞
v_3	$(2, v_2)$	X	X	∞	4	∞
v_5	$(4, v_3)$	X	X	6	X	6
v_4	$(5, v_2)$	X	X	X	X	$5 + 3$

e finalmente

S	etiqueta	v_2	v_3	v_4	v_5	v_6
v_1	0	1	3	∞	∞	∞
v_2	$(1, v_1)$	X	2	5	∞	∞
v_3	$(2, v_2)$	X	X	∞	4	∞
v_5	$(4, v_3)$	X	X	6	X	(6)
v_4	$(5, v_2)$	X	X	X	X	8
v_6	$(6, v_5)$	X	X	X	X	X

Temos por exemplo que a distância entre v_1 e v_6 é 6 e um caminho- v_1v_6 optimal (mais curto) pode ser encontrado andando de v_6 para trás até v_1 seguindo as etiquetas: v_6 tem v_5 na etiqueta, v_5 tem v_3 na etiqueta, v_3 tem v_2 na etiqueta, v_2 tem v_1 na etiqueta, logo um caminho- v_1v_6 optimal é $P : v_1, v_2, v_3, v_5, v_6$.

Árvores geradoras de peso mínimo.

Definição 5.33. Um *círculo* é um percurso fechado (i.e. os vértices finais coincidem). Se num círculo não existirem vértices repetidos este diz-se um *ciclo*.

Definição 5.34. Um grafo sem ciclos diz-se um *grafo acíclico* ou *floresta*. Um grafo acíclico conexo é chamado de *árvore*

Proposição 5.35. Se um grafo $G = (V, E)$ é acíclico então existe pelo menos um vértice com grau menor ou igual a um. Se $|V| \geq 2$ e $E \neq \emptyset$ então existem pelo menos dois vértices de grau um.

Dem. Exercício. □

Corolário 5.36. Se um grafo $G = (V, E)$ é acíclico então $|E| \leq |V| - 1$.

Dem. Exercício. □

Proposição 5.37. Se $G = (V, E)$ é conexo então $|E| \geq |V| - 1$.

Dem. Exercício. □

Corolário 5.38. Se $G = (V, E)$ é uma árvore então $|E| = |V| - 1$.

Dem. Exercício. □

Proposição 5.39. Sejam G_1, G_2 e G_3 três grafos tais que $E(G_1) \subseteq E(G_2) \subseteq E(G_3)$ e $V(G_1) = V(G_2) = V(G_3)$. Se G_2 é uma árvore então G_1 é acíclico (não conexo se $E(G_1) \neq E(G_2)$) e G_3 é conexo (cíclico se $E(G_2) \neq E(G_3)$).

Dem. Exercício. □

Proposição 5.40. Seja $T = (V, E)$ um grafo. Então as seguintes afirmações são equivalentes:

- (i) T é uma árvore.
- (ii) T é conexo e cada aresta de T é uma ponte⁶.
- (iii) Existe um único caminho entre cada par de vértices.
- (iv) T é conexo e $|E| = |V| - 1$.
- (v) T é acíclico e $|E| = |V| - 1$.

Dem. Exercício. □

Definição 5.41. Dado um grafo conexo $G = (V, E)$, dizemos que $T = (V', E')$ é uma *árvore geradora* de G se $V' = V$, $E' \subseteq E$ e $T = (V', E')$ é uma árvore.

Observação 5.42. Todo o grafo conexo possui pelo menos uma árvore geradora. O *problema da árvore geradora de peso mínimo* consiste em determinar num grafo com pesos nas aresta uma árvore geradora com peso mínimo.

Proposição 5.43. (i) Se $T = (V, E)$ é uma árvore e $[v_i, v_j] \notin E$, então o grafo $G = (V, E \cup \{[v_i, v_j]\})$ contém exactamente um ciclo C' .
(ii) Se C' é o único ciclo em (i) e $[v_k, v_l] \in C' \setminus \{[v_i, v_j]\}$, então o grafo $G' = (V, (E \cup \{[v_i, v_j]\}) \setminus \{[v_k, v_l]\})$ é uma árvore.

Dem. (i) Se C' é um ciclo em G então $[v_i, v_j] \in C'$ pois senão C' estaria contido em T e T não seria uma árvore. Como T é uma árvore, existe um único caminho a ligar v_i a v_j , logo, juntamente com a aresta $[v_i, v_j]$, este forma o único ciclo em G .

(ii) Se retirarmos uma aresta a um ciclo de um grafo conexo este permanece conexo. Logo o grafo G' é conexo. Como o número de arestas de G' é igual ao número de arestas de T e T é um árvore (logo $|E| = |V| - 1$) temos que G' é uma árvore. □

Algoritmo de Kruskal:

Dada um grafo conexo $G = (V, E)$ com $|V| = n$ e $|E| = m$ podemos obter uma árvore geradora $T = (V, F)$ com peso mínimo procedendo do seguinte modo.

⁶Uma aresta diz-se uma ponte se ao ser removida o número de componentes conexas do grafo aumenta.

- Passo 0 Ordena-se as aresta por ordem crescente dos seus pesos: $w(e_1) \leq w(e_2) \leq \dots \leq w(e_m)$.
 Inicializa-se com $F := \emptyset$ e $k = 1$.
- Passo 1 Se o grafo $(V, F \cup \{e_k\})$ é acíclico, então redefine-se $F := F \cup \{e_k\}$.
- Passo 2 Se $|F| = n - 1$, então o algoritmo termina $T = (V, F)$ é uma árvore geradora de peso mínimo. Senão, redefine-se $k := k + 1$ e retorna-se ao passo 1.

Observação 5.44. No passo 1 do algoritmo de Kruskal temos que verificar se o grafo obtido é acíclico. Tal pode ser feito etiquetando os vértices pelas componentes conexas em (V, F) . Inicializamos $c : V \rightarrow \mathbb{N}$ com $c(v_i) = i$ (cada vértice forma uma componente conexa de (V, \emptyset)). No passo 1 sejam v_i e v_j os extremos da aresta e_k , se $c(v_i) = c(v_j)$ então o grafo $(V, F \cup \{e_k\})$ não é acíclico e tanto F como c mantêm-se inalterados, se $c(v_i) \neq c(v_j)$ então $(V, F \cup \{e_k\})$ é acíclico e redefinimos $F := F \cup \{e_k\}$ e $c(v) = c(v_i)$ sempre que $c(v) = c(v_j)$.

Teorema 5.45. Para qualquer grafo conexo com pesos nas arestas, o algoritmo de Kruskal termina produzindo uma árvore geradora de peso mínimo.

Dem. É claro que o algoritmo termina (há apenas um número finito de aresta a considerar) e que o output (V, F) é uma árvore geradora (pois é acíclico e $|F| = n - 1$). Falta ver que tem peso mínimo.

Seja $F = \{e_{i_1}, \dots, e_{i_{m-1}}\}$ o output do algoritmo. Seja M uma árvore geradora de peso mínimo que partilhe com a árvore $T = (V, F)$ o maior número possível de arestas.

Se T é uma árvore geradora minimal então $M = T$. Senão existe pelo menos uma aresta de T que não está em M , seja e_{i_k} a aresta nessa condição com menor índice.

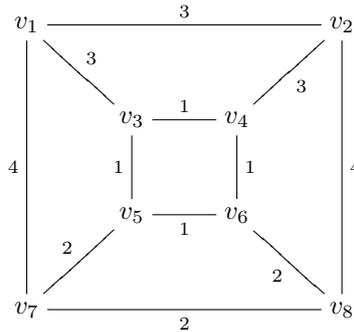
$M \cup \{e_{i_k}\}$ tem um ciclo e este ciclo contém uma aresta e que não está em T (caso contrário todo o ciclo estava contido em T o que é impossível pois T é acíclico). Temos então que $(M \cup \{e_{i_k}\}) \setminus \{e\}$ é uma árvore.

Por M ser minimal $w(e) \leq w(e_{i_k})$.

Por outro lado, como $e \notin T$, e estava disponível quando e_{i_k} foi escolhido, logo $w(e) \geq w(e_{i_k})$.

Portanto $w(e) = w(e_{i_k})$ e $(M \cup \{e_{i_k}\}) \setminus \{e\}$ é uma árvore geradora de peso mínimo com mais arestas em comum com T que M o que contradiz a definição de M . \square

Exemplo 5.46. Considere-se o seguinte grafo

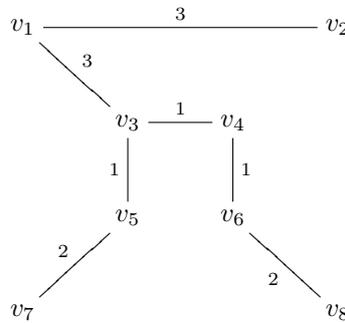


com os pesos das arestas representados sobre estas. Para aplicar o algoritmo de Kruskal podemos proceder do seguinte modo. Vamos construindo uma tabela pondo na primeira coluna as arestas ordenadas por ordem crescente dos seus pesos, na segunda coluna colocamos os respectivos pesos. As seguintes colunas serão indexadas pelos vértices de G e por baixo destes colocamos as etiquetas provisórias das suas componentes conexas $c(v_i)$. Na última coluna colocamos sim ou não consoante a aresta que indexa a linha irá ser incluída ao não no conjunto F , tal acontecerá se e só se $c(v_i) \neq c(v_j)$ (onde v_i e v_j são os vértices extremos da aresta em questão). Se a resposta for não então os $c(v)$'s permanecem inalterados na linha seguinte, se for sim então igualamos as etiquetas $c(v_i)$ e $c(v_j)$. Quando as etiquetas foram todas iguais o algoritmo para.

aresta	peso	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	$\in F?$
$[v_3, v_4]$	1	1	2	(3)	(4)	5	6	7	8	sim
$[v_3, v_5]$	1	1	2	(3)	3	(5)	6	7	8	sim
$[v_4, v_6]$	1	1	2	3	(3)	3	(6)	7	8	sim
$[v_5, v_6]$	1	1	2	3	3	(3)	(3)	7	8	não
$[v_5, v_7]$	2	1	2	3	3	(3)	3	(7)	8	sim
$[v_6, v_8]$	2	1	2	3	3	(3)	3	3	(8)	sim
$[v_7, v_8]$	2	1	2	3	3	3	3	(3)	(3)	não
$[v_1, v_2]$	3	(1)	(2)	3	3	3	3	3	3	sim
$[v_1, v_3]$	3	(1)	1	(3)	3	3	3	3	3	sim
$[v_2, v_4]$	3	1	1	1	1	1	1	1	1	Stop
$[v_1, v_7]$	4									
$[v_2, v_8]$	4									

A árvore gerada obtida será formada pelo conjunto de arestas

$$F = \{[v_3, v_4], [v_3, v_5], [v_4, v_6], [v_5, v_7], [v_6, v_8], [v_1, v_2], [v_1, v_3]\}$$



Fluxos e cortes. Recordemos que um *digrafo* ou *grafo orientado* ou *grafo dirigido* é um grafo onde as arestas têm uma orientação. Uma aresta orientada (também designada por *arco*) e é constituída por um par ordenado de vértices, $e = (u, v)$. Ao primeiro vértice, u , chamamos *vértice inicial* e ao segundo, v , chamamos *vértice final*. Para cada vértice v chamamos de *estrela de saída* ao conjunto, denotado por $\delta^+(v)$, das aresta orientadas com vértice inicial v , e chamamos de *estrela de entrada* ao conjunto, denotado por $\delta^-(v)$, das aresta orientadas com vértice final v .

Definição 5.47. Uma *rede* é um 4-tuplo $\mathcal{N} = (V, E, d, b)$ onde

- (i) (V, E) é um digrafo;
- (ii) $d : E \rightarrow \mathbb{R}_0^+$ é uma *função capacidade* onde $d_e = d(e) \geq 0$ denota a capacidade da aresta e ;
- (iii) $b : V \rightarrow \mathbb{R}$ é uma *função de fornecimento* onde $b_v = b(v)$ denota o fornecimento no vértice v .

Podemos olhar para uma função $x : E \rightarrow \mathbb{R}$ como sendo um vector de variáveis, uma por cada aresta, $x = (x_e : e \in E) \in \mathbb{R}^E$. Sempre que $F \subseteq E$ denotamos $x(F) := \sum_{e \in F} x_e$.

Definição 5.48. Um *fluxo* numa rede $\mathcal{N} = (V, E, d, b)$ é um vector $x \in \mathbb{R}^E$ que satisfaz o seguinte sistema linear:

- (i) $x(\delta^+(v)) - x(\delta^-(v)) = b_v$ para todo o vértice v ;
- (ii) $0 \leq x_e \leq d_e$ para toda a aresta e .

Observação 5.49. Se $b_v = 0$ para qualquer vértice então o fluxo x é chamado de *circulação* e as equações $x(\delta^+(v)) - x(\delta^-(v)) = 0$ mais não são que a primeira lei de Kirchhoff para os circuitos eléctricos (o que entra é igual ao que sai).

Definição 5.50. Dado um grafo orientado $D = (V, E)$ com ordem $|V| = n$ e tamanho $|E| = m \geq 1$ definimos a *matriz de incidência* como sendo uma matriz $A \in \mathbb{R}^{n, m}$ em que as linhas são numeradas

pelos vértices, as colunas são numeradas pelas arestas e a entrada genérica $a_{i,j}$ da matriz é 1 se o vértice i for o vértice inicial da aresta j , -1 se for o vértice final e 0 nos restantes casos.

Observação 5.51. Um fluxo é uma solução do sistema linear

$$Ax = b, \quad 0 \leq x \leq d$$

onde A é a matriz de incidência do digrafo $D = (V, E)$, b é o vector de fornecimento e d é o vector capacidade da rede \mathcal{N} . Assim sendo, o conjunto dos fluxos de uma rede forma um poliedro limitado $\{x \in \mathbb{R}^E : Ax = b, 0 \leq x \leq d\}$.

Observação 5.52. Se uma rede $\mathcal{N} = (V, E, d, b)$ possui pelo menos um fluxo então $\sum_{v \in V} b_v = 0$.

Definição 5.53. Dados dois vértices distintos, uma *fonte* s (i.e. $\delta^-(s) = \emptyset$) e um *poço* t (i.e. $\delta^+(t) = \emptyset$), um fluxo $x \in \mathbb{R}^E$ numa rede $\mathcal{N} = (V, E, d, b)$ é chamado de *fluxo- st* se $b_v = 0$ para todo o vértice v distinto de s e t .

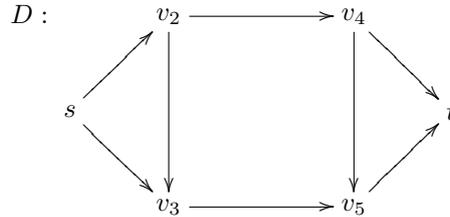
Observação 5.54. Da observação 5.52 resulta que, para um fluxo- st , temos $x(\delta^+(s)) = x(\delta^-(t))$.

Definição 5.55. Ao valor $f(x) = x(\delta^+(s))$ chamamos o *valor* do fluxo x .

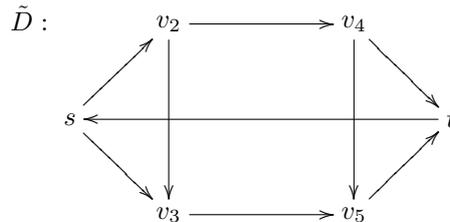
Observação 5.56. Num caso mais geral em que $\delta^-(s) \neq \emptyset$ ou $\delta^+(t) \neq \emptyset$, chamamos valor do fluxo- st a $f(x) = x(\delta^+(s)) - x(\delta^-(s)) (= x(\delta^-(t)) - x(\delta^+(t)))$.

Definição 5.57. Dado um digrafo $D = (V, E)$ e dois vértices distintos s e t definimos o *grafo aumentado* como sendo $\tilde{D} = (V, E \cup \{(t, v)\})$. Ou seja, junta-se artificialmente uma aresta orientada (t, v) ligando o poço t à fonte s .

Exemplo 5.58. O grafo aumentado do digrafo



é o digrafo



Observação 5.59. Uma circulação em \tilde{D} corresponde a um fluxo- st em D e vice-versa.

Dado um digrafo $D = (V, E)$ seja \mathcal{P} o conjunto de todos os caminhos dirigidos⁷ de D , e seja \mathcal{C} o conjunto de todos os ciclos dirigidos de D .

Se $P \in \mathcal{P}$ é um caminho- st então P dá origem a um fluxo- st de valor z dado por $x = z\chi^P$, ou seja,

$$x(e) = \begin{cases} z & \text{se } e \in P \\ 0 & \text{se } e \notin P \end{cases}$$

de igual modo $C \in \mathcal{C}$ induz uma circulação $x = z\chi^C$.

⁷Um percurso (caminho, circuito, ciclo) num digrafo é um percurso (caminho, circuito, ciclo) *dirigido* ou *orientado* se cada uma das suas arestas orientadas for precedida pelo seu vértice inicial e sucedida pelo seu vértice final na sequência que define o percurso.

Consideremos a seguinte função (chamada de *função fluxo de caminho e ciclo*) $g : \mathcal{P} \cup \mathcal{C} \rightarrow \mathbb{R}_+$, chamamos $g(P)$ o fluxo no caminho P e $g(C)$ o fluxo no ciclo C . g pode ser vista como um vector em $PC_+ := \mathbb{R}_+^{\mathcal{P} \cup \mathcal{C}}$. Define-se a transformação linear $T : PC_+ \rightarrow \mathbb{R}_+^E$ dada por

$$T(g) = \sum_{P \in \mathcal{P}} g_P \chi^P + \sum_{C \in \mathcal{C}} g_C \chi^C$$

O próximo resultado (conhecido por **teorema da decomposição de fluxo**) mostra que qualquer fluxo é imagem desta aplicação.

Teorema 5.60. *Cada fluxo $x \in \mathbb{R}^E$ num digrafo $D = (V, E)$ pode ser decomposto como uma soma de fluxos de caminho e fluxos de ciclo, ou seja, $x = T(g)$ para algum $g \in PC_+$.*

De facto, definindo $n := |\{v \in V : b_v > 0\}|$ e $m := |\{e \in E : x_e > 0\}|$, g pode ser escolhido de forma a que

- (i) *se $g_P > 0$ para um dado caminho P então este vai de um vértice v fornecedor ($b_v > 0$) para um vértice u de procura ($b_u < 0$);*
- (ii) *no máximo $n + m$ caminhos e ciclos têm fluxo não-nulo;*
- (iii) *no máximo m ciclos têm fluxo não-nulo.*

Dem. Vamos mostrar a veracidade do teorema por indução em $n + m$. Para $n + m = 0$ o resultado é trivial, $x = 0$.

Seja x um fluxo com $n + m > 0$. Se $n = 0$ temos, pela observação 5.52, que $b_v = 0 \forall v \in V$ logo x é uma circulação. Nesse caso, existe um ciclo C onde todas as arestas têm fluxo positivo (i.e. $x_e > 0 \forall e \in C$), basta pegar numa aresta com fluxo não-nulo (existe pois $m > 0$) e continuando o percurso, saindo em cada vértice que se entra por uma aresta com fluxo positivo⁸, até este se fechar. Seja $\varepsilon = \min\{x_e : e \in C\}$, temos então que $x = \varepsilon \chi^C + x'$ onde x' é um fluxo com pelo menos menos uma aresta com fluxo positivo ($m(x') \leq m - 1$), logo podemos usar a hipótese de indução para concluir a tese do teorema.

Se $n > 0$ então vamos construindo um percurso H começando num vértice de fornecimento s , prologando por uma aresta com fluxo positivo que saia do último que se entrou até que este percurso feche sobre si próprio ou que acabe num vértice t de onde não pode sair (caso em que $x(\delta^+(t)) = 0$ pelo que t será necessariamente um vértice de procura). Se o percurso H fechar sobre si próprio temos um ciclo C como anteriormente e procedemos do mesmo modo. No outro caso temos que H é caminho- st na condição (i) do teorema. Seja $\varepsilon = \min\{x_e : e \in H\} \cup \{b_s\}$, temos então que $x = \varepsilon \chi^H + x'$ onde x' é um fluxo para o vector de fornecimento b' (com $b'_s = b_s - \varepsilon$, $b'_t = b_t + \varepsilon$ e $b'_v = b_v$ para os restantes vértices) com pelo menos menos uma aresta com fluxo positivo ($m(x') \leq m - 1$, caso em que $\varepsilon = \min\{x_e : e \in H\}$) ou com menos um vértice de fornecimento ($n(x') = n - 1$, caso em que $\varepsilon = b_s$ ⁹) logo podemos também usar a hipótese de indução para concluir a tese do teorema. \square

Corolário 5.61. *Uma circulação num digrafo $D = (V, E)$ pode ser decomposta na soma de um número máximo de $|E|$ ciclos dirigidos.*

Definição 5.62. Um *corte* num digrafo $D = (V, E)$ é um conjunto de arcos da forma $\delta^+(S) := \{(v, w) \in E : v \in S \text{ e } w \notin S\}$ para um subconjunto próprio de V (i.e. $\emptyset \subsetneq S \subsetneq V$). Se S for tal que $s \in S$ e $t \notin S$ dizemos que o corte $\delta^+(S)$ é um *corte- st* . A *capacidade* de um corte $C = \delta^+(S)$ é o valor $d(C) = \sum_{e \in C} d_e$ onde d é a função capacidade da rede.

Lema 5.63. *O valor de um qualquer fluxo- st não é maior que a capacidade de um qualquer corte- st .*

Dem. Seja x um fluxo- st com valor $f(x) = x(\delta^+(s))$ e $C = \delta^+(S)$ um corte- st . Como, para todo o $v \in S \setminus \{s\}$, $x(\delta^+(v)) - x(\delta^-(v)) = 0$, temos que

$$f(x) = \sum_{v \in S} x(\delta^+(v)) - x(\delta^-(v))$$

⁸Existe pois se há uma aresta a entrar com fluxo positivo tem de existir uma aresta a sair com fluxo positivo para que o fornecimento seja nulo, $b_v = 0$.

⁹Neste caso temos que $b'_s = 0$ e $b'_t \leq 0$ pois $b_t = -x(\delta^-(t)) \leq -\varepsilon$.

por outro lado

$$\sum_{v \in S} x(\delta^+(v)) - x(\delta^-(v)) = \sum_{e \in \delta^+(S)} x_e - \sum_{e \in \delta^-(S)} x_e = x(\delta^+(S)) - x(\delta^-(S)) \leq x(\delta^+(v)) \leq d(\delta^+(v))$$

Concluimos portanto que o valor $f(x)$ do fluxo x não é maior que a capacidade $d(C)$ do corte C . \square

Definição 5.64. O problema do fluxo máximo consiste em, dado um digrafo com vértices distintos¹⁰ s e t , determinar um fluxo- st , $x \in \mathbb{R}^E$, com valor $f(x) = x(\delta^+(s))$ o maior possível.

Observação 5.65. Existe, para qualquer digrafo com vértices distintos s e t , um fluxo- st com valor máximo pois a função valor $f(x) = \sum_{e \in \delta^+(s)} x_e$ é uma função contínua no conjunto dos fluxos- st , $\{x \in \mathbb{R}^E : x(\delta^+(v)) - x(\delta^-(v)) \forall v \in V \setminus \{s, t\}, 0 \leq x_e \leq d_e \forall e \in E\}$, que é um conjunto compacto (mais concretamente é um poliedro limitado).

Definição 5.66. O problema do corte mínimo consiste em, dado um digrafo com vértices distintos s e t , determinar um corte- st , $C = \delta^+(S)$, com capacidade $d(C) = d(\delta^+(S))$ a menor possível.

Proposição 5.67. O valor máximo dos fluxo- st nunca é maior que a capacidade mínima dos corte- st .

Dem. Exercício. \square

Dado um fluxo- st , x , num digrafo $D = (V, E)$ e um vértice $v \in V \setminus \{s\}$, considere-se P um caminho- sv não necessariamente dirigido em D . Há em P dois tipos de arcos, aqueles que estão orientados no sentido sv , designados por *arcos de avanço*, e aqueles que estão orientados no sentido vs , designados por *arcos de recuo*.

Definição 5.68. Seja P^+ e P^- os conjuntos dos arcos de avanço e dos arcos de recuo respectivamente. Dizemos que um tal caminho- sv , P , é um *caminho x -aumentador* a v se $x_e < d_e$ para todo $o e \in P^+$ e $x_e > 0$ para todo $e \in P^-$. Se $v = t$, então P chama-se um *caminho x -aumentador*.

Lema 5.69. Um fluxo- st , x , é máximo se e só se não existem caminhos x -aumentadores.

Dem. Seja $S(x) = \{v \in V : \text{existem caminhos } x\text{-aumentadores a } v\}$.

Se existem caminhos x -aumentadores então $t \notin S(x)$. Seja P um caminho x -aumentadores e sejam $\varepsilon^+ = \min\{d_e - x_e : e \in P^+\} > 0$, $\varepsilon^- = \min\{x_e : e \in P^-\} > 0$ e $\varepsilon = \min\{\varepsilon^+, \varepsilon^-\} > 0$, então

$$\bar{x}_e = \begin{cases} x_e + \varepsilon & \text{se } e \in P^+ \\ x_e - \varepsilon & \text{se } e \in P^- \\ x_e & \text{se } e \notin P \end{cases}$$

é um fluxo- st com $f(\bar{x}) = f(x) + \varepsilon > f(x)$, logo x não é um fluxo máximo.

Se não existem caminhos x -aumentadores então $t \notin S(x)$. Temos então que $C(x) = \delta^+(S(x))$ é um corte- st . Para qualquer arco $e = (i, j) \in C(x)$ temos $x_e = d_e$ (senão $j \in S(x)$ e $e \notin C(x)$). Do mesmo modo para qualquer arco $e \in \delta^-(S(x)) = \{(i, j) : i \notin S(x) \text{ e } j \in S(x)\}$ temos $x_e = 0$. O valor do fluxo x é

$$f(x) = \sum_{v \in S(x)} x(\delta^+(v)) - x(\delta^-(v)) = x(\delta^+(S(x))) - x(\delta^-(S(x))) = d(\delta^+(S(x))) - 0 = d(C(x))$$

logo, pela proposição 5.67, x é um fluxo máximo. \square

Note-se que na demonstração do lema anterior é provada a existência de um corte- st com capacidade igual ao fluxo máximo. Assim juntamente com a proposição 5.67 temos demonstrado o seguinte resultado, conhecido por **teorema fluxo máximo corte mínimo**.

Teorema 5.70. Para cada digrafo com função capacidade nos arcos e vértices distintos s e t , o valor de fluxo- st máximo é igual à capacidade de corte- st mínima.

¹⁰Geralmente s é uma fonte e t um poço.

O próximo teorema, conhecido por **teorema de Menger**, é consequência deste dois últimos resultados.

Teorema 5.71. *O número máximo de caminhos- st dirigidos disjuntos nos arcos num digrafo é igual ao número mínimo de arcos num corte- st .*

Dem. Seja k o número mínimo de arcos num corte- st e l o número máximo de caminhos- st dirigidos disjuntos nos arcos. Tomamos por função capacidade a função constante igual a 1 (i.e. $d_e = 1 \forall e \in E$).

Temos então que k é a capacidade mínima de um corte- st .

Sejam P_1, \dots, P_l , l caminhos- st dirigidos disjuntos nos arcos. Consideremos o fluxo $x = \sum_{i=1}^l \chi^{P_i}$. Tal fluxo tem valor $f(x) = l$ e, além disso, não possui caminhos x -aumentados (pois um tal caminho x -aumentado seria um caminho- st dirigido que não partilha nenhum arco com P_1, \dots, P_l) logo x é um fluxo máximo e portanto $l = k$. \square

O próximo algoritmo (Ford-Fulkerson) mostra como obter um fluxo- st maximal.

Algoritmo genérico do fluxo máximo:

Passo 0 Inicializa-se com $x := 0$.

Passo 1 Determina-se o conjunto $S(x)$ (i.e. o conjunto dos vértices para os quais existem caminhos x -aumentadores). Se $t \notin S(x)$, termina-se: o fluxo x é máximo e o corte $\delta^+(S(x))$ é mínimo. Senão tomamos um caminho P x -aumentador.

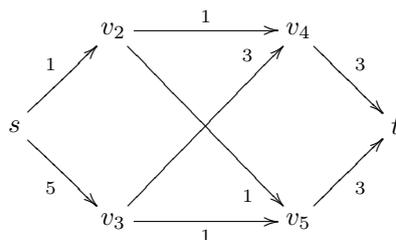
Passo 2 Redefine-se x , somando ε nos arcos de avanço de P e subtraindo ε nos arcos de recuo de P , onde $\varepsilon = \min(\{d_e - x_e : e \in P^+\} \cup \{x_e : e \in P^-\})$. Retorna-se ao passo 1.

Observação 5.72. Se a função capacidade contiver valores irracionais não é líquido que o algoritmo termina. No entanto o próximo resultado garante que tal acontece se as capacidades dos arcos forem todas racionais.

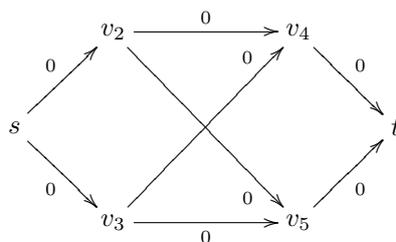
Proposição 5.73. *Sempre que a função capacidade d tiver apenas valores inteiros (racionais) o algoritmo termina em tempo finito com um fluxo com valores inteiros (racionais).*

Dem. Exercício. \square

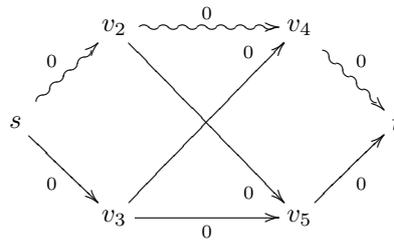
Exemplo 5.74. Consideremos o problema do fluxo- st máximo no seguinte digrafo (com as capacidades das arestas indicadas sobre estas):



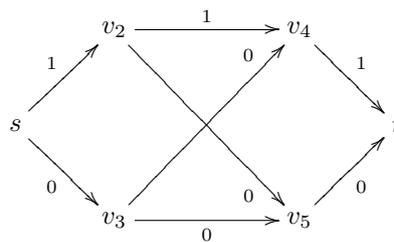
Iniciamos com o fluxo nulo:



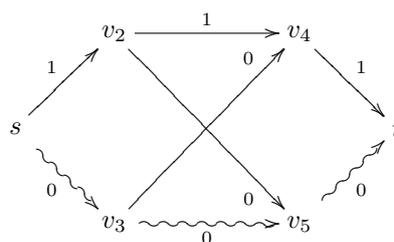
Determinamos o $S(x)$, que neste caso é formado por todos os vértices incluindo t . Logo existe um caminho x -aumentador, por exemplo:



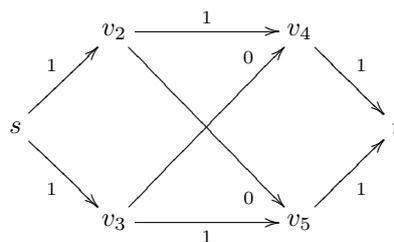
Temos que $\varepsilon = \min\{1, 1, 3\} = 1$, logo ficamos com o novo fluxo:



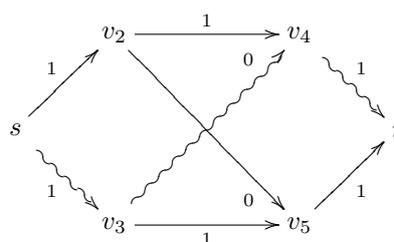
$S(x) = \{s, v_2, v_3, v_4, v_5, t\}$, tomamos novamente um caminho x -aumentador, por exemplo:



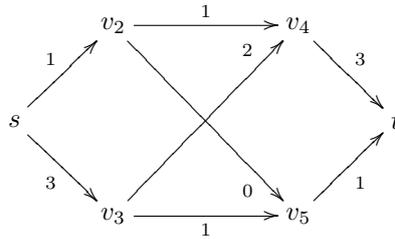
Temos que $\varepsilon = \min\{5, 1, 3\} = 1$, logo ficamos com o novo fluxo:



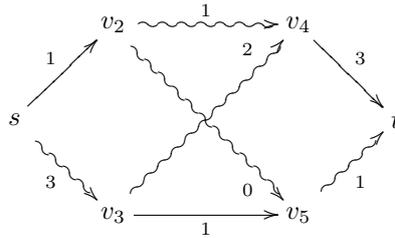
$S(x) = \{s, v_2, v_3, v_4, v_5, t\}$, tomamos novamente um caminho x -aumentador, por exemplo:



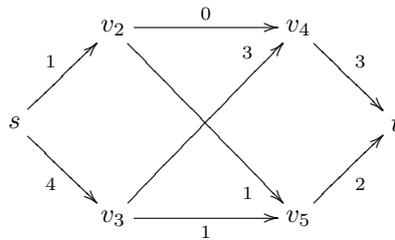
Temos que $\varepsilon = \min\{5 - 1, 3, 3 - 1\} = 2$, logo ficamos com o novo fluxo:



$S(x) = \{s, v_2, v_3, v_4, v_5, t\}$, tomamos novamente um caminho x -aumentador, por exemplo:



Temos que $\varepsilon = \min\{5 - 3, 3 - 2, 1, 1, 3 - 1\} = 1$, logo ficamos com o novo fluxo:



$S(x) = \{s, v_2\} \not\cong t$, logo este fluxo- st é máximo e $C = \delta^+(S(x)) = \{[s, v_2], [v_3, v_4], [v_3, v_5]\}$ é corte- st mínimo. Confirma-se, de facto, que $f(x) = 1 + 4 = d(C) = 1 + 3 + 1$.

Fluxo de rede de custo mínimo.

Definição 5.75. Chama-se *problema do fluxo de rede de custo mínimo* ao problema de determinar de entre os fluxos viáveis, $x \in \{x \in \mathbb{R}^E : Ax = b, 0 \leq x \leq d\}$, de uma dada rede $\mathcal{N} = (V, E, d, b)$ um que tenha $c^T x = \sum_{e \in E} c_e x_e$ mínimo.

Observação 5.76. Trata-se de um caso particular de programação linear:

$$\min\{c^T x : Ax = b, 0 \leq x \leq d\}$$

onde A é a matriz de incidência do digrafo (V, E) , b e d são os vectores de fornecimento e capacidade da rede respectivamente.

Exemplo 5.77. Um problema típico de fluxo de rede de custo mínimo é o de calcular o custo mínimo do transporte de uma determinada mercadoria desde das suas fontes de origem (por exemplo fábricas) até aos locais de consumo.

Exemplo 5.78. O problema de fluxo- st máximo num digrafo $D = (V, E)$ pode ser resolvido como um problema de fluxo de rede de custo mínimo no digrafo aumentado $\tilde{D} = (V, E \cup \{(t, s)\})$. Basta tomar para vector de custo $c = -\chi^{\{(t,s)\}}$, ou seja, $c_{(t,s)} = -1$ e $c_e = 0$ para qualquer $e \in E$.

Vamos considerar para o que se segue que o digrafo da rede $D = (V, E)$ é conexo (quando visto como um grafo não-orientado).

Proposição 5.79. *Seja A a matriz de incidência do digrafo $D = (V, E)$ então a característica de A é $n - 1$ (onde $n = |V|$). Além disso, se \tilde{A} é a matriz obtida a partir de A eliminando uma linha (não interessa qual) então as linhas de \tilde{A} são linearmente independentes.*

Dem. Como a soma das linhas de A dá o vector nulo (pelo modo como a matriz de incidência para digrafos é definida) é óbvio que as linhas de A são linearmente dependentes logo $\text{car}(A) \leq n - 1$.

Tomamos a matriz \tilde{A} obtida da matriz A pela remoção de uma linha (vamos designar por i_0 o vértice correspondente à linha removida) e tomamos uma árvore geradora¹¹ $T = (V, F)$ de D . Em seguida vamos reordenar os restantes vértices, i_1, \dots, i_{n-1} , e as aresta, e_1, \dots, e_m , de modo que:

- (i) $F = \{e_1, \dots, e_{n-1}\}$ (i.e. as primeiras arestas são as que formam a árvore geradora T);
- (ii) $i_k \neq i_0$ é uma folha (i.e. tem grau 1) na árvore obtida de T removendo os vértices i_1, \dots, i_{k-1} , ou seja, $T[V \setminus \{i_1, \dots, i_{k-1}\}]$;
- (iii) e_k é a única aresta incidente a i_k em $T[V \setminus \{i_1, \dots, i_{k-1}\}]$.

Temos que a matriz \tilde{A} fica da forma

$$\begin{bmatrix} \pm 1 & 0 & 0 & \cdots & 0 & * & \cdots & * \\ * & \pm 1 & 0 & \cdots & 0 & * & \cdots & * \\ * & * & \pm 1 & \cdots & 0 & * & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots \\ * & * & * & \cdots & \pm 1 & * & \cdots & * \end{bmatrix}$$

Portanto temos que \tilde{A} tem característica $n - 1$ e como tal A também tem característica $n - 1$. \square

A demonstração anterior mostra-nos que a uma árvore geradora corresponde uma base de \tilde{A} . O próximo resultado diz-nos que essa correspondência é biunívoca.

Proposição 5.80. *Escolhe-se um vértice $i_0 \in V$ (chamado vértice de raiz), e seja \tilde{A} a matriz obtida a partir de A após a remoção da linha correspondente a i_0 . Então existe uma correspondência biunívoca entre as árvores geradoras de D e as bases em \tilde{A} (i.e. submatrizes não-singulares de dimensão $(n - 1) \times (n - 1)$).*

Dem. Seja $F \subseteq E$ um conjunto de aresta de D e \tilde{A}_F a submatriz de \tilde{A} que formadas pelas arestas em F . Vimos na demonstração da proposição anterior que se (V, F) é uma árvore então \tilde{A}_F é não-singular (i.e. é uma base de \tilde{A}). Vejamos agora que se \tilde{A}_F é não-singular então (V, F) é acíclico (logo é uma árvore pois para \tilde{A}_F ser quadrada temos que ter $|F| = n - 1$). Suponhamos que (V, F) tinha um ciclo e_1, e_2, \dots, e_k . Escolhida uma orientação para o ciclo definimos

$$\lambda_i = \begin{cases} 1 & \text{se as orientações do ciclo e do arco } e_i \text{ coincidem} \\ -1 & \text{se as orientações do ciclo e do arco } e_i \text{ são opostas} \end{cases}$$

Temos que se j_1, \dots, j_k forem as colunas de A associadas aos arcos e_1, \dots, e_k então $\sum_{i=1}^k \lambda_i j_i = 0$ logo \tilde{A}_F não seria não-singular. \square

Cosideremos agora o problema de fluxo de rede de custo mínimo. Para simplificar vamos restringir-nos ao caso em que as capacidades são ilimitadas.

$$\min\{c^T x : Ax = b, x \geq 0\}$$

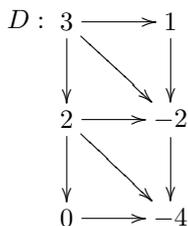
Embora este problema possa ser resolvido usando algoritmo simplex, o facto de A ser uma matriz de incidência de um digrafo permite fazer uso das proposições anteriores para obter um algoritmo mais simples.

Primeiro escolhemos um vértice raiz i_0 e eliminamos a correspondente linha à matriz A . Ficamos com um problema equivalente:

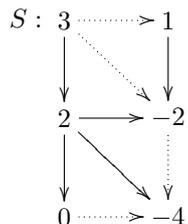
$$\min\{c^T x : \tilde{A}x = \tilde{b}, x \geq 0\}$$

Como vimos as bases de \tilde{A} correspondem às árvores geradoras de $D = (V, E)$. Além disso, dada uma árvore geradora S , a solução básica x^S associada à base \tilde{A}_S pode ser calculada folha a folha. Por exemplo, se no seguinte digrafo

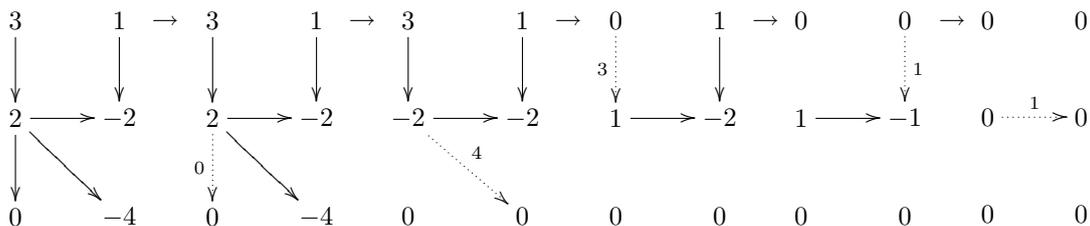
¹¹Existe pois estamos a assumir que D é conexo.



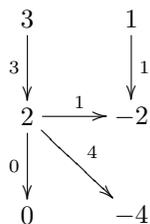
com os fornecimentos indicados no lugar dos vértices, tomarmos a árvore geradora



podemos calcular a fluxo associado do seguinte modo. Procuremos um vértice v de grau 1, se o seu único arco incidente for de entrada então atribuímos a esse arco o fluxo $-b_v$ (i.e. o simétrico do fornecimento do vértice v) e se o arco for de saída então atribuímos o fluxo b_v . Em seguida removemos esse arco e alteramos os fornecimentos nos seus vértices inicial (subtraindo pelo valor do fluxo do arco) e final (somando pelo valor do fluxo do arco).



No final ficamos com



Note-se que a solução básica x^S obtida pode ser ou não viável. A solução básica dual $y^S = c_S^T \tilde{A}_S^{-1}$ é solução do seguinte sistema:

$$\begin{cases} y_i - y_j & \text{se } (i, j) \in S \\ y_{i_0} & = 0 \end{cases}$$

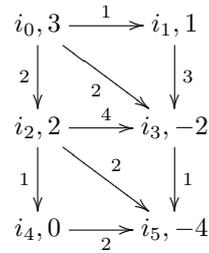
Segue-se então o algoritmo simplex para redes.

Algoritmo Simplex para Redes:

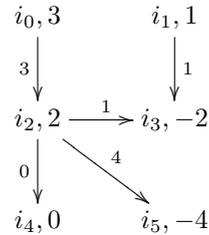
- Passo 0 Seja x^S uma solução básica viável correspondente a uma árvore geradora S .
- Passo 1 Calcula-se as variáveis duais $y_i^S, i \in V$ resolvendo as equações $y_i^S - y_j^S = c_{ij}$ para cada $(i, j) \in S$ (usando $y_{i_0} = 0$). O custo reduzido para as variáveis não-básicas $x_{ij}, (i, j) \notin S$ é então $\bar{c}_{ij} = c_{ij} - (y_i^S - y_j^S)$.
- Passo 2 Se $\bar{c}_{ij} \geq 0, \forall (i, j) \notin S$ então termina-se, a solução corrente é optimal. Senão faz-se o passo 3.

Passo 3 Escolhe-se um arco não-básico $(i, j) \notin S$ com $\bar{c}_{ij} < 0$. Seja C o único ciclo de $S \cup \{(i, j)\}$. Reajusta-se o fluxo aumentando o fluxo no ciclo C tal que x_{ij} seja o máximo possível, até um outro arco $e \in C \setminus \{(i, j)\}$ ficar com fluxo zero ($x_e = 0$). Então muda-se a base (árvore gerado) S para $(S \cup \{(i, j)\}) \setminus \{e\}$ e volta-se ao passo 1.

Exemplo 5.81. Consideremos a seguinte rede, com os fornecimentos indicados ao lado dos vértices e os custos de cada aresta indicados sobre estas:



Tomemos a solução básica associada à árvore geradora S :



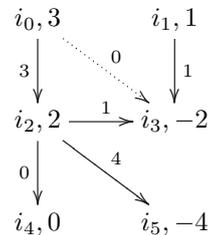
Calculando as variáveis duais $y_i^S, i \in V$, através das equações $y_i^S - y_j^S = c_{ij}, (i, j) \in S, y_{i_0} = 0$, obtemos¹²:

$$(y_{i_0}, y_{i_1}, y_{i_2}, y_{i_3}, y_{i_4}, y_{i_5}) = (0, -3, -2, -6, -3, -4)$$

Os custos reduzidos $\bar{c}_{ij} = c_{ij} - (y_i^S - y_j^S)$ das arestas não-básicas, $(i, j) \notin S$, são:

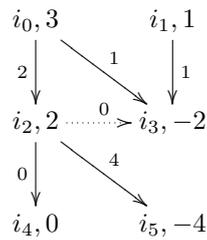
$$\bar{c}_{01} = -2 < 0, \bar{c}_{03} = -4 < 0, \bar{c}_{35} = 3 \geq 0, \bar{c}_{45} = 1 \geq 0$$

Escolhemos então uma variável não-básica x_{ij} com $\bar{c}_{ij} < 0$, por exemplo x_{03} . Aumentemos o fluxo no único ciclo do digrafo $S \cup (i_0, i_3)$

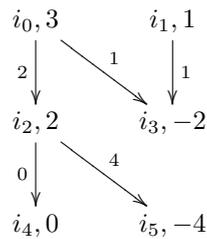


de modo a que o fluxo em (i_0, i_3) seja tão grande quanto possível (o que acontece quando o fluxo se anula num outro arco do ciclo)

¹²Um modo fácil de calcular y_i é tomar o único caminho- i, i_0 e ir somando ou subtraindo os custos dos arcos conforme estes sejam de avanço ou de recuo.



Remove-se então esse arco¹³ obtendo uma nova solução básica:



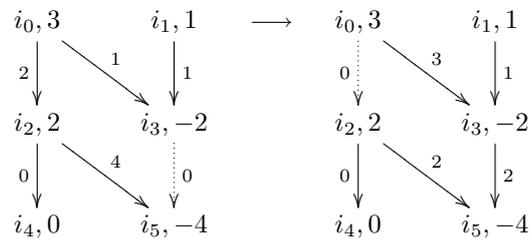
As variáveis duais são

$$(y_{i_0}, y_{i_1}, y_{i_2}, y_{i_3}, y_{i_4}, y_{i_5}) = (0, 1, -2, -2, -3, -4)$$

e os custos reduzidos são

$$\bar{c}_{01} = 2 \geq 0, \bar{c}_{23} = 4 \geq 0, \bar{c}_{35} = -1 < 0, \bar{c}_{45} = 1 \geq 0$$

Repete-se o processo, agora com a variável x_{35} ,



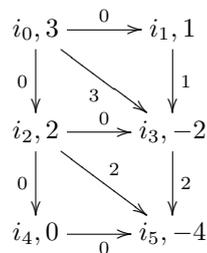
Temos agora que as variáveis duais são

$$(y_{i_0}, y_{i_1}, y_{i_2}, y_{i_3}, y_{i_4}, y_{i_5}) = (0, 1, -1, -2, -2, -3)$$

e os custos reduzidos são

$$\bar{c}_{01} = 2 \geq 0, \bar{c}_{02} = 1 \geq 0, \bar{c}_{23} = 2 \geq 0, \bar{c}_{45} = 1 \geq 0$$

Temos verificada a condição optimal, o algoritmo termina com a solução de custo mínimo:



¹³Na verdade pode haver mais de um arco a ficar com fluxo zero, em todo o caso remove-se apenas um que esteja nessas condições.

Matchings, cocliques e coberturas.

Definição 5.82. Seja $G = (V, E)$ um grafo. Um *coclique* ou *conjunto independente de vértices* é um subconjunto C de V tal que não existe nenhuma aresta $e \in E$ a ligar dois vértices de C (i.e. $e = [v, w] \in E \Rightarrow v \notin C$ ou $w \notin C$).

Uma *cobertura de vértices* é um subconjunto W de V tal que qualquer aresta $e \in E$ é incidente a pelo menos um vértice de W (i.e. $e = [v, w] \in E \Rightarrow v \in W$ ou $w \in W$).

Observação 5.83. $C \subseteq V$ é uma coclique se e só se $V \setminus C$ é uma cobertura de vértices.

Observação 5.84. Se $C \subseteq V$ é uma coclique e $B \subseteq C$ então B é uma coclique.

Se $W \subseteq V$ é uma cobertura de vértices e $W \subseteq X \subseteq V$ então X é uma cobertura de vértices.

Definição 5.85. Um *matching* ou *emparelhamento* ou *conjunto independente de arestas* é um subconjunto M de E tal que não existe duas arestas $e, e' \in M$ que sejam adjacentes (i.e. incidentes ao mesmo vértice).

Um matching diz-se *perfeito* se cobre todos os vértices (i.e. tem tamanho $\frac{1}{2}|V|$). Uma *cobertura de arestas* é um subconjunto F de E tal que qualquer vértice v é incidente a uma aresta em F (i.e. $\forall v \in V \exists e \in F : v \in e$).

Observação 5.86. Uma cobertura de arestas só existe se G não contém vértices isolados.

Observação 5.87. Se $M \subseteq E$ é um matching e $N \subseteq M$ então N é um matching.

Se $F \subseteq E$ é uma cobertura de arestas e $F \subseteq F' \subseteq E$ então F' é uma cobertura de vértices.

Definição 5.88. Para um grafo $G = (V, E)$ define-se os seguintes valores:

- número de coclique de G : $\alpha(G) = \max\{|C| : C \text{ é uma coclique em } G\}$;
- número de cobertura de arestas de G : $\rho(G) = \min\{|F| : F \text{ é uma cobertura de arestas em } G\}$;
- número de cobertura de vértices de G : $\tau(G) = \min\{|W| : W \text{ é uma cobertura de vértices em } G\}$;
- número de matching de G : $\nu(G) = \max\{|M| : M \text{ é um matching em } G\}$.

Observação 5.89. $\alpha(G) \leq \rho(G)$ e $\nu(G) \leq \tau(G)$.

O próximo resultado é conhecido por **teorema de Gallai** e relaciona este vários números.

Teorema 5.90. Para qualquer grafo $G = (V, E)$ sem vértices isolados temos:

$$\alpha(G) + \tau(G) = |V| = \nu(G) + \rho(G)$$

Dem. A primeira igualdade resulta da observação 5.83. De facto, C é uma coclique de tamanho máximo se e só se $V \setminus C$ é uma cobertura de arestas de tamanho mínimo. Logo $\alpha(G) + \tau(G) = |C| + |V \setminus C| = |V|$.

Para a segunda igualdade, seja M um matching de tamanho máximo e seja $X := \{v \in V : v \notin e \forall e \in M\}$ o conjunto dos vértices que não são cobertos pelo matching ($|X| = |V| - 2|M|$). Como G não tem vértices isolados, para cada $v \in X$ pode-se escolher uma aresta e_v incidente a v . Por definição de X , $e_v \notin M$, e além disso, como M é maximal, para vértices distintos $v, w \in X$ temos arestas e_v e e_w distintas. Logo $|\{e_v : v \in X\}| = |X|$. Como $M \cup \{e_v : v \in X\}$ forma uma cobertura de arestas de G com tamanho $|V| - |M|$ temos que $\nu(G) + \rho(G) \leq |V|$.

Por outro lado, seja F uma cobertura de aresta de tamanho mínimo. Vamos produzir um matching do seguinte modo. Damos uma ordenação aos vértices v_1, \dots, v_n , e tomando $F_0 = F$ construímos uma sucessão de conjuntos de arestas do seguinte modo. Se $d_{F_k}(v_{k+1}) \geq 1$ então F_{k+1} é obtido de F_k removendo $d_{F_k}(v_{k+1}) - 1$ arestas incidentes a v_{k+1} (se $d_{F_k}(v_{k+1}) = 0$ então $F_{k+1} = F_k$). Temos então que $M = F_n$ é um matching que se obteve de F removendo no máximo $d_F(v_k)$ em cada passo¹⁴. Assim, o seu tamanho é pelo menos

$$|M| \geq |F| - \sum_{v \in V} (d_F(v) - 1) = |F| - 2|F| + |V| = |V| - |F|$$

Portanto $\nu(G) \geq |V| - |F| = |V| - \rho(G)$, logo $\nu(G) + \rho(G) \geq |V|$. □

¹⁴Note-se que $d_{F_k}(v_{k+1}) \leq d_F(v_{k+1})$ para qualquer k , e que se $d_{F_k}(v_{k+1}) = 0$ então $d_{F_k}(v_{k+1}) \leq d_F(v_{k+1}) - 1$ pois F é uma cobertura de arestas (logo $d_F(v) \geq 1$ para qualquer $v \in V$).

Definição 5.91. Seja M um matching num dado grafo $G = (V, E)$. Um caminho $P : v_0, e_0, v_1, \dots, e_{t-1}, v_t$ em G diz-se M -aumentador se:

- (i) t é ímpar;
- (ii) $e_1, e_3, \dots, e_{t-2} \in M$;
- (iii) $v_0, v_t \notin \bigcup M$ (i.e. v_0 e v_t não são incidentes a nenhuma aresta pertencente a M).

Note-se que a condição (ii) implica que $e_0, e_2, \dots, e_{t-1} \notin M$. Além disso, $M' = (M \setminus \{e_1, e_3, \dots, e_{t-2}\}) \cup \{e_0, e_2, \dots, e_{t-1}\}$ forma um novo matching com $|M'| = |M| + 1$. Isto demonstra parcialmente o seguinte resultado.

Teorema 5.92. *Seja $G = (V, E)$ um grafo e seja M um matching em G . Então M é um matching de cardinalidade máxima (i.e. $|M| = \nu(G)$) se e só se não existem caminhos M -aumentadores.*

Dem. Já vimos atrás que se existe um caminho M -aumentador $P : v_0, e_0, v_1, \dots, e_{t-1}, v_t$ então $M' = (M \setminus \{e_1, e_3, \dots, e_{t-2}\}) \cup \{e_0, e_2, \dots, e_{t-1}\}$ forma um novo matching com $|M'| = |M| + 1$. Logo, se M for um matching de cardinalidade máxima não pode haver caminhos M -aumentadores.

Por outro lado, se M não é de cardinalidade máxima então existe pelo menos um matching M' com $|M'| > |M|$. Tomemos o grafo $G' = G[M \cup M']$ (o subgrafo de G induzido pelo subconjunto de arestas pertencentes a M ou a M'), cada vértice de G' tem grau menor ou igual a dois logo as componentes conexas de G' são caminhos ou ciclos. Como $|M'| > |M|$ pelo menos uma dessas componentes tem mais arestas pertencentes a M' que a M , logo essa componente forma um caminho M -aumentador. \square

Este resultado dá uma maneira de construir um matching de tamanho máximo desde que tenhamos um modo de decidir se existem caminhos M -aumentadores e de como os encontrar. Uma classe de grafos onde isso pode ser feito é a seguinte.

Definição 5.93. Um grafo bipartido é um grafo $G = (V, E)$ tal que V pode ser decomposto em duas cocliques usualmente chamadas de *classes de cor*. Dito de outra forma, $V = V_1 \cup V_2$, $V_1 \cap V_2 = \emptyset$ e $[v, u] \in E \Rightarrow (v \in V_1 \text{ e } u \in V_2) \text{ ou } (v \in V_2 \text{ e } u \in V_1)$.

Definição 5.94. Um grafo bipartido $G = (V, E)$, $V = V_1 \cup V_2$ diz-se *completo* se existem arestas a ligar todos os vértices de uma classe de cor V_1 à outra classe de cor V_2 (i.e. $E = \{[v_1, v_2] : v_1 \in V_1 \text{ e } v_2 \in V_2\}$), e denota-se por K_{n_1, n_2} onde $n_1 = |V_1|$ e $n_2 = |V_2|$.

Existem outras definições alternativas para grafo bipartido que são exemplificadas na seguinte proposição.

Proposição 5.95. *As seguintes afirmações são equivalentes:*

- (i) G é um grafo bipartido;
- (ii) G não tem ciclos de comprimento ímpar;
- (iii) Existe um homomorfismo de G em $K_{1,1} = (\{v_1, v_2\}, \{[v_1, v_2]\})$.

Dem. Exercício. \square

O próximo resultado é conhecido por **teorema de matching de König**.

Teorema 5.96. *Para um grafo bipartido $G = (V, E)$ temos*

$$\nu(G) = \tau(G)$$

Ou seja, a cardinalidade máxima de um matching é igual à cardinalidade mínima de uma cobertura de vértices.

Dem. Pela observação 5.89, basta mostrar que $\nu(G) \geq \tau(G)$. Podemos assumir que $|E| \geq 1$ (se $|E| = 0$ então $\nu(G) = \tau(G) = 0$).

Vejam que

G tem um vértice que é coberto por qualquer matching de tamanho máximo.

Suponhamos por absurdo que tal não acontece. Então, dada $e = [u, v]$ uma aresta de G , existiriam dois matchings de tamanho máximo M e N que não cobririam u e v respectivamente. Seja P a componente conexa de u em $(V, M \cup N)$. Como M e N são matchings P só pode ser à partida ou um caminho (que até pode ter comprimento zero) ou um ciclo. No entanto, como M não cobre u , P só pode ser um caminho com vértice final u . Além disso, P tem comprimento par, pois senão seria um caminho M -aumentador. Como v não é coberto por N , teríamos que $P \cup e$ seria um caminho N -aumentador contrariando o facto de N ser de tamanho máximo.

Vamos agora concluir a demonstração do teorema. Tomamos um vértice u que seja coberto por qualquer matching de tamanho máximo. Então para o grafo $G' := G - u$ temos que $\nu(G') = \nu(G) - 1$. Por indução, temos que $\tau(G') \leq \nu(G')$, logo existe uma cobertura de vértices C de G' de tamanho não superior a $\nu(G')$. Então $C \cup \{u\}$ é uma cobertura de vértices com tamanho não superior a $\nu(G') + 1 = \nu(G)$. \square

Sai imediatamente deste resultado juntamente com o teorema de Gallai o seguinte resultado, conhecido por **teorema de cobertura de arestas de Kőnig**.

Corolário 5.97. *Para um grafo bipartido $G = (V, E)$ sem vértices isolados temos*

$$\alpha(G) = \rho(G)$$

Ou seja, a cardinalidade máxima de uma coclique é igual à cardinalidade mínima de uma cobertura de arestas.

Algoritmo de aumento de matching para grafos bipartidos:

- input Seja $G = (V, E)$ um grafo bipartido e M um matching (se não se tiver nenhum candidato tome-se $M = \emptyset$).
- passo 1 Seja U e W as classes de cores de G . Orienta-se cada aresta $e = [u, w]$ de G (com $u \in U$ e $w \in W$) da seguinte maneira:
 se $e \in M$ então orienta-se e de w para u , $e = (w, u)$;
 se $e \notin M$ então orienta-se e de u para w , $e = (u, w)$.
- passo 2 Seja D o digrafo obtido no passo 1. Considere-se os conjuntos $U' := U \setminus \bigcup M = \{v \in U : v \notin e \ \forall e \in M\}$ e $W' := W \setminus \bigcup M = \{v \in W : v \notin e \ \forall e \in M\}$. Temos que um caminho M -aumentador (se existir) pode ser encontrado achando um caminho dirigido em D de um vértice pertencente a U' para um vértice pertencente a W' . Se existir faz-se o passo 3, senão termina-se, M é um matching de tamanho máximo.
- passo 3 Toma-se um caminho M -aumentador P e redefine-se $M := M \Delta EP$ (onde Δ é diferença simétrica de conjuntos¹⁵ EP é o conjunto das arestas em P), volta-se ao passo 1.

Podemos usar o algoritmo de aumento de matching para grafos bipartidos para resolver problemas de atribuição de tarefas.

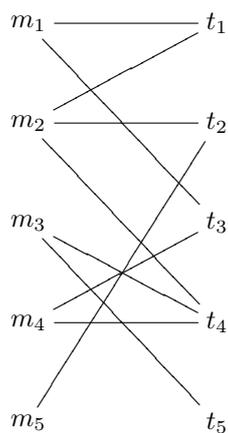
Suponhamos que temos a cumprir n tarefas t_1, \dots, t_n e para tal temos ao dispor k máquinas m_1, \dots, m_k . Cada máquina está apta a efectuar algumas tarefas mas só pode fazer uma por dia. O problema de encontrar uma atribuição de tarefas num dia em que sejam cumpridas o máximo número de tarefas é equivalente a determinar um matching de tamanho máximo no grafo bipartido $G = (V, E)$ onde $V = \{t_1, \dots, t_n\} \cup \{m_1, \dots, m_k\}$, sendo $\{t_1, \dots, t_n\}$ e $\{m_1, \dots, m_k\}$ as duas classes de cores, e $[m_i, t_j] \in E$ se a máquina m_i pode cumprir a tarefa t_j . Tomamos o seguinte exemplo

Exemplo 5.98. Cinco máquinas estão disponíveis para cumprir cinco tarefas de acordo com o seguinte quadro de compatibilidades:

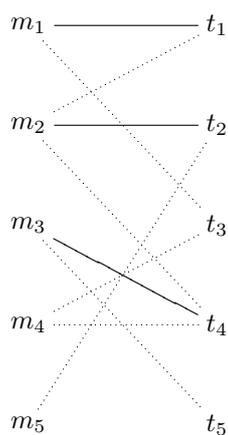
¹⁵ $A \Delta B := (A \cup B) \setminus (A \cap B)$.

	t_1	t_2	t_3	t_4	t_5
m_1	X		X		
m_2	X	X			X
m_3				X	X
m_4			X	X	
m_5		X			

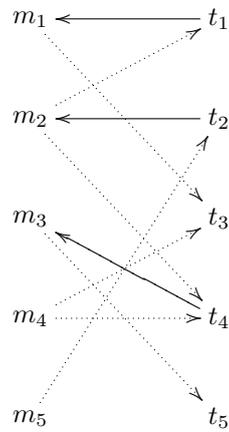
Temos então de determinar um matching de tamanho máximo para o seguinte grafo:



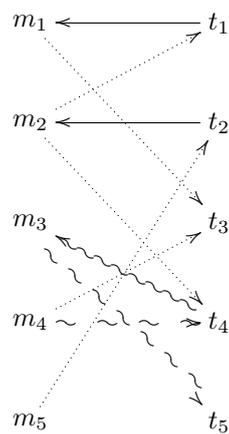
Começando com o matching $M_0 = \{[m_1, t_1], [m_2, t_2], [m_3, t_4]\}$:



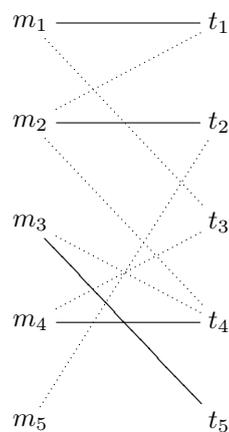
obtemos no passo 1 o digrafo:



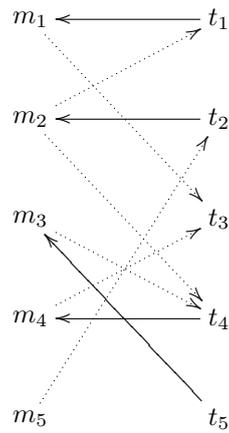
temos $U' = \{m_4, m_5\}$ e $W' = \{t_3, t_5\}$, tomando o caminho dirigido $P : m_4, t_4, m_3, t_5$:



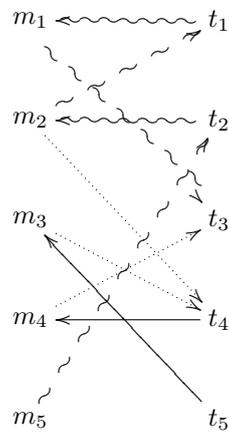
obtemos após o passo 3 o novo matching $M_1 = \{[m_1, t_1], [m_2, t_2], [m_4, t_4], [m_3, t_5]\}$:



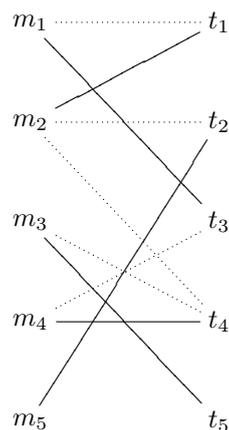
Continuando o algoritmo:



temos $U' = \{m_5\}$ e $W' = \{t_3\}$, tomando o caminho dirigido $P : m_5, t_2, m_2, t_1, m_1, t_3$:



obtemos o matching $M_2 = \{[m_1, t_1], [m_2, t_2], [m_4, t_4], [m_3, t_5]\}$:



que é de tamanho máximo pois não tem caminhos M -aumentados.

Consideremos agora que no exemplo anterior cada máquina m_i efectua a tarefa t_j com uma dada eficiência w_{ij} . Ou seja, temos uma função "peso" $w : E \rightarrow \mathbb{R}$ nas arestas. Deste modo, para cada subconjunto M de E temos um peso associado

$$w(M) := \sum_{e \in M} w(e)$$

O novo problema agora é, de entre os matchings de G , encontrar um que maximize (ou minimize) o peso.

Definição 5.99. Dado um grafo $G = (V, E)$ e uma função peso $w : E \rightarrow \mathbb{R}$, dizemos que um matching M é *extremo* se tem peso máximo sobre os matchings de igual cardinalidade $|M|$.

Temos portanto que um matching de peso máximo é um matching extremo, embora o contrário nem sempre seja válido.

Seja M um matching. Define-se uma função "comprimento" $l : E \rightarrow \mathbb{R}$ do seguinte modo:

$$\begin{aligned} l(e) &:= w(e) \text{ se } e \in M, \\ l(e) &:= -w(e) \text{ se } e \notin M. \end{aligned}$$

Proposição 5.100. *Seja P um caminho M -aumentador de comprimento mínimo. Se M é um matching extremo, então $M' := M \Delta EP$ é também um matching extremo.*

Dem. Queremos ver que $M' := M \Delta EP$ é um matching extremo. Ou seja, que para qualquer matching N com $|N| = |M'| = |M| + 1$, $w(N) \leq w(M')$. Dado N um matching com $|N| = |M| + 1$, $M \cup N$ tem uma componente Q que é um caminho M -aumentador. Como P é um caminho M -aumentador de comprimento mínimo, temos $l(Q) \geq l(P)$. Além disso, como $N \Delta EQ$ é um matching de tamanho $|M|$ e M é extremo, temos $w(N \Delta EQ) \leq w(M)$.

Portanto temos

$$w(N) = w(N \Delta EQ) - l(Q) \leq w(M) - l(P) = w(M').$$

□

Podemos então adaptar o algoritmo de aumento de matching para grafos bipartidos (AAMGB) de modo a permitir determinar um matching de peso máximo. Basta que no passo 3 escolhamos um caminho M -aumentador de comprimento mínimo, que pode ser determinado usando o algoritmo de Dijkstra.

Observação 5.101. O(s) matching(s) com peso mínimo pode(m) não ser de cardinalidade máxima. No entanto o algoritmo produz uma sequência finita de matchings extremos $M_0, M_1, \dots, M_{\nu(G)}$ com $|M_k| = k$, se tomarmos o M_k com maior peso temos um matching de peso máximo.

Temos de ter algum cuidado a aplicar o algoritmo de Dijkstra para determinar o caminho M -aumentador de comprimento mínimo. Em primeiro lugar, queremos determinar o caminho mais curto de entre todos os caminhos M -aumentadores, o que corresponde a aplicar o algoritmo de Dijkstra para digrafos a todos os elementos de U' do passo 2 do AAMGB. Em segundo lugar, temos arestas com comprimentos negativos, pelo que uma hipotética existência de ciclos dirigidos com comprimento negativo faz com que não exista caminho dirigido mais curto.

No entanto, tal não acontece como garante o seguinte teorema.

Teorema 5.102. *Seja M um matching extremo num grafo bipartido $G = (V, E)$. Então o digrafo D (obtido a partir de G de acordo com o passo 1 do AAMGB) não possui ciclos dirigidos de comprimento negativo.*

Dem. Suponhamos que existe um ciclo dirigido C em D com comprimento negativo, $l(C) < 0$. Sendo G bipartido, a sequência de vértices que define C é da forma $u_0, w_1, u_1, \dots, w_t, u_t$ com $u_0 = u_t$ e $u_1, \dots, u_t \in U$ e $w_1, \dots, w_t \in W$. Do modo como as arestas de D foram orientadas temos que $[u_0, w_1], \dots, [u_{t-1}, w_t]$ não pertencem a M e $[w_1, u_1], \dots, [w_t, u_t]$ pertencem a M . Assim $M'' := M \Delta EC$ é um matching de cardinalidade igual à de M e de peso $w(M'') = w(M) - l(C) > w(M)$ o que contradiz o facto de M ser extremo. □

Cliques, cocliques e colorações de vértices.

Definição 5.103. Uma *clique* num grafo $G = (V, E)$ é um subconjunto C de V tal que quaisquer dois elementos de C são vértices adjacentes. O *número de clique* de G , denotado por $\omega(G)$, é a cardinalidade máxima das cliques de G .

Observação 5.104. O conjunto C é um clique do grafo G se e só se C é uma coclique do grafo complementar \bar{G} . Logo $\omega(G) = \alpha(\bar{G})$.

Observação 5.105. Se $C \subseteq V$ é uma clique e $B \subseteq C$ então B é uma clique.

Definição 5.106. Uma k -coloração (de vértices) de um grafo $G = (V, E)$ é uma decomposição de V em k cocliques C_1, \dots, C_k . O número cromático (de vértices) de G , denotado¹⁶ por $\chi(G)$, é menor k para o qual existe uma k -coloração de G . Um grafo G diz-se k -colorável (nos vértices) se existe uma k -coloração, ou seja $\chi(G) \leq k$.

Em geral é difícil determinar se um grafo é k -colorável. Seguem-se alguns resultados.

Proposição 5.107. Para qualquer grafo de G temos $\chi(G) \geq \omega(G)$.

Dem. Trivial □

Proposição 5.108. Se H é um subgrafo de G então $\chi(H) \leq \chi(G)$.

Dem. Exercício. □

Proposição 5.109. Se $H \rightarrow G$ é um homomorfismo de grafos então $\chi(H) \leq \chi(G)$.

Dem. Exercício. □

Proposição 5.110. Se $\Delta(G)$ o maior grau dos vértices de $G = (V, E)$ (i.e. $\Delta(G) = \max\{d(v) : v \in V\}$), então

$$\chi(G) \leq \Delta(G) + 1$$

Dem. Exercício. □

Este resultado pode ser melhorado de duas formas. A primeira é apresentada no seguinte resultado, conhecido por **teorema de Brooks**.

Teorema 5.111. Seja G um grafo conexo, então $\chi(G) \leq \Delta(G)$ a menos que G seja um grafo completo ou $\Delta(G) = 2$ e G seja um circuito de comprimento ímpar.

Dem. O teorema é claramente válido para $\Delta = 0, 1, 2$. Suponhamos que existem grafos tais que $\Delta \geq 3$ e $\chi \geq \Delta + 1$. Seja G um grafo nestas condições com ordem mínima. Seja v_0 um vértice de G e G' o grafo resultante da remoção de v_0 . Então G' é $\Delta(G')$ -colorável (em particular é $\Delta(G')$ -colorável). Isto implica que $d(v_0) = \Delta(G)$, pois se não G era $\Delta(G)$ -colorável.

Do facto de G ter ordem mínima resulta o seguinte:

- 1 - Em qualquer Δ -coloração de G' os vértices adjacentes a v_0 são de cores diferentes.

Caso contrário G seria Δ -colorável.

Seja $u_1, u_2, \dots, u_\Delta$ os vértices adjacentes a v_0 . Seja $1, 2, \dots, \Delta$ as suas cores respectivas em G' . Seja $G(i, j)$ o subgrafo induzido pelos vértices de cores i e j . Temos então o seguinte:

- 2 - Os vértices u_i e u_j são da mesma componente conexa de $G(i, j)$.

Caso contrário, por trocas das cores i e j na componente conexa de u_i ficamos com os dois vértices u_i e u_j com a mesma cor, o que contradiz a primeira afirmação.

Seja C_{ij} a componente conexa de $G(i, j)$ que contém u_i e u_j .

- 3 - C_{ij} é um caminho de u_i a u_j .

Supõe-se que o grau de u_i em C_{ij} é maior que 1. Então u_i é adjacente a pelo menos dois vértices de cor j . Como em G' , $d(u_i) \leq \Delta - 1$, podemos recolorir u_i com uma nova cor k diferente de i , o que contradiz a primeira afirmação.

Portanto $d(u_i) = 1$ e similarmente $d(u_j) = 1$.

O grau em C_{ij} dos outros vértices é 2. Caso contrário existiria um vértice u de grau superior a 2 em C_{ij} , e podemos supor que u é o primeiro vértice no caminho com tal propriedade. Então u é adjacente a pelo menos três vértices da mesma cor (se u é da cor i os outros são da cor j). Como

¹⁶No texto "A Course in Combinatorial Optimization" de A. Schrijver o número cromático é denotado por $\gamma(G)$.

$d(u) \leq \Delta$, u pode ser colorido com outra cor, o que faz com que u_i e u_j não pertençam à mesma componente conexa de C_{ij} contradizendo a segunda afirmação.

Portanto C_{ij} é um caminho de u_i a u_j .

4 - Para qualquer i, j e k , C_{ij} e C_{ik} não têm vértices comuns excepto u_i .

Seja $u \neq u_i$ comum a C_{ij} e C_{ik} . Então u é colorido pela cor i , e é adjacente a dois vértices de cor j e a dois de cor k . Como $d(u) \leq \Delta$, u pode ser recolorido com uma cor l distinta de i, j e k . Isto faz desconectar u_i e u_j , contradizendo a segunda afirmação.

Vamos agora contradizer esta última afirmação.

Se G não é um grafo completo então existem dois vértices u_1 e u_2 que não são adjacentes. Então o caminho C_{12} contém um vértice u diferente de u_2 e adjacente a u_1 . Supondo que trocávamos as cores 1 e 3 no caminho C_{13} (que existe pois $\Delta(G) \geq 3$), então temos uma nova coloração de G' , u_1 fica com cor 3 e u_3 com cor 1. Neste caso as novas componentes conexas C'_{12} e C'_{23} contêm o vértice comum u , contradizendo a quarta afirmação.

Com isto terminamos a desmonstração. □

Para a maior parte dos grafos *regulares* (i.e. grafos em que todos os vértices têm o mesmo grau), este resultado dá-nos um majorante do número cromático bem próximo deste. No entanto, para grafos em que os graus dos vértices variam bastante é mais útil o seguinte resultado.

Proposição 5.112. *Seja $\delta(G) = \min\{d(v) : v \in V\}$ e $sw(G) = \max\{\delta(G) : H \text{ é subgrafo de } G\}$, temos que*

$$\chi(G) \leq sw(G) + 1.$$

Dem. Exercício. □

Uma classe de grafos onde a proposição anterior é, em geral, mais útil que o teorema de Brooks é a classe dos grafos mergulhados numa dada superfície M .

Definição 5.113. Um grafo $G = (V, E)$ diz-se *mergulhável* numa superfície M se pode ser "desenhado" na superfície M . Dito de uma forma mais rigorosa, existem $n = |V|$ pontos distintos v_1, \dots, v_n da superfície M (que representam os vértices de G) e para cada aresta $e = [v_i, v_j]$ existe um caminho $\gamma_e : [0, 1] \rightarrow M$ (que representa a aresta e) tal que $\gamma_e(0) = v_i$ e $\gamma_e(1) = v_j$. Além disso, temos que o interior de qualquer caminho, $\gamma_e]0, 1[$, representando uma aresta e não intersecta nem o conjunto dos vértices nem os outros caminhos que representam as arestas.

Se um grafo for mergulhável no plano \mathbb{R}^2 diz-se *planar*.

Observação 5.114. Por projecção estereográfica um grafo é planar se e só se é mergulhável na esfera.

Uma condição necessária para um grafo conexo ser planar, deduzida pela fórmula de Euler, é a seguinte.

Proposição 5.115. *Seja G um grafo planar conexo com $n > 2$ vértices e m arestas, então $m \leq 3n - 6$.*

Dem. Exercício. □

Corolário 5.116. *Se G é um grafo planar então $\delta(G) \leq 5$.*

Dem. Exercício. □

Como consequência temos que qualquer grafo planar é colorável com seis cores.

Proposição 5.117. *Se G é um grafo planar então $\chi(G) \leq 6$.*

Dem. Resulta imediatamente da proposição 5.112 e do corolário 5.116 (note-se que um subgrafo de um grafo planar é um grafo planar). □

Este resultado não é optimal, não há grafos planares que necessitem de cinco ou mais cores. É isto que diz o famoso **teorema das quatro cores**.

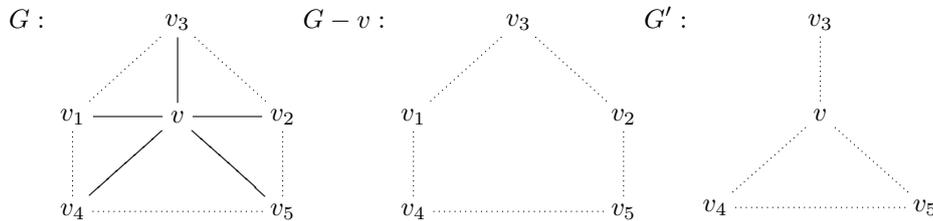
Teorema 5.118. *Se G é um grafo planar então $\chi(G) \leq 4$.*

Este resultado (nada trivial) não pode ser melhorado pois existem grafos planares com número cromático igual a quatro, por o grafo completo K_4 . A sua demonstração só foi possível com o recurso a meios computacionais de modo a analisar centenas de configurações locais. É, no entanto, muita mais fácil demonstrar o seguinte resultado mais fraco.

Teorema 5.119. *Se G é um grafo planar então $\chi(G) \leq 5$.*

Dem. Para demonstrar tal por indução basta ver que a propriedade $\chi(G) \leq 5$ é hereditária para grafos planar.

Dado um grafo planar G , assumamos que $\chi(H) \leq 5$ para qualquer grafo planar H com menos vértices que G . Seja v um vértice de G com grau mínimo (sabemos pelo corolário 5.116 que $d(v) \leq 5$). Podemos colorir os vértices de $G - v$ com 5 cores. Se $d(v) \leq 4$ então podemos estender essa coloração para o vértice v pois há pelo menos uma cor livre, logo $\chi(G) \leq 5$. No caso em que $d(v) = 5$ temos que existem dois vértices vizinhos v_1 e v_2 que não são adjacentes¹⁷. Nesse caso podemos construir um novo grafo planar G' com menos vértices que G , identificando em $G - v$ os vértice v_1 e v_2 .



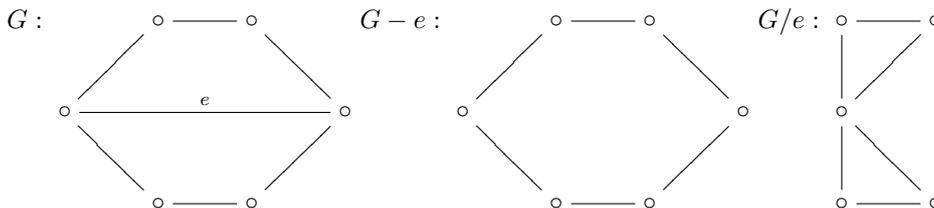
Uma 5-coloração de G' induz uma 5-coloração de $G - v$ onde v_1 e v_2 recebem a mesma cor, portanto pode ser estendida a G . \square

Definição 5.120. Dado um grafo $G = (V, E)$ e uma aresta e , denota-se por $G - e$ o grafo obtido por *eliminação* da aresta e . Ou seja, $G - e = (V, E \setminus \{e\})$.

Definição 5.121. Dado um grafo $G = (V, E)$ e uma aresta $e = [u, v]$, denota-se por G/e o grafo obtido por *contração* da aresta e . Tal consiste em remover os vértices extremos u e v da aresta e e substituir-los por um novo vértice w que é adjacentes a todos os outros vértices que eram adjacentes a u ou v . Ou seja, $G/e = (V', E')$ onde $V' = (V \setminus \{u, v\}) \cup \{w\}$ e $E' = \{[v_i, v_j] \in E : v_i, v_j \notin \{u, v\}\} \cup \{[v_i, w] : [v_i, u] \in E \text{ ou } [v_i, v] \in E\}$.

Ao grafo G/e chamamos uma *contração elementar* de G .

Exemplo 5.122.



Definição 5.123. Se um grafo H é obtido do grafo G por uma sequência de contrações elementares dizemos que H é uma *contração* de G . Um *menor* de G é um grafo que é contração de um subgrafo de G .

Note-se que se um grafo G é planar então qualquer menor de G também é planar. Portanto para mostrar que um grafo não é planar basta encontrar um menor que não o seja. O próximo resultado, conhecido como **teorema de Wagner** e que não iremos demonstrar, reduz essa procura apenas a dois grafo, K_5 e $K_{3,3}$.

¹⁷Se os cinco vértices vizinhos fossem adjacentes entre si, teríamos que o grafo completo K_5 seria planar e que não é verdade como facilmente se deduz da proposição 5.115.

Teorema 5.124. *Um grafo é planar se e só se não contém o K_5 ou o $K_{3,3}$ como menor.*

A seguinte conjectura mantém-se em aberto e relaciona o número cromático de grafo com o seus menores. É conhecida por **conjectura de Hadwiger**.

Conjectura 5.125. *Se $\chi(G) \geq n$ então G contém o K_n como menor.*

Esta conjectura está em aberto para $n \geq 7$. Para $n = 1$ e $n = 2$ é trivial e o caso $n = 3$ é consequência imediata da proposição 5.95, os casos $n = 5$ e $n = 6$ provou-se ser equivalentes ao teorema das quatro cores.

Definição 5.126. Chamamos *polinómio cromático* do grafo G à função $P_G : \mathbb{N} \rightarrow \mathbb{N}$ que a cada natural k atribui o número de k -colorações de G .

Propriedades:

- (1) Se $k < \chi(G)$ então $P_G(k) = 0$;
- (2) $P_{K_n}(k) = k(k-1)(k-2) \cdots (k-n)$;
- (3) $P_{\bar{K}_n}(k) = k^n$ onde \bar{K}_n é o complementar de K_n (i.e. o grafo vazio com n vértices);
- (4) $P_{G \sqcup H} = P_G P_H$ onde $G \sqcup H$ é a união disjunta dos grafos G e H ;
- (5) $P_G = P_{G-e} - P_{G/e}$;
- (6) P_G é um polinómio mónico de grau igual à ordem de G .

Colorações de arestas.

Definição 5.127. Uma k -coloração de arestas de um grafo $G = (V, E)$ é uma decomposição de E em k matchings M_1, \dots, M_k . O *índice cromático* ou *número de coloração de arestas* de G , denotado¹⁸ por $\chi'(G)$, é menor k para o qual existe uma k -coloração de arestas de G .

Observação 5.128. $\chi'(G) \geq \Delta(G)$.

Por outro lado, o próximo resultado, conhecido como **teorema de Vizing**, dá-nos um majorante para o índice cromático bastante próximo do anterior minorante.

Teorema 5.129. *Para qualquer grafo (simples) $G = (V, E)$ temos que $\chi'(G) \leq \Delta(G) + 1$.*

Deste modo os grafos dividem-se em duas classes, os grafos com $\chi' = \Delta$ e os grafos com $\chi' = \Delta + 1$.

De entre os grafos que pertencem à primeira classe estão os *grafos cúbicos* (i.e. grafos regulares de grau três) planares sem pontes. Tal afirmação é equivalente ao teorema das quatro cores, como foi demonstrado por Tait.

Teorema 5.130. *Qualquer grafo cúbico planar sem pontes é tricolorável nas arestas.*

Dentro da classe dos grafo com $\chi' = \Delta$ também estão os grafos bipartidos como mostra o seguinte resultado, chamado de **teorema da coloração de aresta de König**.

Teorema 5.131. *Para um grafo bipartido $G = (V, E)$ temos que $\chi'(G) = \Delta(G)$.*

Dem. Se $\Delta(G) = 2$, então G é uma união de caminhos e ciclos (de ordem par pois G é bipartido) logo as arestas de G podem ser coloridas com 2 cores. Se $\Delta(G) = 1$ então E é um matching de G logo $\chi'(G) = 1$.

No caso geral, considere-se uma $\Delta(G)$ -coloração parcial de arestas de G com o número máximo possível de arestas coloridas de modo que arestas adjacentes tenham cores diferentes. Suponhamos que existe uma aresta e que não foi colorida (senão já temos o resultado pretendido). Essa aresta liga dois vértices v_1 e v_2 . Como estamos a colorir com $\Delta(G)$ cores tanto v_1 como v_2 não são incidentes a pelo menos um cor cada. Podemos supor que v_1 não é incidente a uma cor (digamos vermelho) e v_2 não é incidente a outra¹⁹ cor (digamos azul). Sendo assim, tomamos o subgrafo de G induzido pelas arestas de cor azul e vermelho mais a aresta e . Tal subgrafo H é bipartido e tem

¹⁸No texto "A Course in Combinatorial Optimization" de A. Schrijver o número coloração de arestas é denotado por $\chi(G)$.

¹⁹Se v_1 e v_2 não fossem incidentes a uma mesma cor, essa cor poderia ser usada para colorir a aresta e .

$\Delta(H) \leq 2$, logo pode ser colorido com as cores vermelho e azul de modo a colorir também a aresta e . Obtemos então uma nova coloração com mais arestas coloridas que a anterior contradizendo o facto de esta colorir o número máximo possível de arestas. \square

Esta demonstração dá-nos um algoritmo para colorir um as arestas de um grafo bipartido em tempo polinomial. Além disso, podemos obter uma k -coloração num grafo $G = (V, E)$ com $\chi'(G) \leq k$ de modo que as cardinalidades de cada par de matchings (classes de cores) difiram no máximo em um. Ou seja, cada classe de cor tem $\lfloor \frac{|E|}{k} \rfloor$ ou $\lceil \frac{|E|}{k} \rceil$ arestas. O modo de o conseguir é tomar duas classes de cor M_i e M_j em que as cardinalidades difiram em mais que um ($|M_i| > |M_j| + 1$), e trocar as cores numa das componentes conexas do subgrafo induzido pelas classes que tenhas mais arestas em M_i que em M_j .

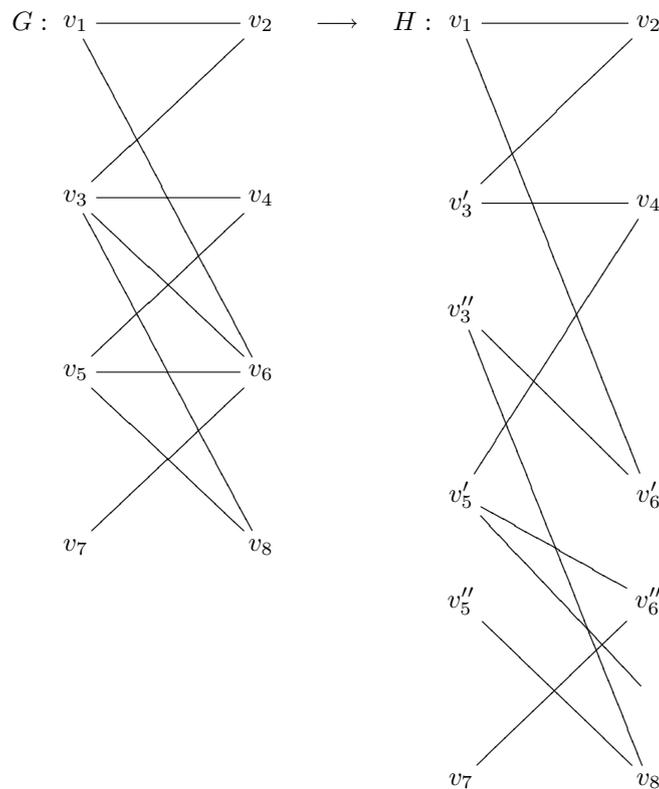
Corolário 5.132. *Dado um grafo G seja $\xi(G)$ o número máximo de cobertura de arestas disjuntas duas a duas. Se G é bipartido então*

$$\xi(G) = \delta(G)$$

Ou seja, o número máximo de cobertura de arestas disjuntas duas a duas é igual ao grau mínimo de G .

Dem. Como cada vértice tem que ser coberto por uma cobertura de arestas é fácil ver que $\xi(G) \leq \delta(G)$.

Por outro lado, podemos construir de G um grafo bipartido H em que os graus dos vértices não são superiores a $\delta(G)$. Basta ir dividindo cada vértice v com grau maior que $\delta(G)$ em dois vértices v' e v'' , um com grau $\delta(G)$ e o outro com grau $d(v) - \delta(G)$ até não restar vértices com grau superior a $\delta(G)$.

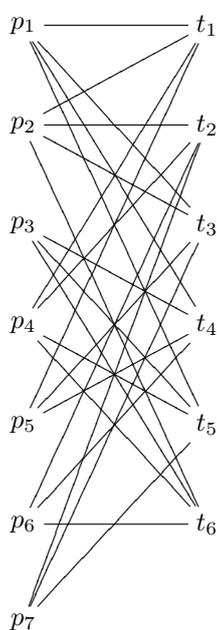


Temos então que, pelo teorema de coloração de arestas de Kőnig, que H possui uma δ -coloração de arestas. Cada classe de cor corresponde a uma cobertura de arestas em G logo há pelo menos $\delta(G)$ coberturas de arestas disjuntas duas a duas, ou seja, $\xi(G) \geq \delta(G)$. \square

Exemplo 5.133. Consideremos o problema de encontrar um horário o mais compacto possível, que seja compatível com o seguinte quadro de atribuição de professores (p_1, \dots, p_7) por turmas (t_1, \dots, t_6) :

	t_1	t_2	t_3	t_4	t_5	t_6
p_1	X		X	X	X	
p_2	X	X	X			X
p_3				X	X	X
p_4	X	X			X	X
p_5	X		X	X		
p_6		X		X		X
p_7		X	X		X	

Tal problema consiste em colorir com o menor número de cores possível as arestas do seguinte grafo bipartido:



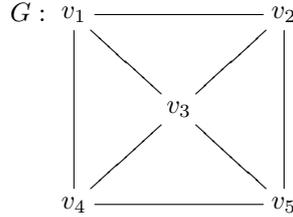
Pelo teorema de Kőnig, como $\Delta = 4$, bastam quatro cores. Logo, é possível construir um horário em que as aulas sejam todas dadas em quatro horas e além disso, pode ser feito de modo a não haver mais de seis $(\lceil \frac{|E|}{4} \rceil)$ aulas em cada hora. Por exemplo:

	sala 1	sala 2	sala 3	sala 4	sala 5	sala 6
1ª hora	(p_1, t_1)	(p_2, t_2)	(p_5, t_3)	(p_6, t_4)	(p_7, t_5)	(p_4, t_6)
2ª hora	(p_2, t_1)	(p_4, t_2)	(p_1, t_3)	(p_5, t_4)	(p_3, t_5)	(p_6, t_6)
3ª hora	(p_5, t_1)	(p_7, t_2)	(p_2, t_3)	(p_1, t_4)	(p_4, t_5)	(p_3, t_6)
4ª hora	(p_4, t_1)	(p_6, t_2)	(p_7, t_3)	(p_3, t_4)	(p_1, t_5)	(p_2, t_6)

Grafos perfeitos.

Definição 5.134. Um grafo G diz-se *perfeito* se $\omega(G') = \chi(G')$ para qualquer subgrafo induzido G' de G .

Exemplo 5.135. O seguinte grafo



é perfeito. Todos o seus subgrafos induzidos têm número cromático igual ao número de clique.

$S \subseteq V$	$G' = G[S]$	$\omega(G')$	$\chi(G')$
$\{v_1\}, \{v_2\}, \{v_3\}, \{v_4\}, \{v_5\}$	\bullet	1	1
$\{v_1, v_2\}, \{v_1, v_3\}, \{v_1, v_4\}, \{v_2, v_3\}, \{v_2, v_5\}, \{v_3, v_4\}, \{v_3, v_5\}, \{v_4, v_5\}$	$\bullet \text{ --- } \bullet$	2	2
$\{v_1, v_5\}, \{v_2, v_4\}$	$\bullet \quad \bullet$	1	1
$\{v_1, v_2, v_3\}, \{v_1, v_3, v_4\}, \{v_2, v_3, v_5\}, \{v_3, v_4, v_5\}$	$\bullet \text{ --- } \bullet$ $\quad \diagdown$ $\quad \bullet$	3	3
$\{v_1, v_2, v_4\}, \{v_1, v_2, v_5\}, \{v_1, v_3, v_5\}, \{v_3, v_4, v_5\}, \{v_2, v_3, v_4\}, \{v_2, v_4, v_5\}$	$\bullet \text{ --- } \bullet \text{ --- } \bullet$	2	2
$\{v_1, v_2, v_3, v_4\}, \{v_1, v_2, v_3, v_5\}, \{v_1, v_3, v_4, v_5\}, \{v_2, v_3, v_4, v_5\}$	$\bullet \text{ --- } \bullet$ $\quad \diagdown$ $\quad \bullet$ $\quad \bullet$	3	3
$\{v_1, v_2, v_4, v_5\}$	$\bullet \text{ --- } \bullet$ $\quad \bullet$ $\quad \bullet$	2	2
$\{v_1, v_2, v_3, v_4, v_5\}$	$\bullet \text{ --- } \bullet$ $\quad \diagdown$ $\quad \bullet$ $\quad \bullet$ $\quad \bullet$	3	3

Outros exemplos de grafos perfeitos são os grafos completos e os seus complementares, bem como os grafos bipartidos.

Teorema 5.136. Um grafo é perfeito se e só se $\omega(G')\alpha(G') \geq |V(G')|$ para cada subgrafo induzido G' de G .

Dem. A condição necessária é fácil de verificar. Se G é perfeito e G' é um subgrafo induzido então $\omega(G') = \chi(G')$ logo $\omega(G')\alpha(G') = \chi(G')\alpha(G') \geq |V(G')|$ pois $V(G')$ pode ser decomposto em $\chi(G')$ cocliques todas elas de tamanho não superior a $\alpha(G')$.

Para demonstrar a condição recíproca. Seja $G = (V, E)$ um grafo não perfeito que satisfaça a condição $\omega(G')\alpha(G') \geq |V(G')|$ para todo o subgrafo induzido G' de G , com cardinalidade mínima. Logo para qualquer subgrafo induzido G' de G temos $\omega(G') = \chi(G')$ se $G' \neq G$ e $\omega(G) < \chi(G)$.

Seja $\omega := \omega(G)$ e $\alpha := \alpha(G)$ e assumamos $V(G) = \{1, \dots, n\}$.

Seja C_0 uma coclique de tamanho α . Por minimalidade de G , o subgrafo de G induzido por $V(G) \setminus \{v\}$ é perfeito para qualquer vértice v , logo $V(G) \setminus \{v\}$ pode ser decomposto em ω cocliques (pois $\chi(G[V(G) \setminus \{v\}]) = \omega(G[V(G) \setminus \{v\}]) \leq \omega$). Fazendo isto para cada $v \in C_0$, obtemos

$\alpha\omega + 1$ cocliques (não necessariamente distintas entre si) $C_0, C_1, \dots, C_{\alpha\omega}$ tais que cada vértice de G é coberto exactamente por α destas cocliques. (*)

Para cada $i = 0, \dots, \alpha\omega$ existe uma clique K_i de tamanho ω tal que $K_i \cap C_i = \emptyset$. Senão o subgrafo G' de G induzido por $V(G) \setminus C_i$ seria tal que $\omega(G') < \omega$ o que implicaria, pelo facto de G' ser perfeito, que $\chi(G') \leq \omega - 1$, logo $\chi(G) \leq \omega$ o que contradiz o que já foi dito acerca de G .

Além disso, temos que se $i \neq j$ então $|K_i \cap C_j| = 1$. Isto acontece porque, por (*), cada um dos ω vértices em K_i é coberto por α das cocliques $C_0, \dots, C_{\alpha\omega}$. Como cada clique só intersecta um coclique no máximo em um vértice, há $\alpha\omega$ cocliques a intersectar K_i . Ou seja, todas as cocliques $C_0, \dots, C_{\alpha\omega}$ à excepção de C_i intersectam K_i em um vértice ($|K_i \cap C_j| = 1 \forall i \neq j$).

Consideremos agora as matrizes de incidência $M, N \in \{0, 1\}^{\alpha\omega+1, n}$ de $C_0, \dots, C_{\alpha\omega}$ e $K_0, \dots, K_{\alpha\omega}$ respectivamente. Ou seja, $M = [m_{i,j}]$ com $m_{i,j} = 1 \Leftrightarrow j \in C_i$ e $N = [n_{i,j}]$ com $n_{i,j} = 1 \Leftrightarrow j \in K_i$. Temos então que $MN^T = J - I$ onde J é a matriz $(\alpha\omega + 1) \times (\alpha\omega + 1)$ com todas as entradas iguais a 1 e I é a matriz identidade.

Teríamos então que $|V(G)| = n \geq \text{car}(M) \geq \text{car}(MN^T) = \alpha\omega + 1 > \alpha\omega$ o que contradiz a definição de G . \square

Como corolário deste resultado temos o **teorema do grafo perfeito de Lovász**.

Corolário 5.137. *O complemento de um grafo perfeito é também um grafo perfeito.*

Dem. Basta ver que, se $G[U]$ é o subgrafo de G induzido pelo conjunto $U \subseteq V(G)$ então o seu complemento é o subgrafo de \bar{G} (complemento de G) induzido pelo mesmo conjunto U . Ou seja, $\bar{G}[U] = \bar{G}[U]$. Depois segue naturalmente que $\omega(\bar{G}[U]) = \alpha(G[U])$ e $\alpha(\bar{G}[U]) = \omega(G[U])$, logo $\omega(\bar{G}[U])\alpha(\bar{G}[U]) \geq |V(\bar{G}[U])|$. \square

Note-se que, como o complemento do complemento de um grafo é o próprio grafo, verificar se um grafo é perfeito é equivalente a verificar se o seu complemento é perfeito. Tal pode ser útil quando o complemento de um grafo é mais simples (por ter menos arestas ou ser bipartido por exemplo) que o próprio grafo. Em particular, temos que o complemento de qualquer grafo bipartido é perfeito.

Corolário 5.138. *O complemento de um grafo bipartido é um grafo perfeito. Equivalentemente, o número de cobertura de arestas $\rho(G)$ de um grafo bipartido G sem vértices isolados é igual ao número de coclique $\alpha(G)$ de G . Ou seja, $\rho(G) = \alpha(G)$.*

Note-se que isto não é mais que o teorema de cobertura de arestas de König.

Dem. Exercício. \square

Também o teorema de coloração de arestas de König resulta como corolário do teorema do grafo perfeito de Lovász.

Corolário 5.139. *O grafo linha²⁰ $L(G)$ de um grafo bipartido G é um grafo perfeito. Equivalentemente, o número de coloração de arestas $\chi'(G)$ de um grafo bipartido G é igual ao grau máximo $\Delta(G)$ de G . Ou seja, $\chi'(G) = \Delta(G)$.*

Dem. Exercício. \square

Definição 5.140. Dado um conjunto parcialmente ordenado (X, \leq) , um subconjunto C de X é chamado de *cadeia* se todos os seus membros são comparáveis (i.e. $\forall x, y \in C, x \leq y$ ou $y \leq x$). Um subconjunto A de X é chamado de *anticadeia* se todos os seus membros não são comparáveis entre si (i.e. $\forall x, y \in A, x \not\leq y$ e $y \not\leq x$).

Definição 5.141. Dado um conjunto (finito) parcialmente ordenado (X, \leq) , definimos o *grafo de comparabilidade* ao grafo $G = (V, E)$ cujos vértices são os elementos de X (i.e. $V = X$) e dois elementos distintos são adjacentes se e só se são comparáveis (i.e. $[x, y] \in E \Leftrightarrow x < y$ ou $y < x$).

Teorema 5.142. *Num conjunto parcialmente ordenado finito (X, \leq) o tamanho máximo de uma cadeia é igual ao número mínimo de anticadeias necessárias para cobrir X .*

²⁰Dado um grafo $G = (V, E)$ o seu grafo linha $L(G)$ é o grafo cujos vértices são as arestas de G e dois vértices (em $L(G)$) são adjacentes sse as arestas de G que eles representam também o são.

Dem. Para cada elemento x de X definimos $\phi(x)$ como sendo o tamanho máximo de uma cadeia que tenha x como elemento máximo. Seja k o valor máximo de ϕ em X . Para $i = 1, \dots, k$ seja A_i o conjunto dos elementos x com $\phi(x) = i$. Então A_1, \dots, A_k são k anticadeias cobrindo X .

Por outro lado, como qualquer anticadeia intersecta qualquer cadeia em no máximo um elemento, temos que são necessárias pelo menos k anticadeias para cobrir X .

Portanto temos que o tamanho máximo de uma cadeia em X (que é igual a k) será igual o número mínimo de anticadeias necessárias para cobrir X . \square

Note-se que num grafo de comparabilidade as cliques são as cadeias e as cocliques são anticadeias. Portanto o que o resultado anterior diz é que para um grafo de comparabilidade $\omega = \chi$. Como um subgrafo induzido de um grafo de comparabilidade é um grafo de comparabilidade, temos o seguinte resultado.

Corolário 5.143. *Qualquer grafo (finito) de comparabilidade é um grafo perfeito.*

Resulta então do teorema do grafo perfeito o seguinte resultado, conhecido por **teorema da decomposição de Dilworth**.

Corolário 5.144. *Num conjunto parcialmente ordenado finito (X, \leq) o tamanho máximo de uma anticadeia é igual ao número mínimo de cadeias necessárias para cobrir X .*

Dem. Basta ver que se G é o grafo de comparabilidade deste conjunto parcialmente ordenado. Então o tamanho máximo de uma anticadeia é igual a $\omega(\overline{G})$ e o número mínimo de cadeias necessárias para cobrir X é igual a $\chi(\overline{G})$, onde \overline{G} é o completo de G . Logo se G é perfeito então \overline{G} também o é e portanto $\omega(\overline{G}) = \chi(\overline{G})$. \square

O teorema do grafo perfeito de Lovász não é a última palavra acerca dos grafos perfeitos. De facto, Chudnovsky, Robertson, Seymour e Thomas demonstraram em 2002 um resultado conjecturado por Berge [1963] e que ficou conhecido como **teorema forte do grafo perfeito**. Este diz o seguinte.

Teorema 5.145. *Um grafo é perfeito se e só se G não contém nenhum ciclo de ordem ímpar C_{2k+1} com $k \geq 2$ ou o seu complemento \overline{C}_{2k+1} como subgrafo induzido.*

6. MATRÓIDES

A teoria dos matróides pode ser vista como uma abordagem combinatória dos conceitos de álgebra linear como o de conjunto independente, base, dimensão entre outros. A teoria unifica conceitos de álgebra linear e da teoria de grafos.

Definição 6.1. Um *matróide* $M = (X, \mathcal{I})$ consiste num conjunto finito X e numa família \mathcal{I} de subconjuntos de X que satisfaz os seguintes axiomas:

- (i) $\emptyset \in \mathcal{I}$;
- (ii) se $Y \in \mathcal{I}$ e $Z \subseteq Y$ então $Z \in \mathcal{I}$;
- (iii) se $Y, Z \in \mathcal{I}$ e $|Y| < |Z|$ então $Y \cup \{x\} \in \mathcal{I}$ para algum $x \in Z \setminus Y$.

Um subconjunto Y de X é chamado de *independente* se $Y \in \mathcal{I}$ e chamado de *dependente* se $Y \notin \mathcal{I}$.

Exemplo 6.2. Dado um conjunto finito de vectores X num dado espaço linear, podemos formar um matróide $M = (X, \mathcal{I})$ tomando por \mathcal{I} o conjunto dos subconjuntos de X linearmente independentes. Em particular uma matriz A com entradas num corpo \mathbb{K} pode ser vista como um matróide $M[A] = (X, \mathcal{I})$ onde X é o conjunto das suas colunas e \mathcal{I} é o conjunto dos conjuntos de colunas de A linearmente independentes. Um matróide deste tipo (ou isomorfo²¹ a um tal matróide) é chamado de *matróide linear*.

Exemplo 6.3. Dado um grafo $G = (V, E)$ se tomarmos o conjunto $\mathcal{I} \subseteq 2^E$ formado pelos conjuntos de arestas E' tais que (V, E') forma uma floresta (i.e. um grafo acíclico) então $M(G) = (E, \mathcal{I})$ é um matróide, chamado de *matróide cíclico de G* . Qualquer matróide obtido deste modo, ou isomorfo a tal, é chamado de *matróide gráfico*.

²¹Dois matróides (X, \mathcal{I}) e (X', \mathcal{I}') dizem-se *isomorfos* se existir uma bijecção $\phi : X \rightarrow X'$ tal que $Y \in \mathcal{I}$ se e só se $\phi(Y) \in \mathcal{I}'$.

Observação 6.4. Os matróides gráficos são matróides lineares. De facto, o matróide cíclico de um dado grafo pode ser visto como o matróide associado à sua matriz de incidência vista como uma matriz com entradas no corpo \mathbb{F}_2 (corpo de ordem 2).

Definição 6.5. Dado um matróide (X, \mathcal{I}) e um subconjunto Y de X , dizemos que um subconjunto independente B de Y é uma *base* de Y se B é maximal (em relação à ordem de inclusão) de entre os subconjuntos independentes de Y . Ou seja, para cada $Z \in \mathcal{I}$ com $B \subseteq Z \subseteq Y$ temos $Z = B$.

A uma base de X chamamos simplesmente *base*.

Proposição 6.6. *A condição (iii) da definição dos conjuntos independentes pode ser substituída pela afirmação:*

Para qualquer subconjunto Y de X , quaisquer bases de Y têm a mesma cardinalidade.

Dem. Exercício. □

Definição 6.7. Dado um matróide $M = (X, \mathcal{I})$ e um subconjunto Y de X , definimos a *característica* de Y , $r_M(Y)$, como sendo a cardinalidade das suas bases.

Consideremos o seguinte problema de optimização. Dada uma função peso $w : X \rightarrow \mathbb{R}$ e uma família \mathcal{I} de subconjuntos de X queremos determinar um conjunto $Y \in \mathcal{I}$ que maximize $w(Y) := \sum_{y \in Y} w(y)$.

Uma tentativa de resolução será usar um *algoritmo ganancioso* que consiste em começar com o conjunto vazio e ir seleccionando y_1, \dots, y_r sucessivamente como se segue. Se y_1, \dots, y_k já foram seleccionados, escolhemos $y \in X$ tal que:

- (i) $y \notin \{y_1, \dots, y_k\}$ e $\{y_1, \dots, y_k, y\} \in \mathcal{I}$,
- (ii) y é de todos os elementos de X que satisfazem condição (i) aquele que tem peso $w(y)$ máximo.

Paramos quando não existir elementos y de X que satisfazem a condição (i). Ou seja, paramos quando o conjunto obtido for uma base.

Teorema 6.8. *Seja X um conjunto finito e \mathcal{I} uma família de subconjuntos de X tal que*

- (i) $\emptyset \in \mathcal{I}$;
- (ii) se $Y \in \mathcal{I}$ e $Z \subseteq Y$ então $Z \in \mathcal{I}$;

Então (X, \mathcal{I}) é um matróide se e só se o algoritmo ganancioso termina com um conjunto $Y \in \mathcal{I}$ com peso $w(Y)$ máximo, para cada função peso $w : X \rightarrow \mathbb{R}_0^+$.

Dem. Suponhamos que (X, \mathcal{I}) é um matróide e seja $w : X \rightarrow \mathbb{R}_0^+$ uma função peso qualquer. Como os pesos são todos não-negativos o peso máximo será atingido em bases. Vamos chamar *ganancioso* a conjunto independente Y se estiver contido numa base de peso máximo. Basta mostrar que se Y é ganancioso e $x \in X \setminus Y$ é tal que $w(x)$ é o maior valor possível com $Y \cup \{x\} \in \mathcal{I}$ então $Y \cup \{x\}$ é também ganancioso. Como Y é ganancioso está contido numa base B de peso máximo. Se $x \in B$ então $Y \cup \{x\}$ também é ganancioso. Se $x \notin B$, então existe uma base B' contendo $Y \cup \{x\}$. Logo $B' = (B \setminus \{x'\}) \cup \{x\}$ para algum $x' \in B \setminus Y$. Como $w(x)$ é escolhido como sendo máximo, $w(x) \geq w(x')$, portanto $w(B') \geq w(B)$. Logo B' é uma base de peso máximo pelo que $Y \cup \{x\}$ é ganancioso.

Suponhamos agora que (X, \mathcal{I}) não é um matróide. Então existem $Y, Z \in \mathcal{I}$ tais que $|Y| < |Z|$ e $Y \cup \{z\} \notin \mathcal{I}$ para qualquer $z \in Z \setminus Y$. Seja $k := |Y|$, consideremos a seguinte função peso $w : X \rightarrow \mathbb{R}_0^+$ dada por:

$$w(x) = \begin{cases} k + 2 & \text{se } x \in Y \\ k + 1 & \text{se } x \in Z \setminus Y \\ 0 & \text{se } x \in X \setminus (Y \cup Z) \end{cases}$$

A solução obtida neste caso pelo algoritmo ganancioso será uma base com o mesmo peso que Y pois vai escolhendo primeiro todos os elementos de Y (que são os que têm peso maior) e depois como não podemos juntar elementos de $Z \setminus Y$ resta-nos juntar elementos com peso nulo. Ou seja, o algoritmo leva-nos a uma base de peso $|Y|(k+2) = k(k+2)$. No entanto, qualquer base contendo

Z tem peso de pelo menos $w(Z) \geq |Z|(k+1) \geq (k+1)(k+1) > k(k+2)$. Logo o algoritmo não produz, neste caso, uma solução optimal. \square

Axiomas equivalentes para matróides.

Definição 6.9. Dado um matróide (X, \mathcal{I}) , um subconjunto de X diz-se um *circuito* se for um conjunto dependente minimal para ordem de inclusão. Ou seja, C é um circuito se $C \notin \mathcal{I}$ e $(Y \subsetneq C \Rightarrow Y \in \mathcal{I})$.

Temos então que uma vez dado X e definidos os conjuntos independentes \mathcal{I} , podemos definir as bases \mathcal{B} , os circuitos \mathcal{C} e a função característica $r : 2^X \rightarrow \mathbb{N}_0$. No entanto, podemos recuperar os conjuntos independentes \mathcal{I} a partir das bases \mathcal{B} , dos circuitos \mathcal{C} ou da função característica $r : 2^X \rightarrow \mathbb{N}_0$.

De facto, temos a seguinte proposição.

Proposição 6.10. *Seja (X, \mathcal{I}) um matróide e \mathcal{B} o conjunto das bases, \mathcal{C} o conjunto dos circuitos e r a função característica. Então o conjunto dos conjuntos independentes \mathcal{I} pode ser definido do seguinte modo:*

- (1) $\mathcal{I} = \{Y : Y \subseteq B \text{ para algum } B \in \mathcal{B}\};$
- (2) $\mathcal{I} = \{Y : \text{não existe } C \in \mathcal{C} \text{ tal que } C \subseteq Y\};$
- (3) $\mathcal{I} = \{Y : r(Y) = |Y|\}.$

Assim podemos definir o conceito de matróide alternativamente através uma axiomática apropriada para definir as bases ou os circuitos ou a função característica. Temos então os seguintes resultados.

Teorema 6.11. *Seja X um conjunto finito e \mathcal{B} uma família de subconjuntos de X . Definindo $\mathcal{I} = \{Y : Y \subseteq B \text{ para algum } B \in \mathcal{B}\}$ temos que (X, \mathcal{I}) é um matróide se e só se*

- B1 $\mathcal{B} \neq \emptyset$;*
- B2 se $B_1, B_2 \in \mathcal{B}$ e $x \in B_1 \setminus B_2$ então $(B_1 \setminus \{x\}) \cup \{y\} \in \mathcal{B}$ para algum $y \in B_2 \setminus B_1$.*

Dem. Exercício. \square

Teorema 6.12. *Seja X um conjunto finito e \mathcal{C} uma família de subconjuntos de X . Definindo $\mathcal{I} = \{Y : \text{não existe } C \in \mathcal{C} \text{ tal que } C \subseteq Y\}$ temos que (X, \mathcal{I}) é um matróide se e só se*

- C1 $\emptyset \notin \mathcal{C}$;*
- C2 se $C_1 \in \mathcal{C}$ e $C_2 \subseteq C_1$ então $C_1 = C_2$;*
- C3 se $C_1, C_2 \in \mathcal{C}$ e $e \in C_1 \cup C_2$ então existe $C_3 \in \mathcal{C}$ tal que $C_3 \subseteq (C_1 \cup C_2) \setminus \{e\}$.*

Dem. Exercício. \square

Teorema 6.13. *Dado um conjunto finito X e uma função $r : 2^X \rightarrow \mathbb{N}_0$. Definindo $\mathcal{I} = \{Y : r(Y) = |Y|\}$ temos que (X, \mathcal{I}) é um matróide se e só se*

- R1 $0 \leq r(Y) \leq |Y|$;*
- R2 se $Y \subseteq Z$ então $r(Y) \leq r(Z)$;*
- R3 se $Y, Z \subseteq X$ então $r(Y \cap Z) + r(Y \cup Z) \leq r(Y) + r(Z)$.*

Dem. Exercício. \square

Exemplos de matróides e construções de novos matróides. Já foram dados os exemplos dos matróides lineares e dos matróides gráficos.

Um matróide linear $M = (X, \mathcal{I})$ é representado por um conjunto finito X de vectores num espaço linear. Os subconjuntos independentes são precisamente os subconjuntos linearmente independentes, a função característica é dada por $r(Y) = \dim(\text{span}Y)$ e uma base $B \subseteq X$ é um conjunto tal que $r(B) = |B| = r(X)$. Note-se que dois conjuntos distintos de vectores podem representar o mesmo matróide (a menos de isomorfismo). Um matróide M diz-se \mathbb{K} -representável se pode ser representado por um conjunto de vectores num espaço linear sobre o corpo \mathbb{K} . O facto de um matróide poder ser representado num dado corpo não implica que possa ser representado em qualquer corpo.

Um matr oide grfico $M = (E, \mathcal{I})$   representado por um grafo $G = (V, E)$. Os subconjuntos independentes so os subconjuntos de arestas que induzem um subgrafo acclico, os circuitos de M so os ciclos do grafo G , a funo caracterstica   dada por $r(Y) = |V(G')| - 1$ onde G'   o subgrafo induzido por Y e uma base $B \subseteq E$ induz uma floresta que cobre todos os vrtices no isolados de G . Note-se que dois grafos distintos podem representar o mesmo matr oide (a menos de isomorfismo).

Outra classe importante de matr oides so os *matr oides transversais*. Seja (X_1, \dots, X_m) uma famlia indexada de subconjuntos, no necessariamente distintos, de um conjunto finito X . Um conjunto $Y \subseteq X$ diz-se um *transversal* (de X_1, \dots, X_m) se existe em bijeo $\phi : \{1, \dots, m\} \rightarrow Y$ tal que $\phi(j) \in X_j$ para todo o $j = 1, \dots, m$. Um subconjunto de um transversal (de X_1, \dots, X_m) diz-se um *transversal parcial* (de X_1, \dots, X_m). Seja \mathcal{I} o conjunto dos transversais parciais de X_1, \dots, X_m ento (X, \mathcal{I}) forma um matr oide. Um matr oide obtido deste modo, ou isomorfo a tal,   chamado de *matr oide transversal*. As bases de (X, \mathcal{I}) so precisamente os transversais de X_1, \dots, X_m . Um matr oide transversal pode ser representado por um grafo bipartido. Seja $G = (V, E)$ um grafo bipartido cujo o conjunto vrtices   $V = \{1, \dots, m\} \cup X$ e o conjunto de arestas   $E = \{[i, x] : 1 \leq i \leq m \text{ e } x \in X_i\}$, temos que $Y \subseteq X$   um transversal parcial se e so se cobre um matching de G (i.e. existe um matching em G cujo cada aresta termina num elemento de Y).

Entre os matr oides transversais encontram-se os *matr oides uniformes*. Dados dois inteiros no-negativos m e n , com $m \leq n$, chamamos matr oide uniforme $U_{m,n}$ a um matr oide (X, \mathcal{I}) com $|X| = n$ e $\mathcal{I} = \{Y \subseteq X : |Y| = m\}$. Este pode ser visto como um matr oide transversal associado a X com famlia indexada (X_1, \dots, X_m) com $X_1 = \dots = X_m = X$.

Podemos construir outros matr oides atravs dos seguintes procedimentos.

Proposio 6.14. *Seja M um matr oide sobre X com conjunto de bases \mathcal{B} . Ento o conjunto $\mathcal{B}^* := \{X \setminus B : B \in \mathcal{B}\}$   o conjunto base de um matr oide, denotado por M^* , sobre X . O matr oide M^*   chamado de matr oide dual de M .*

Dem. Exerccio. □

O matr oide dual de matr oide grfico so chamados de *matr oide cogrficos*. Os duais de matr oide uniforme so tambm uniformes, de facto, temos $U_{k,n}^* = U_{n-k,n}$.

Definio 6.15. *Seja $M = (X, \mathcal{I})$ um matr oide, Y um subconjunto de X e $\mathcal{I}' := \{Z \subseteq Y : Z \in \mathcal{I}\}$, ento $M' = (Y, \mathcal{I}')$ um matr oide, chamado de *restrio* de M a Y . Se $Y = X \setminus Z$ com $Z \subseteq X$, dizemos que M'   obtido por *eliminao* de Z , e denotamos M' por $M \setminus Z$.*

Definio 6.16. *Seja $M = (X, \mathcal{I})$ um matr oide, Z um subconjunto de X . Ao matr oide $(M^* \setminus Z)^*$ chamamos de *contrao* de M por Z e denotamos por M/Z .*

Definio 6.17. *Um matr oide M' obtido a partir de M por uma srie de contraoes e eliminaoes   chamado de *menor* de M .*