Sumário

Carlos Humes Jr. e Anaflora P. C. Humes

Revisado: Leônidas O. Brandão (Fev. e Maio de 2000)

27 de outubro de 2000

Capítulo 2

A estrutura do conjunto de pontos viáveis (caracterização de poliedros)

2.1 Introdução

Neste capítulo, iniciaremos o estudo de programação linear propriamente dita. Na medida em que o título apropriado para o material que se segue é: "caracterização de poliedros", a afirmação que começamos a estudar programação linear pode parecer falsa.

Retornaremos ao espírito da introdução. Nesta procuramos motivar o ponto de que a pergunta "o problema é fácil (difícil)?" tem "a priori" a resposta "depende". Mais do que isto, a resposta é "depende da estrutura".

Neste sentido, a pergunta "problemas de otimização são fáceis (difíceis)?" deve ser respondida com "depende da estrutura". A nós, parece que linearidade é uma forte estrutura ou pelo menos, uma estrutura bem conhecida.

No caso de programação linear, esta forte estrutura aparece tanto na função objetivo como no conjunto de pontos viáveis, que será sempre descrito por equações e inequações lineares. Esta constatação leva naturalmente a busca de resultados uteis para a solução do problema de programação linear, baseados na "estrutura linear". Dentro desta orientação definimos:

Definição 2.1 *Diz-se que P é um* **poliedro** [1] *se P coincide com a intersecção de um número finito de hiperplanos e semi-espaços.*

Esta definição algebricamente caracteriza poliedro como a solução de um sistema linear com um número finito de equações e inequações. (Cada equação do sistema linear define um hiperplano e cada inequação um semi-espaço.) Assim, um exemplo de poliedro é o conjunto $Y = \{\boldsymbol{y} \in \mathbb{R}^n \ / \ \mathbf{A}\boldsymbol{y} = \boldsymbol{b} \text{ e } \mathbf{B}\boldsymbol{y} \geqq \boldsymbol{d}\}$.

Note que a omissão do qualificador "finito" na definição acima, levar-nos-ia a classificar círculos como poliedros.

Uma classe particular de poliedros é a dos polígonos (supostos limitados). Uma pergunta interessante é como caracterizar polígonos no $\mathbb{R}^2$ (ou poliedros limitados "usuais" no $\mathbb{R}^3$), usando o mínimo de informação possível.

No caso do triângulo ABC da figura 2.1.(a), o triângulo é univocamente determinado pelos vértices. Já o pentágono $DEFGH$ não está univocamente determinado pelo conjunto de seus vértices, como pode ser visto

[1] Alguns autores usam esta definição e seguem com a definição "politopo é poliedro limitado". Outros autores usam a definição acima para politopos e definem "poliedro é um politopo limitado". Optamos pela definição acima e a não dar nome específico a poliedros limitados

nas figuras 2.1.(b),2.1.(c) e 2.1.(d). A resposta “natural” de que um polígono é determinado pelos seus vértices, é válida se considerarmos implicitamente polígonos convexos. Ou seja, se eliminarmos os casos de polígonos com pontas protuberantes ou “vazios” internos.

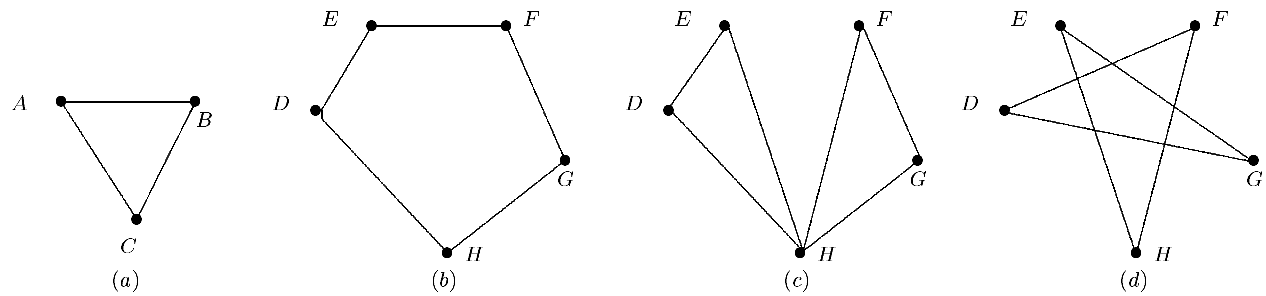


Figura 2.1: Definição de polígonos a partir de seus vértices

A consideração de convexidade está presente na definição de poliedros (definição 2.1) implicitamente. Como pré-requisito para caracterização de poliedros abordaremos na próxima secção o conceito e propriedades básicas de conjuntos convexos.

2.2 Conjuntos Convexos e Cones

Definição 2.2 *Um conjunto $A \subset \mathbb{R}^n$ é dito* **convexo** *se*

$$\forall(\lambda, \boldsymbol{x}^1, \boldsymbol{x}^2) \in [0,1] \times A \times A \Rightarrow (\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) \in A.$$

Em termos geométricos, a definição de conjunto convexo equivale a dizer que qualquer segmento de reta com extremos no conjunto, está contido no conjunto. Seja por esta visão, ou por simples inspecção, é trivial que o qualificador $(\forall\lambda \in [0,1])$ pode ser substituido indiscriminadamente por “$\forall\lambda \in (0,1)$”, “$\forall\lambda \in [0,1)$” ou “$\forall\lambda \in (0,1]$” sem que a definição seja afetada.

Os seguintes conjuntos são exemplos de conjuntos convexos:

(i) ϕ

(ii) $\mathbb{R}^n$

(iii) Subespaços lineares

(iv) Hiperplanos

(v) Semiespaços

(vi) Conjunto unitário[2]

(vii) $X = \{\boldsymbol{x} \in \mathbb{R}^n / \ \mathbf{A}\boldsymbol{x} = \boldsymbol{b} \ e \ \boldsymbol{x} \geqq \mathbf{0}\}$

(viii) A bola (aberta ou fechada) de centro $\bar{\boldsymbol{x}}$ ($\bar{\boldsymbol{x}} \in \mathbb{R}^n$) e raio r ($r > 0$), denotada $B(\bar{\boldsymbol{x}}, r)$

Todos estes exemplos são de fácil verificação ou até mesmo uma afirmação é caso particular de outra (por exemplo, (iii) implica (i) e (ii)). Por utilizarem argumentos usuais e para exemplificar provas derivadas diretamente da definição, provaremos os exemplos (iii), (vii) e (viii).

(iii) Todo subespaço linear L é convexo.

Como L é fechado em relação à multiplicação por escalar e à soma,

[2]Os únicos conjuntos convexos não vazios e enumeráveis são os conjuntos unitários, i.é, aqueles formados por um único elemento.

$$\forall(\lambda, \boldsymbol{x}^1, \boldsymbol{x}^2) \in \mathbb{R} \times L \times L, \text{ temos que } (\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) \in L.$$

Particularizando para $\lambda \in [0,1]$, segue que L é convexo. ■

(vii) $X = \{\boldsymbol{x} \in \mathbb{R}^n / \ \mathbf{A}\boldsymbol{x} = \boldsymbol{b} \text{ e } \boldsymbol{x} \geqq \mathbf{0}\}$ é convexo.

Sejam $(\boldsymbol{x}^1, \boldsymbol{x}^2) \in X \times X \ \ e \ \ \lambda \in [0,1]$.

$$\left.\begin{array}{rcl} \boldsymbol{x}^1 \geqq 0 \text{ e } \lambda & \geqq 0 \Longrightarrow & \lambda \boldsymbol{x}^1 \geqq 0 \\ \boldsymbol{x}^2 \geqq 0 \ e\ (1-\lambda) & \geqq 0 \Longrightarrow & (1-\lambda)\boldsymbol{x}^2 \geqq 0 \end{array}\right\} \quad \lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2 \geqq 0 \qquad (a)$$

Mais ainda

$$\begin{aligned} \mathbf{A}(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) &= \lambda \mathbf{A}\boldsymbol{x}^1 + (1-\lambda)\mathbf{A}\boldsymbol{x}^2 && \text{(por linearidade)} \\ &= \lambda \boldsymbol{b} + (1-\lambda)\boldsymbol{b} && (\text{pois } \boldsymbol{x}^i \in X \Longrightarrow \mathbf{A}\boldsymbol{x}^i = \boldsymbol{b}, i = 1,2) \\ &= \boldsymbol{b} && (c) \end{aligned}$$

De (a) e (c) segue que X é convexo. ■

(viii) $B(\overline{\boldsymbol{x}}, r) = \{\boldsymbol{x} \in \mathbb{R}^n \ / \ \|\overline{\boldsymbol{x}} - \boldsymbol{x}\| \leqq r\}$ é convexa.

Sejam $(\boldsymbol{x}^1, \boldsymbol{x}^2) \in B(\overline{\boldsymbol{x}}, r) \times B(\overline{\boldsymbol{x}}, r) \ \ e \ \ \lambda \in [0,1]$

$$\begin{aligned} \|\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2 - \overline{\boldsymbol{x}}\| &= \|\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2 - \lambda\overline{\boldsymbol{x}} - (1-\lambda)\overline{\boldsymbol{x}}\| \\ &\leqq \|\lambda(\boldsymbol{x}^1 - \overline{\boldsymbol{x}})\| + \|(1-\lambda)(\boldsymbol{x}^2 - \overline{\boldsymbol{x}})\| \\ &\leqq \|\lambda\| r + \|1-\lambda\| r \\ &= r \end{aligned}$$

portanto $B(\overline{\boldsymbol{x}}, r)$ é convexa. ■

Uma forma mais simples de derivar resultados e provar convexidade de conjuntos é usar propriedades de convexos. Vejamos algumas destas propriedades:

Propriedade 1 *Intersecção de conjuntos convexos é um conjunto convexo.*

Prova Seja a família de conjuntos $\{A^i\}_{i \in I}$, tal que, A^i é convexo, para todo $i \in I$. Assim,

$\forall(\boldsymbol{x}^1, \boldsymbol{x}^2) \in D \times D, \ D := \bigcap_{i \in I} A^i, \ \forall\lambda \in [0,1]$ segue que

$\forall i \in I, (\boldsymbol{x}^1, \boldsymbol{x}^2) \in A^i$, e pela convexidade de A^i, $(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) \in A^i$.

Portanto

$(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) \in D = \bigcap_{i \in I} A^i$,

isto é, D é convexo. ■

Exercício 2.1 *Usando a propriedade 1, prove que* $C = \{\boldsymbol{x} \in \mathbb{R}^n \|x - \boldsymbol{x}^0\| \leqq r\}$ *é convexo.*

Desta propriedade e da convexidade de hiperplanos e semiespaços segue que <u>todo poliedro é convexo</u>, em particular o apresentado no exemplo (vii).

É importante notar que:

1) A e B convexos $\not\Rightarrow A \bigcup B$ convexo.
   Basta tomar $A = \{\overline{\boldsymbol{x}}\}$ e $B = \{\overline{\boldsymbol{y}}\}$ com $\overline{\boldsymbol{x}} \neq \overline{\boldsymbol{y}}$.

2) A convexo $\not\Rightarrow$ complemento de A é convexo.
   Basta tomar $A = \{\mathbf{0}\}$.

Propriedade 2 *Combinação linear de conjuntos convexos é um conjunto convexo.*

Sem perda de generalidade, provar a propriedade 2 é equivalente a provar que para quaisquer subconjuntos convexos de $\mathbb{R}^n$ (A e B) e para quaisquer escalares (α e β)

$$C = (\alpha A + \beta B) \text{ é um conjunto convexo, onde}$$
$$(\alpha A + \beta B) := \{\boldsymbol{z} \;/\; \boldsymbol{z} = \alpha \boldsymbol{a} + \beta \boldsymbol{b},\ \boldsymbol{a} \in A \text{ e } \boldsymbol{b} \in B\}.$$

Prova Sejam $(\boldsymbol{x}^1, \boldsymbol{x}^2) \in C \times C$ e $\lambda \in [0,1]$.

Então, como

$\boldsymbol{x}^i = \alpha \boldsymbol{a}^i + \beta \boldsymbol{b}^i$, com $\boldsymbol{a}^i \in A$ *e* $\boldsymbol{b}^i \in B,\ i = 1, 2,$

temos que

$$\begin{aligned}\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2 &= \lambda(\alpha \boldsymbol{a}^1 + \beta \boldsymbol{b}^1) \ +\ (1-\lambda)(\alpha \boldsymbol{a}^2 + \beta \boldsymbol{b}^2)\\ &= \alpha(\lambda \boldsymbol{a}^1 + (1-\lambda)\boldsymbol{a}^2) \ +\ \beta(\lambda \boldsymbol{b}^1 + (1-\lambda)\boldsymbol{b}^2)\\ &= (\alpha v + \beta w) \ \in\ C\end{aligned}$$

onde $\boldsymbol{v} \in A$ e $\boldsymbol{w} \in B$ pela convexidade de A e B .

Portanto, $C = \alpha A + \beta B$ é convexo. ∎

Esta propriedade admite duas particularizações de grande importância.

Propriedade 3 *Soma ou subtração* ($A - B = A + (-1)B$) *de conjuntos convexos é um conjunto convexo.*

Propriedade 4 *Convexidade mantem-se com translação.*
(Uma translação de A é o conjunto $A + \{\overline{\boldsymbol{x}}\}$.)

Outras propriedades interessantes cuja prova fica a cargo do leitor são:

Propriedade 5 *Produto cartesiano de conjuntos convexos é um conjunto convexo.*

Propriedade 6 *A imagem de uma transformação linear L, $L : \mathbb{R}^n \to \mathbb{R}^m$, aplicada a um conjunto convexo $A \subseteq \mathbb{R}^n$, $L(A)$, é convexa.*

Propriedade 7 *Se A é convexo então o interior de A é convexo.*

Propriedade 8 *Se A é convexo fechado com interior não vazio então A coincide com o fecho do seu interior.*

Uma particular classe de conjuntos convexos que serão úteis em nosso estudo é a dos cones convexos. Inicialmente definamos cones:

Definição 2.3 *$B \subset \mathbb{R}^n$ é dito ser um* **cone** *se, e somente se for fechado em relação à multiplicação por escalar não negativo, i.é, $\forall \lambda \geqq 0,\ \forall b \in B,\ (\lambda b) \in B$.*

Em termos geométricos, um conjunto é dito cone se qualquer semi reta com inicio na origem e passando por algum ponto do cone, pertence ao cone.

Exemplos e propriedades de cones

(i) $\mathbb{R}^n$, 0, $\emptyset$ são cones triviais;

(ii) Todo subespaço linear é um cone;

(iii) O hiperplano H é um cone $\Longleftrightarrow$ o hiperplano H contém a origem.

(iv) $C = \{\boldsymbol{h} \in \mathbb{R}^n \ / \ \mathbf{A}\boldsymbol{h} = \mathbf{0} \ e \ h \geqq 0\}$ é um cone;

(v) Duas retas distintas que se interceptam na origem é um cone;

(vi) B é um cone limitado $\Longleftrightarrow B \equiv \{0\}$ ou $B \equiv \emptyset$;

(vii) Intersecção e união de cones é cone;

(viii) Soma de cones é cone;

(ix) Origem pertence ao cone se cone é não vazio;

(x) Translação de cones não necessariamente é cone;

(xi) Nem todo cone é convexo;

(xii) A união do complemento de um cone com a origem é um cone;

(xiii) Qualquer cone é ilimitado se contiver mais que um elemento.

Observemos que o fato de duas retas que se interceptem na origem formarem um cone (exemplo (v)), não corresponde a intuição usual de cone, baseada em cones gerados por conjuntos convexos.

Definição 2.4 *Dado $A \subset \mathbb{R}^n$, A não vazio, definimos* **cone gerado** *por A, denotado $C(A)$, como*

$$C(A) := \{\boldsymbol{x} \in \mathbb{R}^n \ / \ \exists(\boldsymbol{a}, \alpha) \in A \times \mathbb{R}_+ \ : \ \boldsymbol{x} = \alpha\boldsymbol{a}\} \ .$$

Em termos geométricos, o cone gerado por A é a união de todas as semi retas iniciando na origem e passando por algum ponto de A. Se A for um conjunto convexo, vale a propriedade:

Propriedade 9 *A convexo $\Longrightarrow C(A)$ é um cone convexo.*

Prova Seja $C := C(A)$. Devemos demonstrar que: $\forall(\boldsymbol{c}^1, \boldsymbol{c}^2) \in C \times C$, $\lambda_1 \in (0,1)$ e $\lambda_2 \in (0,1)$, para $\lambda_1 + \lambda_2 = 1$[3], teremos $\boldsymbol{c}_\lambda := \lambda_1\boldsymbol{c}^1 + \lambda_2\boldsymbol{c}^2 \in C$. Também suporemos $\boldsymbol{c}^1 \neq \boldsymbol{c}^2$ (caso a igualdade ocorra segue diretamente o resultado).

Como C é gerado por A, devem existir $(\boldsymbol{a}^1, \boldsymbol{a}^2) \in A \times A$ e $(\alpha_1, \alpha_2) \in \mathbb{R}_+ \times \mathbb{R}_+$ para os quais:

$$\boldsymbol{c}^1 = \alpha_1\boldsymbol{a}^1, \tag{2.1}$$
$$\boldsymbol{c}^2 = \alpha_2\boldsymbol{a}^2. \tag{2.2}$$

Como $\lambda_1 \neq 0$ e $\lambda_2 \neq 0$, sejam

$$\beta_1 := \frac{\alpha_1}{\lambda_1\alpha_1 + \lambda_2\alpha_2} \geqq 0 \ \text{ e } \ \beta_2 := \frac{\alpha_2}{\lambda_1\alpha_1 + \lambda_2\alpha_2} \geqq 0. \tag{2.3}$$

Assim, tomando $\boldsymbol{c} := \lambda_1\beta_1\boldsymbol{a}^1 + \lambda_2\beta_2\boldsymbol{a}^2$, segue que

$$\boldsymbol{c} \in A, \tag{2.4}$$

[3] Os casos $\lambda_1 = 0$ e $\lambda_2 = 0$ são triviais.

pois, $\lambda_1\beta_1 + \lambda_2\beta_2 \overset{(2.3)}{=} \lambda_1 \frac{\alpha_1}{\lambda_1\alpha_1+\lambda_2\alpha_2} + \lambda_2 \frac{\alpha_2}{\lambda_1\alpha_1+\lambda_2\alpha_2} = 1$ e A é convexo por hipótese.

Além disso,

$$\begin{aligned} \boldsymbol{c}_\lambda &= \lambda_1\boldsymbol{c}^1 + \lambda_2\boldsymbol{c}^2 \overset{(2.1)}{=} \lambda_1\alpha_1\boldsymbol{a}^1 + \lambda_2\boldsymbol{c}^2 \overset{(2.2)}{=} \lambda_1\alpha_1\boldsymbol{a}^1 + \lambda_2\alpha_2\boldsymbol{a}^2 \\ &\overset{(2.3)}{=} (\lambda_1\alpha_1 + \lambda_2\alpha_2)(\lambda_1\beta_1\boldsymbol{a}^1 + \lambda_2\beta_2\boldsymbol{a}^2) = (\lambda_1\alpha_1 + \lambda_2\alpha_2)\boldsymbol{c}. \end{aligned}$$

Portanto, de (2.4) e da definição de cone, segue a tese. ■

Notemos que a convexidade de $C(A)$ não implica na convexidade de A.

Exercício 2.2 *Verifique através de um contra-exemplo que $C(A)$ convexo $\not\Longrightarrow$ A convexo.*

Exercício 2.3 *Se D é um cone convexo não vazio, prove que $D \equiv C(D \cap B(\mathbf{0}, 1))$*

Propriedade 10 *Se B é um cone, então: B é convexo $\iff$ $B \equiv B + B$.*
(i.é, B é cone convexo $\iff$ B é um cone fechado em relação à soma)

Prova Como $\mathbf{0} \in B$ (para $\lambda = 0$, $\lambda\boldsymbol{b} \in B$) segue que $B \subseteq B + B$. Assim, resta provar que:

$$\langle B \text{ é convexo} \rangle \iff B + B \subseteq B.$$

($\Rightarrow$) Sendo B convexo e cone,

$$\forall(\boldsymbol{b}^1, \boldsymbol{b}^2) \in B \times B \overset{B \text{ é convexo}}{\Longrightarrow} \frac{1}{2}\boldsymbol{b}^1 + \frac{1}{2}\boldsymbol{b}^2 \in B \overset{B \text{ é cone}}{\Longrightarrow} 2(\frac{1}{2}\boldsymbol{b}^1 + \frac{1}{2}\boldsymbol{b}^2) \in B \Longrightarrow \boldsymbol{b}^1 + \boldsymbol{b}^2 \in B,$$

portanto $B + B \subseteq B$

($\Leftarrow$) Sendo $B \equiv B + B$,

$$\forall(\lambda, \boldsymbol{b}^1, \boldsymbol{b}^2) \in [0,1] \times B \times B \overset{B \text{ é cone}}{\Longrightarrow} \lambda\boldsymbol{b}^1 \in B \text{ e } (1-\lambda)\boldsymbol{b}^2 \in B \overset{B+B\subseteq B}{\Longrightarrow} \lambda\boldsymbol{b}^1 + (1-\lambda)\boldsymbol{b}^2 \in B.$$

portanto B é convexo. ■

Cones convexos são conjuntos extremamente semelhantes a subespaços lineares. Subespaços lineares são conjuntos fechados em relação à soma e multiplicação por escalar e cones convexos são conjuntos fechados em relação à soma e multiplicação por escalar não negativo.

Exercício 2.4 *Prove que se $C(A)$ é convexo então $C(A)$ é o conjunto dos pontos que são combinações lineares de elementos de A com coeficientes não negativos.*

Uma categoria que nos interessa de cones convexos é a dos cones poliedrais:

Definição 2.5 *C é um* **cone poliedral** *se, e somente se C é um cone e C é um poliedro.*

Exemplos de cones poliedrais

(i) $\{\mathbf{0}\}, \emptyset, \mathbb{R}^n$, subespaços lineares;

(ii) $\mathbb{R}^n_+ = \{\boldsymbol{x} \in \mathbb{R}^n \;/\; \boldsymbol{x} \geqq \mathbf{0}\}$;

(iii) $C = \{h \in \mathbb{R}^n \;/\; \mathbf{A}\boldsymbol{h} = 0 \text{ e } \boldsymbol{h} \geqq \mathbf{0}\}$.

2.3 Combinações Convexas e Casco Convexo

As definições apresentadas na secção anterior podem ser vistas como generalizações de subespaços lineares. Para tornar mais clara esta observação, repetimos abaixo, de forma suscinta, as caracterizações de subespaço linear, cone convexo e conjunto convexo:

(a) X é subespaço linear

$$\text{se } \left\{ \begin{array}{l} (i)\ \forall (x,y) \in X \times X \\ (ii)\ \forall (\alpha,\beta) \in \mathbb{R} \times \mathbb{R} \end{array} \right\}, \text{ então } (\alpha x + \beta y) \in X$$

(b) X é cone convexo

$$\text{se } \left\{ \begin{array}{l} (i)\ \forall (x,y) \in X \times X \\ (ii)\ \forall (\alpha,\beta) \in \mathbb{R}_+ \times \mathbb{R}_+ \end{array} \right\}, \text{ então } (\alpha x + \beta y) \in X$$

(c) X é conjunto convexo

$$\text{se } \left\{ \begin{array}{l} (i)\ \forall (x,y) \in X \times X \\ (ii)\ \forall (\alpha,\beta) \in \mathbb{R}_+ \times \mathbb{R}_+ \ /\ \alpha + \beta = 1 \end{array} \right\}, \text{ então } (\alpha x + \beta y) \in X$$

Da mesma forma que subespaços lineares são caracterizados por combinação linear de um número finito de elementos, no exercício 2.4 caracterizamos cones convexos através de combinação linear a coeficientes não negativos. Para o estudo de conjuntos convexos, introduzimos o conceito análogo:

Definição 2.6 *Um ponto* $\overline{\boldsymbol{x}}$ *é dito ser* **combinação convexa** [4] *de* $\{\boldsymbol{x}^i\}_{i=1}^p$ *se existem*

$$\left\{ \begin{array}{l} \alpha_i \in \mathbb{R},\ i \in \{1,2,\ldots,p\}, \\ \alpha_i \geqq 0,\ i \in \{1,2,\ldots,p\}, \quad e \quad \sum\limits_{i=1}^{p} \alpha_i = 1 \end{array} \right\} \quad \textit{tais que} \quad \overline{\boldsymbol{x}} = \sum_{i=1}^{p} \alpha_i \boldsymbol{x}^i.$$

Em termos geométricos, $\overline{\boldsymbol{x}}$ é combinação convexa de p pontos $\boldsymbol{x}^i$, se existirem massas m^i que associadas aos correspondentes pontos, definam o centro de gravidade (CG) coincidente com $\overline{\boldsymbol{x}}$. Isto é,

$$CG = \sum_{i=1}^{p} \frac{m^i}{M} \boldsymbol{x}^i = \overline{\boldsymbol{x}} = \sum_{i=1}^{p} \alpha_i \boldsymbol{x}^i$$

onde $M = \sum_{i=1}^{p} m^i$, e $\alpha_i = \frac{m^i}{M}$.

Lembrando que m^i são massas, segue que $\alpha_i \geqq 0$ e $\sum_{i=1}^{p} \alpha_i = 1$.

É conveniente notar que os coeficientes que definem a combinação convexa não são necessariamente únicos.

Utilizando a definição acima podemos caracterizar conjuntos convexos através de:

Teorema 2.1 *A é convexo se e somente se qualquer combinação convexa de pontos de A é um ponto de A, ou mais precisamente*

$$A \textit{ é convexo} \iff \left\{ \left\langle \begin{array}{l} \forall p \in \mathbb{N} \\ \forall \{\boldsymbol{x}^i\}_{i=1}^p,\ \boldsymbol{x}^i \in A,\ i \in \{1,\ldots,p\} \\ \forall \{\alpha_i\}_{i=1}^p,\ \alpha_i \geqq 0 \quad e \quad \sum_{i=1}^{p} \alpha_i \ = 1, \end{array} \right\rangle \Longrightarrow \sum_{i=1}^{p} \alpha_i \boldsymbol{x}^i \ \in \ A. \right.$$

[4] A definição 2.6 implicitamente nos restringe a considerar somente combinação convexa de um número inteiro de pontos.

Prova A prova no sentido ($\Longleftarrow$) é trivial, pois p=2 é a definição de convexidade. A prova no sentido ($\Longrightarrow$) será feita por indução:

Base de indução: (p=1) trivial.

Passo da indução: Suponhamos que qualquer combinação convexa de p (ou menos) elementos de A seja um ponto de A e vamos analisar combinações convexas de $p+1$ elementos de A. Isto é, seja $\bar{\boldsymbol{x}} = \sum_{i=1}^{p+1} \alpha_i \boldsymbol{x}^i$ onde $\boldsymbol{x}^i \in A$, $\alpha_i \geqq 0$ para $i \in \{1, \ldots, p+1\}$ e $\sum_{i=1}^{p+1} \alpha_i = 1$.

(i) Se $\alpha_{p+1} = 1$, então $\alpha_1 = \alpha_2 = ... = \alpha_p = 0$ e $\bar{\boldsymbol{x}} = \boldsymbol{x}^{p+1} \in A$.

(ii) Se $\alpha_{p+1} \in [0,1)$, então, tomando $\beta_i = \frac{1}{1-\alpha_{p+1}} \alpha_i$, segue que

$$\begin{aligned} \bar{\boldsymbol{x}} &= \textstyle\sum_{i=1}^{p} \alpha_i \boldsymbol{x}^i \;+\; \alpha_{p+1} \boldsymbol{x}^{p+1} \\ &= (1-\alpha_{p+1}) \textstyle\sum_{i=1}^{p} \beta_i \boldsymbol{x}^i \;+\; \alpha_{p+1} \boldsymbol{x}^{p+1} \\ &= (1-\alpha_{p+1}) y \;+\; \alpha_{p+1} \boldsymbol{x}^{p+1}. \end{aligned}$$

sendo $y = \sum_{i=1}^{p} \beta_i \boldsymbol{x}^i$. Por construção, y é combinação convexa de $\{\boldsymbol{x}^i\}_{i=1}^{p}$ com pesos $\{\beta_i\}_{i=1}^{p}$, pois:

$$\beta_i = [(1-\alpha_{p+1})^{-1} \alpha_i] \geqq 0, \quad \text{e}$$

$$\begin{aligned} \sum_{i=1}^{p} \beta_i &= [(1-\alpha_{p+1})^{-1} (\sum_{i=1}^{p} \alpha_i)] \\ &= (1-\alpha_{p+1})^{-1} (1-\alpha_{p+1}) \\ &= 1. \end{aligned}$$

Portanto, por hipótese de indução, $y \in A$. Como $\alpha_{p+1} \in [0,1)$, $(y, \boldsymbol{x}^{p+1}) \in A \times A$ e A é convexo, segue que $\bar{\boldsymbol{x}} \in A$, o que completa a demonstração. ■

Uma forma mais elegante de apresentar este resultado pode ser obtida através da introdução do conceito de casco convexo:

Definição 2.7 *O* **casco convexo** *de um conjunto Y, denotado [Y], é o conjunto das combinações convexas dos elementos de Y, i.é:*

$$[Y] \stackrel{\text{def}}{=} \left\{ \boldsymbol{x} \in \mathbb{R}^n \;/\; \left\langle \begin{array}{l} \exists p \in \mathbb{N}, \;\exists \{\boldsymbol{x}^i\}_{i=1}^{p} \subset Y, \;\; \exists \{\alpha_i\}_{i=1}^{p} \subset \mathbb{R} \\ \textit{tais que } \alpha_i \geqq 0, \;\; i \in \{1, \ldots, p\}, \;\; \sum_{i=1}^{p} \alpha_i = 1 \;\;\; e \;\;\; x = \sum_{i=1}^{p} \alpha_i \boldsymbol{x}^i. \end{array} \right\rangle \right\}$$

Em termos geométricos, o casco convexo de um conjunto Y é o lugar geométrico dos centros de gravidade de quaisquer conjuntos finitos de pontos de Y com quaisquer massas associadas.

Da definição de casco convexo de um conjunto, seguem as seguintes propriedades:

(i) $Y \subseteq [Y]$;

(ii) $Z \subseteq Y \implies [Z] \subseteq [Y]$;

(iii) $[Y]$ é convexo;

(iv) Y é convexo $\iff [Y] \subseteq Y \iff Y \equiv [Y]$;

(v) $[\,[Y]\,] = [Y]$;

(vi) Z convexo e $Y \subseteq Z \Longrightarrow [Y] \subseteq Z$.

A figura 2.2 contém alguns exemplos de conjuntos e seus respectivos cascos convexos.

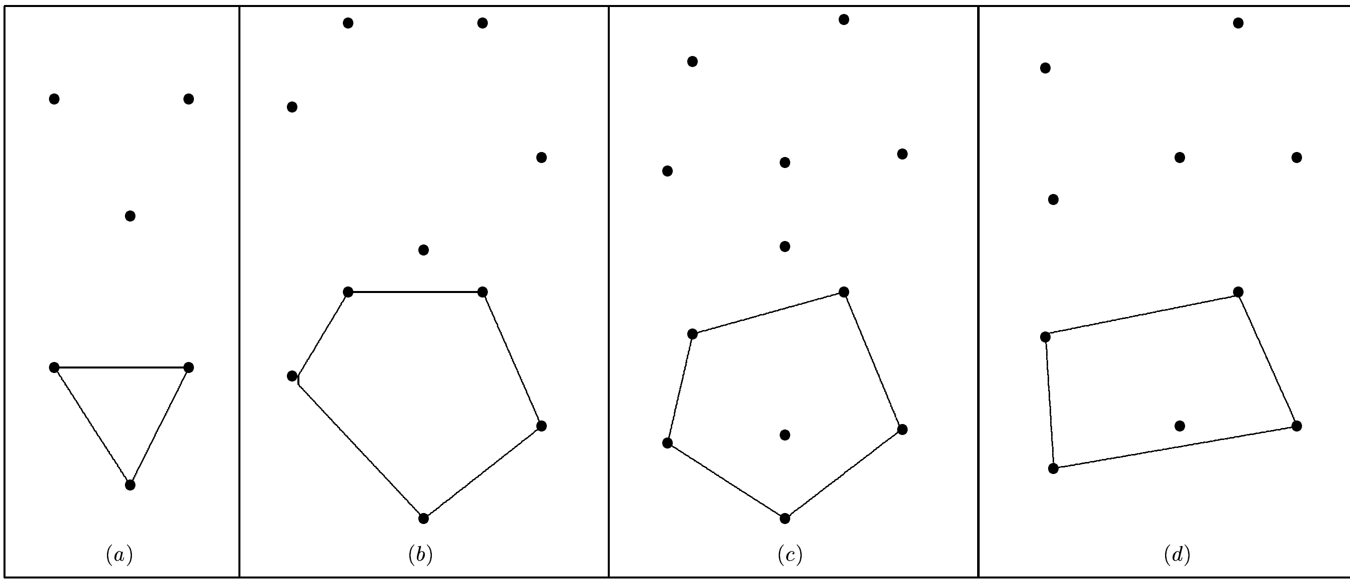


Figura 2.2: Alguns exemplos de geração de casco convexo.

As propriedades (i) e (ii) são consequências imediatas da definição, assim como a propriedade (iii), apesar da prova desta ser menos imediata.

A propriedade (iv) é uma importante caracterização de conjuntos convexos. A propriedade (v) pode ser generalizada como:

(v') $[Y]^p = [Y]$, onde, $[Y]^p$ é definido por:

$$\begin{cases} a)\ [Y]^1 = [Y] \\ b)\ [Y]^{i+1} = [\ [Y]^i\ ]. \end{cases}$$

Tanto (v) quanto (v') tem suas provas baseadas no seguinte argumento:
por (iii), $[Y]$ é convexo, portanto, por (iv), $[[Y]] \equiv [Y]$.

Exercício 2.5 *Prove as propriedades (i) a (vi) e (v').*

Exercício 2.6 *Prove que o casco convexo de um número finito de pontos é um conjunto limitado.*

Exercício 2.7 *Qualquer que seja* $\boldsymbol{c} \in \mathbb{R}^n$, *prove que para o casco convexo de um número finito de pontos,* $A := \{\boldsymbol{a}^1, \boldsymbol{a}^2, \ldots, \boldsymbol{a}^k\}$, $\boldsymbol{c}'\boldsymbol{a} = \sum_{i=1}^{n} c_i a_i$ *é limitado* ($\forall \boldsymbol{a} \in [A]$).

Exercício 2.8 *A recíproca dos dois exercícios anteriores valem ? (se sim, demonstre, caso contrário apresente contra-exemplo).*

A propriedade (vi), de casco convexo, permite uma caracterização alternativa de casco convexo, utilizada como definição por alguns autores, a ser explorada na próxima subseção.

2.3.1 Caracterização de Casco Convexo

O *casco convexo* $[Y]$, de um conjunto Y, é o menor (no sentido de inclusão) convexo que contém Y.

Entenda-se A é menor que B no sentido de inclusão, se $A \subset B$. Portanto as formas precisas de enunciar esta caracterização são:

{$[Y]$ é a intersecção de todos os convexos que contém Y }, ou

{$Y \subseteq A$ e A convexo $\Longrightarrow$ $[Y] \subseteq A$}.

Esta interpretação ou caracterização alternativa justifica os exemplos gráficos da figura 2.2. Note-se que a operação casco convexo dos 5 pontos da figura 2.2.b gera o pentágono como esperado.

Portanto, voltando à idéia de caracterização de polígonos e lembrando que todo poliedro é convexo, a nossa pergunta sobre o mínimo de informação necessária para caracterizar um poliedro pode ser reformulada como: "qual o conjunto minimal M tal que o poliedro convexo X possa ser descrito por $[M]$?".

Claramente tal conjunto minimal deverá conter todos os pontos do conjunto original que não podem ser expressos como combinação convexa de outros pontos do conjunto original. Se verificarmos na figura 2.2, estes pontos são os vértices do triângulo, do pentágono e do quadrilátero.

Por sua importância, estes pontos merecem uma definição:

Definição 2.8 *O ponto a pertencente ao conjunto convexo A, é dito* **ponto extremo** *de A se $\nexists(b,c,\lambda) \in A \times A \times (0,1)$ tal que $a = \lambda b + (1-\lambda)c$ com $c \neq a \neq b$*

ou equivalentemente: a é ponto extremo de A se, e somente se

$$\left.\begin{array}{l} a = \lambda b \ + \ (1-\lambda)c \\ 0 < \lambda < 1 \\ (b,c) \in A \times A \end{array}\right\} \Longrightarrow b \equiv c \equiv a$$

Exemplos

(i) Todo conjunto unitário é formado por um ponto extremo;

(ii) $\{\boldsymbol{x} \in \mathbb{R}^2 / x_1 + x_2 \leqq 1 \ e \ \boldsymbol{x} \geqq \mathbf{0}\}$ tem como pontos extremos $\binom{0}{0}, \binom{1}{0}$ e $\binom{0}{1}$;

(iii) $\mathbb{R}^n_+ = \{\boldsymbol{x} \in \mathbb{R}^n / \boldsymbol{x} \geqq \mathbf{0}\}$ tem um único ponto extremo que é a origem;

(iv) $\{\boldsymbol{x} \in \mathbb{R}^n / \|\boldsymbol{x}\|_2 \leqq 1\}$ tem como pontos extremos $\{\boldsymbol{x} \in \mathbb{R}^n / \|\boldsymbol{x}\|_2 = 1\}$;

(v) $\{\boldsymbol{x} \in \mathbb{R}^2 / \|\boldsymbol{x}\|_1 \leqq 1\}$ tem quatro pontos extremos;

(vi) $\{\boldsymbol{x} \in \mathbb{R}^2 / \ 1 \geqq x_1 \geqq 0\}$ não tem pontos extremos;

(vii) qualquer conjunto aberto não tem pontos extremos.

2.4 Vértices de Poliedros Canônicos

No caso particular de poliedros, usamos o termo **vértice** para designar ponto extremo e, dado um poliedro P, denotamos

$$V(P) = \{\boldsymbol{x} \in P \ / \ \boldsymbol{x} \text{ é vértice de } P\}.$$

Nos exemplos apresentados de polígonos (como da figura 2.2), podemos verificar que o poliedro P (polígono) pode ser representado por $P \equiv [V(P)]$ e que o número de elementos de $V(P)$ (que denotaremos por $\#V(P)$) é finito.

Porém se considerarmos outros poliedros, que não polígonos limitados, a afirmação "$P \equiv [V(P)]$" não é necessariamente verdadeira. Tomando $P = \mathbb{R}^n_+$, a razão da não validade da afirmação é bastante clara, pois $\mathbb{R}^n_+$ é ilimitado e o casco convexo de um número finito de pontos é um conjunto limitado (ver exercício 2.6 e exemplo (iii) acima.).

O exemplo (vi) acima ($P = \{\boldsymbol{x} \in \mathbb{R}^2 \ / \ 1 \geqq x_1 \geqq 0\}$) é mais intrigante, pois $V(P)$ é vazio. Entretanto, se reescrevermos as restrições que descrevem o poliedro P na forma canônica, obteremos o poliedro P_e, dado por:

$$P_e = \{\boldsymbol{z} \in \mathbb{R}^4 \ / \ \boldsymbol{z} \geqq 0 \text{ e } \ z_1 + z_4 = 1\}$$

onde $z_1 = x_1$
$z_2 - z_3 = x_2 \qquad (z_2 = x_2^+ \text{ e } \ z_3 = x_2^+)$
$z_4 = 1 - x_1 \qquad (z_4 = \boldsymbol{x}^r)$.

Para este poliedro, $V(P_e) = \{(0,0,0,1)^T \ e \ (1,0,0,0)^T\}$.

Por este exemplo, procuramos ilustrar que a própria existência de vértices é uma propriedade que pode estar associada à descrição algébrica do poliedro.

Por esta razão e por facilidade de apresentação unificada, doravante estudaremos somente a caracterização de vértices de poliedros correspondentes a conjuntos de pontos viáveis de programas lineares na forma canônica. Isto é, utilizaremos como hipótese desta secção que *os poliedros em consideração serão descritos algebricamente na forma canônica*

$$X = \{\boldsymbol{x} \in \mathbb{R}^n \ / \ \mathbf{A}\boldsymbol{x} = \boldsymbol{b} \text{ e } \ \boldsymbol{x} \geqq \mathbf{0}\}.$$

Para tornar clara tal hipótese, o título da secção é "vértices de poliedros canônicos" como abreviação de "vértices de poliedros correspondentes a conjuntos de pontos viáveis de programas lineares na forma canônica".

Para caracterizar os vértices do poliedro X na forma canônica, uma primeira idéia pode ser verificar o que caracterizaria um ponto do poliedro que não é vértice.

Para um conjunto convexo Y, um ponto de Y não é ponto extremo se e só se ele estiver contido no interior (relativo) de um segmento de reta com extremos no conjunto. Não há perda de generalidade em assumir que $x \in Y$ não é ponto extremo de Y se e só se ele for ponto médio de um segmento (não trivial) [5] cujos extremos são pontos de Y. Isto é,

Fato 2.1 *Dado Y convexo,*

$$\langle \boldsymbol{y} \in Y \ e \ \ \boldsymbol{y} \notin V(Y)\rangle \iff \langle \exists \boldsymbol{h} \in \mathbb{R}^n \ : \ \boldsymbol{h} \neq \mathbf{0} \ e \ \ (\boldsymbol{y}+\boldsymbol{h}, \boldsymbol{y}-\boldsymbol{h}) \in Y \times Y\rangle.$$

Prova

($\Longleftarrow$) Sejam $(\lambda, \boldsymbol{y}^1, \boldsymbol{y}^2) \in]0,1[\times Y \times Y$, definidos por $\lambda := 1/2$, $\boldsymbol{y}^1 := \boldsymbol{y}+\boldsymbol{h}$ e $\boldsymbol{y}^2 := \boldsymbol{y}-\boldsymbol{h}$. Portanto $\boldsymbol{y} \notin V(Y)$, uma vez que $\boldsymbol{y}^1 \neq \boldsymbol{y}^2$, $\lambda \in]0,1[$ e $\boldsymbol{y} = \lambda \boldsymbol{y}^2 + (1-\lambda)\boldsymbol{y}^1$.

($\Longrightarrow$) Sendo

$$\boldsymbol{h}^1 := \boldsymbol{y}^1 - \boldsymbol{y}, \ \boldsymbol{h}^2 := \boldsymbol{y}^2 - \boldsymbol{y} \ \text{ e } \ \boldsymbol{h} \in \arg\min\{\|\boldsymbol{h}^1\|_2, \|\boldsymbol{h}^2\|_2\},$$

mostaremos que $(\boldsymbol{y}+\boldsymbol{h}, \boldsymbol{y}-\boldsymbol{h}) \in Y \times Y$.

Seja $\overline{\lambda} \in]0,1[$ tal que

$$\boldsymbol{y} = \overline{\lambda}\boldsymbol{y}^2 + (1-\overline{\lambda})\boldsymbol{y}^1 = \boldsymbol{y}^1 + \overline{\lambda}(\boldsymbol{y}^2 - \boldsymbol{y}^1). \qquad (\star)$$

Daí segue que

$$\boldsymbol{h}^1 = -\overline{\lambda}(\boldsymbol{y}^2 - \boldsymbol{y}^1), \qquad (\star\star)$$

pois, $\boldsymbol{h}^1 = -\overline{\lambda}(\boldsymbol{y}^2 - \boldsymbol{y}^1) \Leftrightarrow \boldsymbol{y} - \boldsymbol{h}^1 = \boldsymbol{y} + \overline{\lambda}(\boldsymbol{y}^2 - \boldsymbol{y}^1) \Leftrightarrow \boldsymbol{y} - \boldsymbol{h}^1 = \boldsymbol{y}^1 - \boldsymbol{h}^1 + \overline{\lambda}(\boldsymbol{y}^2 - \boldsymbol{y}^1) \Leftrightarrow \boldsymbol{y} = \boldsymbol{y}^1 + \overline{\lambda}(\boldsymbol{y}^2 - \boldsymbol{y}^1)$.

Sem perda de generalidade, suporemos $\boldsymbol{h} = \boldsymbol{h}^1$. Neste caso, é fácil ver que $\overline{\lambda} \in]0, 1/2]$, pois, considerando a direção $\boldsymbol{y}^2 - \boldsymbol{y}^1$, $\boldsymbol{y}$ se encontra antes do ponto médio do segmento $[\boldsymbol{y}^1, \boldsymbol{y}^2]$.

Deste modo, tomando $\lambda := 2\overline{\lambda}$, teremos $\lambda \in]0,1]$ e da convexidade de Y (usada na última identidade)

$$\boldsymbol{y} - \boldsymbol{h}^1 \overset{(\star\star)}{=} \boldsymbol{y} + \overline{\lambda}(\boldsymbol{y}^2 - \boldsymbol{y}^1) \overset{(\star)}{=} \boldsymbol{y}^1 + 2\overline{\lambda}(\boldsymbol{y}^2 - \boldsymbol{y}^1) = \boldsymbol{y}^1 + \lambda(\boldsymbol{y}^2 - \boldsymbol{y}^1) \in Y.$$

[5] Um segmento é dito não trivial quando seus extremos são distintos.

Como $\boldsymbol{y} + \boldsymbol{h}^1 = \boldsymbol{y}^1$ e $\boldsymbol{h} = \boldsymbol{h}^1$, segue a tese $(\boldsymbol{y} + \boldsymbol{h}^1, \boldsymbol{y} - \boldsymbol{h}^1) \in Y \times Y$. ■

O fato 2.1 substitui a caracterização de pontos extremos em termos de intervalos pela existência de direções que definem os mesmos intervalos. A prova deste fato é imediata.

Esta caracterização, em termos de direções, será doravante utilizada. Para o caso em estudo (poliedros na forma canônica), estas direções possuem a seguinte caracterização algébrica:

Fato 2.2 *Para* $\boldsymbol{x} \in X$, $X := \{\boldsymbol{x} \in \mathbb{R}^n \ / \ \mathbf{A}\boldsymbol{x} = \boldsymbol{b} \ e \ \boldsymbol{x} \geqq \mathbf{0}\}$,

$$\{\exists \overline{\lambda} > 0 \ /(\boldsymbol{x} + \overline{\lambda}\boldsymbol{h}, x - \overline{\lambda}\boldsymbol{h}) \in X \times X\} \iff \{\boldsymbol{x} \in X, \ \mathbf{A}\boldsymbol{h} = \mathbf{0} \quad e \quad \langle x_j = 0 \implies h_j = 0\rangle\}.$$

Prova

$(\Longrightarrow)$ i) $\mathbf{A}(\boldsymbol{x} + \overline{\lambda}\boldsymbol{h}) = \mathbf{A}(\boldsymbol{x} - \overline{\lambda}\boldsymbol{h}) = \boldsymbol{b} \iff \mathbf{A}\boldsymbol{h} = -\mathbf{A}\boldsymbol{h} = \mathbf{0}$ e $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$;

ii) $\left\{ \begin{array}{c} (\boldsymbol{x} + \overline{\lambda}\boldsymbol{h}) \geqq \mathbf{0} \\ (\boldsymbol{x} - \overline{\lambda}\boldsymbol{h}) \geqq \mathbf{0} \end{array} \right\} \implies \boldsymbol{x} \geqq \mathbf{0}$;

iii) $x_j = 0, \ (x + \overline{\lambda}\boldsymbol{h})_j \geqq 0, \ (x - \overline{\lambda}\boldsymbol{h})_j \geqq 0 \implies h_j = 0$.

(Observe que $(x_j = 0 \implies h_j = 0)$ equivale a $(h_j \neq 0 \implies x_j \neq 0)$).

De i), ii) e iii) segue a primeira parte da prova.

$(\Longleftarrow)$ i) Se $\boldsymbol{h} = \mathbf{0}$ o resultado é trivial.

ii) Se $\boldsymbol{h} \neq \mathbf{0}$,

a) $\forall \lambda \in \mathbb{R}, \quad \mathbf{A}(\boldsymbol{x} + \lambda\boldsymbol{h}) = \mathbf{A}\boldsymbol{x} + \lambda\mathbf{A}\boldsymbol{h} = \boldsymbol{b} + \lambda\mathbf{0} = Vb$;

b) $x_j = 0$: $\forall \lambda \in \mathbb{R}, \ (\boldsymbol{x} + \lambda h)_j = x_j + \lambda h_j = 0 + \lambda 0 = 0$;

c) $x_j > 0$ e $h_j > 0$: $(\boldsymbol{x} + \lambda h)_j \geqq 0 \iff \lambda \geqq -\frac{x_j}{h_j}(< 0)$;

d) $x_j > 0$ e $h_j < 0$: $(\boldsymbol{x} + \lambda h)_j \geqq 0 \iff \lambda \leqq -\frac{x_j}{h_j}(> 0)$.

Para $\boldsymbol{h} \neq \mathbf{0}$ é necessário que $\boldsymbol{x} \neq \mathbf{0}$, pois $(\boldsymbol{x} = \mathbf{0} \implies \boldsymbol{h} = \mathbf{0})$.

Portanto definimos

$\lambda_- = \sup_{h_j > 0} \ -\frac{x_j}{h_j}$ (sup ao invés de max para considerar o caso $\boldsymbol{h} \leqq \mathbf{0}$, quando $\lambda_+ = \sup \ \emptyset = -\infty$)

$\lambda_+ = \inf_{h_j < 0} \ -\frac{x_j}{h_j}$ (inf ao invés de min para considerar o caso $\boldsymbol{h} \geqq \mathbf{0}$, quando $\lambda_- = \inf \ \emptyset = +\infty$)

$\overline{\lambda} = min\{-\lambda_-, \lambda_+\} \in \mathbb{R}_{++}$ isto é, $\overline{\lambda} \in \mathbb{R}$ e $\overline{\lambda} > 0$.

Assim, da construção de $\overline{\lambda}$, teremos

$$(\boldsymbol{x} + \overline{\lambda}\boldsymbol{h}, \boldsymbol{x} - \overline{\lambda}\boldsymbol{h}) \in X \times X,$$

o que completa a demonstração. ■

No fato acima, o escalar $\overline{\lambda}$ ($\overline{\lambda} > 0$), não altera a direção no sentido geométrico. Este fator de escala é importante para garantirmos que, dada uma direção conveniente, não "andamos demais" ao longo desta.

Esta observação está ligada a $L = \{(\boldsymbol{x} + \lambda\boldsymbol{h}) \ / \ \lambda \in \mathbb{R}\}$ definir uma reta, se $\boldsymbol{h} \neq \mathbf{0}$. Acompanhando a prova do fato 2.2, temos que a intersecção da reta L com o poliedro X é dada pelo segmento de reta $\{(\boldsymbol{x} + \lambda\boldsymbol{h}) \ / \ \lambda \in [\lambda_-, \lambda_+]\}$, se λ_- e λ_+ forem finitos. Caso contrário, $L \bigcap X$ é uma semi-reta, pois ou λ_- ou λ_+ é finito. Em qualquer caso com $\boldsymbol{h} \neq \mathbf{0}$, $(\boldsymbol{x} + \overline{\lambda}\boldsymbol{h})$ ou $(\boldsymbol{x} - \overline{\lambda}\boldsymbol{h})$ será ponto extremo de $L \bigcap X$.

Um resultado que será utilizado repetidas vezes é que este ponto extremo tem pelo menos uma componente nula a mais do que x. Para formalizar esta constatação, introduzimos a notação:

Notação 2.1 *Para* $\boldsymbol{y} \in \mathbb{R}^n$, $I(\boldsymbol{y}) = \{j \in \{1, 2, ...n\} \ / \ y_j \neq 0\}$. *Ou seja,* $I(\boldsymbol{y})$ *é o conjunto dos índices cujas componentes* y_i *sejam estritamente positivas* ($y_i > 0$).

Esta notação e as observações acima conduzem ao seguinte:

Exercício 2.9 *Prove que*

$$\left.\begin{array}{r} (\boldsymbol{x}, \boldsymbol{h}) \in \mathbb{R}^n \times \mathbb{R}^n \\ \boldsymbol{x} \in X, \quad \boldsymbol{h} \neq \boldsymbol{0} \\ \mathbf{A}\boldsymbol{h} = \boldsymbol{0} \ e \ \ I(\boldsymbol{h}) \subseteq I(\boldsymbol{x}) \end{array}\right\} \Longrightarrow \left\{\begin{array}{l} \exists \overline{\lambda} > 0 \ \textit{tal que}: \\ i) \ (\boldsymbol{y}, \boldsymbol{z}) = (\boldsymbol{x} + \overline{\lambda}\boldsymbol{h}, \boldsymbol{x} - \overline{\lambda}\boldsymbol{h}) \in X \times X \\ ii) \ I(\boldsymbol{y}) \subseteq I(\boldsymbol{x}), \ I(\boldsymbol{z}) \subseteq I(\boldsymbol{x}) \ \ e \\ \quad \#I(\boldsymbol{x}) > \min\{\#I(\boldsymbol{y}), \#I(\boldsymbol{z})\}. \end{array}\right.$$

Uma ilustração numérica do exercício 2.9 é:

Exercício 2.10 *Seja* $\overline{\boldsymbol{x}} = (1, 1, 1, 1, 0)^T \ \in X$, $\overline{\boldsymbol{h}} = (1, -5, -1, 5, 0)^T$ *e*

$$X = \left\{ \ x \in \mathbb{R}^5 \ / \ \begin{bmatrix} 1 & 0 & 1 & 0 & -10 \\ 0 & 1 & 0 & 1 & 5 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \ \boldsymbol{x} \geqq \boldsymbol{0} \ \right\}.$$

i) verifique que $\mathbf{A}\overline{\boldsymbol{h}} = \boldsymbol{0}$ *e* $I(\overline{\boldsymbol{h}}) \subset I(\overline{\boldsymbol{x}})$.

ii) Calcule a intersecção da reta L *com o poliedro* X *onde*

$L = \{\boldsymbol{x} \in \mathbb{R}^5 \ / \ \boldsymbol{x} = (1, 1, 1, 1, 0)^T + \lambda(1, -5, -1, 5, 0)^T, \ \lambda \in \mathbb{R}\}$.

iii) Calcule $\boldsymbol{y} \in L \bigcap X$ *tal que* $\#I(\boldsymbol{y}) < \#I(\overline{x})$ *e verifique que* $\boldsymbol{y} = \overline{\boldsymbol{x}} + \overline{\lambda \boldsymbol{h}}$

($\overline{\lambda}$ *como na prova do fato 2.2).*

iv) Mostre que o maior segmento de reta, simétrico em relação a $\overline{\boldsymbol{x}}$ *e contido em* $L \bigcap X$ *é*

$\{\boldsymbol{x} \in \mathbb{R}^5 \ / \ \boldsymbol{x} = (\overline{\boldsymbol{x}} + \lambda\overline{\boldsymbol{h}}) : \ \lambda \in [\lambda_-, \lambda_+]\}$

Estas construções baseadas nos fatos 2.1 e 2.2, são dirigidas à caracterização de pontos do poliedro que não são vértices (ou equivalentemente à vértices). Na medida em que esta caracterização é baseada na existência (ou não) de direções convenientes, introduzimos a notação:

Notação 2.2 $H_x = \{\boldsymbol{h} \in \mathbb{R}^n \ / \ \mathbf{A}\boldsymbol{h} = \boldsymbol{0} \ e \ \ I(\boldsymbol{h}) \subset I(\boldsymbol{x})\}$[6].

Exercício 2.11 *Prove que* H_x *é um subespaço linear e portanto um cone convexo* ($\boldsymbol{x} \in X := \{\boldsymbol{x}\mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geqq \boldsymbol{0}\}$).

Podemos então enunciar o seguinte resultado,

Fato 2.3 *Para todo* $\boldsymbol{x} \in X$, $\boldsymbol{x} \in V(X) \iff H_x \equiv \{\boldsymbol{0}\}$, *que equivale a forma,*

$$\boldsymbol{x} \notin V(X) \iff H_x \not\equiv \{\boldsymbol{0}\}.$$

Prova Dado $\boldsymbol{x} \in X$

$$\begin{array}{ll} \boldsymbol{x} \notin V(X) & \iff \exists \overline{\boldsymbol{h}} \in \mathbb{R}^n \ / \ \overline{\boldsymbol{h}} \neq \boldsymbol{0} \ \text{ e } \ (\boldsymbol{x} + \overline{\boldsymbol{h}}, \boldsymbol{x} - \overline{\boldsymbol{h}}) \in X \times X \quad \text{(pelo fato 2.1)} \\ & \iff \exists \overline{\boldsymbol{h}} \neq \boldsymbol{0} \ / \exists \overline{\lambda} > 0, \ (\boldsymbol{x} + \overline{\lambda \boldsymbol{h}}, \boldsymbol{x} - \overline{\lambda \boldsymbol{h}}) \in X \times X \quad (\overline{\lambda \boldsymbol{h}} = \boldsymbol{h}) \\ & \iff \exists \boldsymbol{h} \neq \boldsymbol{0} \ / \mathbf{A}\boldsymbol{h} = \boldsymbol{0} \ \text{ e } \ \langle x_j = 0 \implies h_j = 0 \rangle \quad \text{(pelo fato 2.2)} \\ & \iff \exists \boldsymbol{h} \neq \boldsymbol{0} \ / \boldsymbol{h} \in H_x. \end{array}$$

Ou seja, $\boldsymbol{x} \in V(X) \iff H_x = \{\boldsymbol{0}\}$. ∎

[6] $H_x = \{\boldsymbol{h} \in N(\mathbf{A}) \ / \ I(\boldsymbol{h}) \subset I(\boldsymbol{x})\}$, onde $N(\mathbf{A})$ é o espaço nulo de $\mathbf{A}$.

2.4.1 Caracterização de Vértices

Lembrando que $\mathbf{A}\boldsymbol{h} = \sum_{j=1}^{n} h_j \boldsymbol{a}^j = \sum_{j \in I(\boldsymbol{h})} h_j \boldsymbol{a}^j$, é fácil verificar o seguinte encadeamento de caracterizações equivalentes[7],

Teorema 2.2 *Dado* $\boldsymbol{x} \in X$,

$$\begin{aligned} \boldsymbol{x} \in V(X) \quad & \Longleftrightarrow \quad H_x \equiv \{\mathbf{0}\} \\ & \Longleftrightarrow \nexists \boldsymbol{h} \in \mathbb{R}^n \ / \ \boldsymbol{h} \neq \mathbf{0}, \quad \textstyle\sum_{j \in I(\boldsymbol{x})} h_j \boldsymbol{a}^j = \mathbf{0} \\ & \Longleftrightarrow \{\boldsymbol{a}^j\}_{j \in I(\boldsymbol{x})} \textit{ são linearmente independentes.} \end{aligned}$$

Prova Imediata do resultado anterior e da definição de dependência linear. ■

Este encadeamento prova um resultado básico que é chamado de **caracterização de vértices por colunas linearmente independentes** , que pode ser apresentado resumidamente como: dado $\boldsymbol{x} \in X = \{\boldsymbol{x} \in \mathbb{R}^n \ / \ \mathbf{A}\boldsymbol{x} = \boldsymbol{b} \text{ e } \boldsymbol{x} \geqq \mathbf{0}\}$

$$\boldsymbol{x} \in V(X) \quad \Longleftrightarrow \quad \{\boldsymbol{a}^j\}_{j \in I(\boldsymbol{x})} \text{ são l.i. (linearmente independentes).}$$

Esta caracterização, básica para o desenvolvimento que se segue, tem como um de seus vários corolários, o abaixo:

Fato 2.4 *O poliedro* $X = \{\boldsymbol{x} \in \mathbb{R}^n \ / \ \mathbf{A}\boldsymbol{x} = \boldsymbol{b} \ \text{ e } \ \boldsymbol{x} \geqq \mathbf{0}\}$ *tem um número finito de vértices.*

Prova Basta apresentar um limitante superior finito, gerado por todos os possíveis vértices,

$$\begin{aligned} \#V(X) \quad & \leqq \#\{J \subseteq \{1,2,\ldots,n\} / \{\boldsymbol{a}^j\}_{j \in J} \text{ é l.i.}\} \\ & \leqq \#\{J \subseteq \{1,2,\ldots,n\} / \#J \leqq m\} \\ & \leqq \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{m}. \end{aligned}$$

■

Exercício 2.12 *Se* $\mathbf{A}$ *tem característica* p ($p \leqq m$) *então* $\#V(X) \leqq \binom{n}{p}$.

Exercício 2.13 *Prove as equivalências:* $\langle \mathbf{0} \in X \rangle \iff \langle \mathbf{0} \in V(X) \rangle \iff \langle \{\mathbf{0}\} \equiv V(X) \rangle$.

O resultado acima utiliza a hipótese do poliedro X estar na forma canônica, isto é, $X = \{\boldsymbol{x} \in \mathbb{R}^n \ / \ \mathbf{A}\boldsymbol{x} = \boldsymbol{b} \text{ e } \boldsymbol{x} \geqq \mathbf{0}\}$.

Considerando novamente o exemplo $Y = \{x \in \mathbb{R}^2 \ / \ 1 \geqq x_1 \geqq 0\}$, para o qual $V(Y) \equiv \emptyset$, é fácil ver que $0 \in Y$ o que mostra que a validade do resultado apresentada no exercício 2.13 depende da representação algébrica do poliedro.

Uma das razões para escolhermos a forma canônica é a existência de resultados fortes ligando o poliedro nesta forma ao conjunto de seus vértices. Por exemplo:

Fato 2.5 *Dado* $X = \{\boldsymbol{x} \in \mathbb{R}^n \ / \ \mathbf{A}\boldsymbol{x} = \boldsymbol{b} \ \text{ e } \ \boldsymbol{x} \geqq \mathbf{0}\}$

$$X \neq \emptyset \iff V(X) \neq \emptyset.$$

Prova

[7] Vale lembrar que $\boldsymbol{a}^j$ é a j-ésima coluna da matriz $\mathbf{A}$.

($\Longleftarrow$) Trivial pois $V(X) \subseteq X$.

($\Longrightarrow$) Se $X \not\equiv \emptyset$, então (pelo axioma da escolha), seja $\boldsymbol{x}^0 \in X \subseteq \mathbb{R}^n$.

Podemos gerar a seguinte sequência $\{\boldsymbol{x}^i\}_{i=0}^{n+1}$:

i) se $\boldsymbol{x}^i \in V(X)$, $\boldsymbol{x}^{i+1} = \boldsymbol{x}^i$;

ii) se $\boldsymbol{x}^i \notin V(X)$, $\boldsymbol{x}^{i+1} = \boldsymbol{y}^i \in X \;:\; \#I(\boldsymbol{y}^i) \;<\; \#I(\boldsymbol{x}^i)$.

A existência de $\boldsymbol{y}^i$ como em ii) é garantida pelo exercício 2.9 (página 14).

Claramente $\#I(\boldsymbol{x}^0) \leqq n$ e, $\boldsymbol{x}^i \in V(X)$ ou $\#I(\boldsymbol{x}^i) \leqq n - i$.

Como $\#I(\boldsymbol{x}^{n+1}) \geqq 0$ temos que $\boldsymbol{x}^{n+1} \in V(X)$. ■

A prova construtiva apresentada acima é algorítmica, no sentido que a partir de um ponto do poliedro é gerada uma sequência de pontos, também do poliedro. Após no máximo n iterações, esta sequência produz um vértice.

Se $\boldsymbol{x}^i$ não for vértice, o conjunto de colunas associadas às componentes não nulas de $\boldsymbol{x}^i$ ($\{\boldsymbol{a}^j\}_{j \in I(\boldsymbol{x}^i)}$) é linearmente dependente. Em cada iteração do processo é gerado $\boldsymbol{x}^{i+1}$ com menos componentes não nulas do que $\boldsymbol{x}^i$ ($I(\boldsymbol{x}^{i+1})$ estritamente incluso em $I(\boldsymbol{x}^i)$). Portanto a cada iteração ou detetamos um vértice ou reduzimos o conjunto de colunas $\{\boldsymbol{a}^j\}_{j \in I(\boldsymbol{x}^k)}$, buscando um subconjunto de colunas linearmente independentes, e portanto, correspondente à um vértice.

2.4.2 Método Conceitual para Obtenção de Vértice (MCOV)

Esta visão e os resultados anteriores, sugerem que o problema de encontrar um vértice do poliedro X, a partir de um ponto qualquer deste poliedro, possui uma solução algorítmica simples, como pode ser visto por:

Inicialização: $i \longleftarrow 0, \;\; \boldsymbol{x}^0 \in X$

Passo ($\star$): Encontre, se existir, $\boldsymbol{h}^i$ tal que:

$\boldsymbol{h}^i \neq 0$

$\boldsymbol{h}^i \in H_{\boldsymbol{x}^i} = \{\boldsymbol{h} \in \mathbb{R}^n \, \mathbf{A}\boldsymbol{h} = \mathbf{0} \text{ e } I(\boldsymbol{h}) \subseteq I(\boldsymbol{x}^i)\}$.

Caso não exista, $\boldsymbol{x}^i \in V(X)$: pare.

Se $\exists h^i_j > 0$, faça:

$\lambda_- = \max_{h^i_k > 0} -\frac{x^i_k}{h^i_k}$ $\quad \left(x^i_j + \lambda_- h^i_j = h^i_j (\frac{x^i_j}{h^i_j} + \max\limits_{h^i_k > 0} \frac{-x^i_k}{h^i_k}) = h^i_j (\frac{x^i_j}{h^i_j} - \min\limits_{h^i_k > 0} \frac{x^i_k}{h^i_k}) \geqq 0 \right)$

$\boldsymbol{x}^{i+1} \longleftarrow \boldsymbol{x}^i + \lambda_- \boldsymbol{h}^i$

$i \longleftarrow i + 1$

retorne ao passo ($\star$).

Caso contrário, faça:

$\lambda_+ = \min_{h^i_k < 0} -\frac{x^i_k}{h^i_k}$ $\quad \left(x^i_j + \lambda_+ h^i_j = h^i_j (\frac{x^i_j}{h^i_j} + \min\limits_{h^i_k < 0} \frac{-x^i_k}{h^i_k}) = h^i_j (\frac{x^i_j}{h^i_j} - \max\limits_{h^i_k < 0} \frac{x^i_k}{h^i_k}) \overset{h^i_j < 0}{\geqq} 0 \right)$

$\boldsymbol{x}^{i+1} \longleftarrow \boldsymbol{x}^i + \lambda_+ \boldsymbol{h}^i$

$i \longleftarrow i + 1$

retorne ao passo ($\star$).

■

A transformação deste método conceitual em algoritmo é obtida pela clara discriminação de como construir $\boldsymbol{h}^i \in H_{\boldsymbol{x}^i}$, tal que $\boldsymbol{h}^i \neq \mathbf{0}$. Isto corresponde a encontrar, se existir, uma solução não trivial para o sistema homogêneo de equações lineares:

$$\textstyle\sum_{j \in J} h_j \boldsymbol{a}^j = \mathbf{A}_J \boldsymbol{h}_J = \mathbf{0},$$

onde $J = I(\boldsymbol{x}^i)$.

Exercício 2.14 *Descreva detalhadamente o algoritimo indicado no parágrafo anterior, utilizando a solução apresentada em* **??**.

O método conceitual (MCOV), além de ser a base para a solução de problemas que veremos adiante, será utilizado, na próxima secção, para a construção de caracterização finita de poliedros.

2.5 Caracterização de Poliedros Canônicos

A caracterização finita de poliedros é a resposta à questão levantada anteriormente sobre existência de um conjunto minimal M tal que X possa ser descrito a partir de M através da operação casco convexo. Uma interpretação geométrica de (MCOV) será a base usada para derivar tal caracterização.

A cada iteração do (MCOV), se não detetarmos um vértice, construimos uma direção tal como a utilizada no fato 2.2. Relembrando a interpretação geométrica deste fato, dado um ponto z do poliedro (z não vértice), existe uma reta L cuja intersecção com o poliedro X contém z no seu interior (relativo). Como já discutido anteriormente, a reta L intercepta o poliedro ou em um segmento de reta ou em uma semireta. Os extremos do segmento de reta, ou a origem da semireta, são pontos do poliedro "mais próximos" de serem vértices, pois possuem menos componentes não nulas do que z.

Usando estas direções podemos mostrar que o primeiro conjunto não vazio do tipo $X_q := \{\boldsymbol{x} \in X : \#I(\boldsymbol{x}) \leqq q\}$, com um número minimal de componentes não nulas, está contido no conjunto dos vértices de X.

Lema 2.1 *Se $p = min\{q : X_q \neq \emptyset\}$, então $X_p \subseteq V(X)$.*

Prova Por contradição, suponha existir $\boldsymbol{x} \in X_p$ tal que $\boldsymbol{x} \notin V(X)$. Neste caso existe $\boldsymbol{h} \neq \boldsymbol{0}$ para o qual

$$\mathbf{A}\boldsymbol{h} = \boldsymbol{0} \quad \text{e} \quad (\boldsymbol{x} - \boldsymbol{h}, \boldsymbol{x} + \boldsymbol{h}) \in X \times X.$$

Neste caso existe i tal que $h_i > 0$ (se $\boldsymbol{h} \leq \boldsymbol{0}$, tome $-\boldsymbol{h}$ no lugar de $\boldsymbol{h}$). Assim, se tomarmos $\lambda = \min\limits_{h_j > 0} \frac{x_j}{h_j} > 0$, teremos $\boldsymbol{x} - \lambda\boldsymbol{h} \in X$, pois

- $\mathbf{A}(\boldsymbol{x} - \lambda\boldsymbol{h}) = \mathbf{A}\boldsymbol{x} - \lambda\mathbf{A}\boldsymbol{h} = \mathbf{A}\boldsymbol{x} = \boldsymbol{b}$; e
- $\boldsymbol{x} - \lambda\boldsymbol{h} \geqq \boldsymbol{0}$ $(\lambda \in \mathbb{R}_+)$: o caso $h_k = 0$ é trivial, assim examinaremos
  - $h_k < 0$: $x_k - \lambda h_k > x_k > 0$ (do fato 2.2, $h_k \neq 0 \Rightarrow x_k > 0$); e
  - $h_k > 0$: $x_k - \lambda h_k = x_k - \min\limits_{h_j>0}\left\{\frac{x_j}{h_j}\right\} h_k = h_k \left(\frac{x_k}{h_k} - \min\limits_{h_j>0}\left\{\frac{x_j}{h_j}\right\}\right) \geqq 0.$

Além disso, do fato 2.2, $x_j = 0 \Rightarrow h_j = 0$ e portanto teríamos

$$\#I(\boldsymbol{x} - \lambda\boldsymbol{h}) < \#I(\boldsymbol{x}) = p,$$

mas isso contraria a escolha de p. Logo não pode existir um tal $\boldsymbol{x} \in X_p \setminus V(X)$, concluindo daí a tese. ∎

Agora podemos demonstrar um importante resultado, que poliedros canônicos são somas diretas do casco convexo de um número finito de pontos mais um cone.

Teorema 2.3 *Se $X = \{\boldsymbol{x} \in \mathbb{R}^n / \ \mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ e $\boldsymbol{x} \geqq \boldsymbol{0}\}$ e $C = \{\boldsymbol{h} \in \mathbb{R}^n / \ \mathbf{A}\boldsymbol{h} = \boldsymbol{0}, \ \boldsymbol{h} \geqq \boldsymbol{0}\}$, então*

$$X \ \equiv \ [\ V(X)\ ] \ + \ C.$$

Prova Devemos demonstrar que: (i) $[V(X)] + C \subseteq X$ e (ii) $X \subseteq [V(X)] + C$.

(i) Se $\boldsymbol{x} \in [V(X)] + C$, então existem

$$\boldsymbol{y} = \sum_{i=1}^{k} \alpha_i \boldsymbol{x}^i \in [V(X)] \qquad \left(\sum_{i=1}^{k} \alpha_i = 1, \boldsymbol{\alpha} \geqq \mathbf{0} \text{ e } \{\boldsymbol{x}^i\}_{i=1}^{k} \subseteq V(X) \right) \quad \text{e} \quad \boldsymbol{h} \in C,$$

tais que $\boldsymbol{x} = \boldsymbol{y} + \boldsymbol{h}$. Daí,

- $\mathbf{A}\boldsymbol{x} = \mathbf{A}(\boldsymbol{y} + \boldsymbol{h}) \overset{\mathbf{A}\boldsymbol{h}=\mathbf{0}}{=} \mathbf{A}\boldsymbol{y} = \sum_{i=1}^{k} \alpha_i \mathbf{A}\boldsymbol{x}^i \overset{\mathbf{A}\boldsymbol{x}^i=\boldsymbol{b}}{=} \sum_{i=1}^{k} \alpha_i \boldsymbol{b} = \boldsymbol{b}$;
- $\boldsymbol{x} = \boldsymbol{y} + \boldsymbol{h} \overset{\boldsymbol{h}\geqq\mathbf{0}}{\geqq} \boldsymbol{y} = \sum_{i=1}^{k} \alpha_i \boldsymbol{x} \geqq \mathbf{0}$.

Portanto, $\boldsymbol{x} \in X$.

(ii) Suporemos $X \not\equiv \emptyset$ e $X \not\equiv [V(X)]$, pois estes casos são diretos. Faremos a demonstração por indução no índice q de $X_q := \{\boldsymbol{x} \in X : \#I(\boldsymbol{x}) \leqq q\}$. Seja novamente $p := \min\{q : X_q \not\equiv \emptyset\}$.

Base: Do lema 2.1, vale $X_p \equiv V(X)$ e daí segue que $X_p \subseteq [V(X)] + C$.

Passo: Supor que $X_i \subseteq [V(X)] + C$, para todo $p \leqq i \leqq k$.

Seja $\boldsymbol{x} \in X_{k+1}$. Se $\boldsymbol{x} \in V(X)$ já segue que $\boldsymbol{x} \in [V(X)] + C$ (basta tomar $\boldsymbol{h} = \mathbf{0}$).
Se $\boldsymbol{x} \notin V(X)$, existe $\boldsymbol{h} \neq \mathbf{0}$, para o qual $\mathbf{A}\boldsymbol{h} = 0$, $I(\boldsymbol{h}) \subseteq I(\boldsymbol{x})$ e $(\boldsymbol{x} - \boldsymbol{h}, \boldsymbol{x} + \boldsymbol{h}) \in X \times X$ (fato 2.2), analisaremos separadamente três casos, a partir da seguinte construção:

(1) Se $\boldsymbol{h} \not\geqq \mathbf{0}$, mostraremos que $\boldsymbol{x} + \lambda_1 \boldsymbol{h} \geqq \mathbf{0}$, $\lambda_1 = \min\limits_{h_j<0} \frac{-x_j}{h_j}$.
Se $h_j \geqq 0$, então $x_j + \lambda h_j \geqq x_j \geqq 0$, $\forall \lambda > 0$.
Se $h_j < 0$, então $x_j + \lambda_1 h_j = h_j(\frac{x_j}{h_j} + \min_{h_l<0} \frac{-x_l}{h_l}) = h_j(\frac{x_j}{h_j} - \max_{h_l<0} \frac{x_l}{h_l}) \geqq 0$. Deste modo, $I(\boldsymbol{x} + \lambda_1 \boldsymbol{h}) \subsetneqq I(\boldsymbol{x})$, e por hipótese de indução $\boldsymbol{x} + \lambda_1 \boldsymbol{h} \in [V(X)] + C$.

(2) Se $\boldsymbol{h} \not\leqq \mathbf{0}$, mostraremos que $\boldsymbol{x} - \lambda_2 \boldsymbol{h} \geqq \mathbf{0}$, $\lambda_2 = \min\limits_{h_j>0} \frac{x_j}{h_j}$.
Se $h_j > 0$, então $x_j - \lambda_2 h_j = h_j(\frac{x_j}{h_j} - \min_{h_l>0} \frac{x_l}{h_l}) \geqq 0$.
Se $h_j \leqq 0$, então $x_j - \lambda h_j \geqq x_j \geqq 0$, $\forall \lambda > 0$.
Do mesmo modo, $I(\boldsymbol{x} - \lambda_2 \boldsymbol{h}) \subsetneqq I(\boldsymbol{x})$, e por hipótese de indução $\boldsymbol{x} - \lambda_2 \boldsymbol{h} \in [V(X)] + C$.

Existem três possibilidade,

* Se $\boldsymbol{h} \geq \mathbf{0}$, teremos $\boldsymbol{h} \in C$ e, usando o item (2) acima, $\boldsymbol{y} := \boldsymbol{x} - \lambda_2 \boldsymbol{h} \in [V(X)] + C$. Deste modo,

$$\boldsymbol{x} = \boldsymbol{y} + \lambda_2 \boldsymbol{h} \in [V(X)] + C.$$

* Se $\boldsymbol{h} \leq \mathbf{0}$, teremos $-\boldsymbol{h} \in C$ e, usando o item (1) acima, segue que $\boldsymbol{y} := \boldsymbol{x} + \lambda_1 \boldsymbol{h} \in [V(X)] + C$. Deste modo, $-\lambda_1 \boldsymbol{h} \in C$ e

$$\boldsymbol{x} = \boldsymbol{y} - \lambda_1 \boldsymbol{h} \in [V(X)] + C.$$

* Se $\boldsymbol{h} \not\geqq \mathbf{0}$ e $\boldsymbol{h} \not\leqq \mathbf{0}$, concluimos que

$$\boldsymbol{x} + \lambda_1 \boldsymbol{h} \in [V(X)] + C \quad \text{e} \quad \boldsymbol{x} - \lambda_2 \boldsymbol{h} \in [V(X)] + C,$$

portanto $\boldsymbol{x} \in [\{\boldsymbol{x} - \lambda_2 \boldsymbol{h}, \boldsymbol{x} + \lambda_1 \boldsymbol{h}\}]$, e daí segue que[8]
$\boldsymbol{x} \in [V(X)] + C$. ■

Este teorema tem uma particularização importante:

[8] Note que, se C é cone, $[C] = C$ e $[A + C] = [A] + C$.

Corolário 2.1 *Seja* $X = \{\boldsymbol{x} \in \mathbb{R}^n \ / \ \ \mathbf{A}\boldsymbol{x} = \boldsymbol{b} \ e \ \ \boldsymbol{x} \geqq \mathbf{0}\}$

$$X \text{ é limitado} \iff X \equiv [V(X)].$$

Prova

($\Longrightarrow$) Se $X = \emptyset$, então segue a tese. Se $X \neq \emptyset$ é limitado, então $\nexists \boldsymbol{h} \neq \mathbf{0}$ em C (caso contrário $\boldsymbol{x} + \lambda\boldsymbol{h} \in X$, para todo $\lambda\mathbb{R}_+$ e $\boldsymbol{x} \in X$). Logo $C \equiv \{\mathbf{0}\}$ e daí, pelo teorema acima, segue que $X \equiv [V(X)]$.

($\Longleftarrow$) Esta outra parte é baseada em que o conjunto de vértices é finito (fato 2.4) e que o casco convexo de um número finito de pontos é limitado (exercício 2.6). ■

O fato acima corresponde a enunciar o resultado intuitivo que um poliedro convexo limitado (vazio ou não) é descrito de forma minimal pelo conjunto de seus vértices. Mais ainda, está ligado a constatação de que um poliedro (não vazio) é limitado se, e só se, o cone associado ao poliedro reduz-se à origem.

Exercício 2.15 *Prove que:*

(i) $C \equiv \{\mathbf{0}\} \Longrightarrow X$ *limitado*

(ii) X *limitado não vazio* $\Longrightarrow C \equiv \{\mathbf{0}\}$

(iii) X *limitado* $\not\Longrightarrow C \equiv \{\mathbf{0}\}$

Portanto, o conjunto de vértices (finito) não é informação suficiente para caracterização (finita) de poliedros somente no caso ilimitado ($C \not\equiv \{0\}$).

A idéia básica desta caracterização é que conhecendo o cone poliedral em uma região limitada, o restante do cone poliedral pode ser descrito através de fatores de escala (de forma similar ao exercício 2.3). Ora, se a região limitada for um poliedro, a intersecção do cone poliedral com este poliedro limitado será um poliedro limitado e passível de descrição como casco convexo de um número finito de pontos. Esta idéia intuitiva é formalmente expressa como:

Fato 2.6 *Se* $C = \{\boldsymbol{h} \in \mathbb{R}^n / \ \mathbf{A}\boldsymbol{h} = \mathbf{0} \ \ e \ \ \boldsymbol{h} \geqq \mathbf{0}\}$ *e*

$$C^1 = \{\boldsymbol{h} \in C / \ \|h\|_1 = 1\} \equiv \{\boldsymbol{h} \in \mathbb{R}^n / \ \mathbf{A}\boldsymbol{h} = \mathbf{0}, \ \ \boldsymbol{h} \geqq \mathbf{0} \ \ e \ \ \textstyle\sum_{j=1}^n h_j = 1\},$$

então $\quad C \not\equiv \{0\} \iff C^1 \not\equiv \emptyset$

$\qquad\qquad\qquad\quad \iff C \equiv C(C^1)$

Prova

(i) $C \not\equiv \{\mathbf{0}\} \ \Longrightarrow \ C^1 \not\equiv \emptyset$

Trivial, pois se $\boldsymbol{h} \in C$ e $\boldsymbol{h} \neq 0$, $\|\boldsymbol{h}\|_1^{-1}\boldsymbol{h} \in C^1$.

(ii) $C^1 \not\equiv \emptyset \ \Longrightarrow \ C \equiv C(C^1)$

$C(C^1) \subseteq C$: Como C é cone e $C^1 \subseteq C$, segue que $C(C^1) \subseteq C$.

$C \subseteq C(C^1)$: Tome um $\boldsymbol{h} \in C$ qualquer:

* Se $\boldsymbol{h} = \mathbf{0}$, podemos tomar $\boldsymbol{h} = 0\overline{\boldsymbol{h}}$, $\overline{\boldsymbol{h}} \in C^1 \not\equiv \emptyset$.
* Se $\boldsymbol{h} \neq \mathbf{0}$, definimos $\overline{\boldsymbol{h}} := \frac{1}{\|\boldsymbol{h}\|_1}\boldsymbol{h}$, e daí $\overline{\boldsymbol{h}} \in C^1$.

  Como $C(C^1)$ é o cone gerado por C^1, qualquer que seja $\lambda \geqq 0$, $\lambda\overline{\boldsymbol{h}} \in C(C^1)$. Assim, tomando $\lambda := \|\boldsymbol{h}\|_1$ teremos $\boldsymbol{h} = \lambda\overline{\boldsymbol{h}} \in C(C^1)$.

(iii) $C \equiv C(C^1) \ \Longrightarrow \ C \not\equiv \{\mathbf{0}\}$

Como $\mathbf{0} \in C \not\equiv \emptyset$, $C \equiv C(C^1) \ \Longrightarrow \ C^1 \not\equiv \emptyset$ e para todo $\boldsymbol{h} \in C^1$, $\boldsymbol{h} \in C$ e $\boldsymbol{h} \neq \mathbf{0}$. ■

Observe que, utilizando o fato 2.6, C fica completamente descrito através de C^1. Se $C^1 \equiv \emptyset$ então $C \equiv \{0\}$, caso contrário $C \equiv C(\ C^1 \ )$. Além disso, como C^1 é um poliedro limitado, ele é completamente descrito por seus vértices.

2.5.1 Caracterização finita de poliedros

Definição 2.9 *Os vértices de* C^1 *são chamados* **raios extremais** *de* C.

Exercício 2.16 *Mostre que* $\boldsymbol{h} \in C$, $\boldsymbol{h} = \sum_{i=0}^{q} \beta_i \boldsymbol{r}^i$ *onde* $\beta_i \geqq 0$, $\boldsymbol{r}^0 = \mathbf{0}$ *e* $\{\boldsymbol{r}^i\}_{i=1}^{q} \equiv V(C^1)$.

Este resultado pode ser considerado como uma caracterização de cones poliedrais.

O encadeamento de resultados acima nos permite enunciar a

Caracterização finita de poliedros: Sejam $X = \{\boldsymbol{x} \in \mathbb{R}^n \ / \ \mathbf{A}\boldsymbol{x} = \boldsymbol{b} \text{ e } \boldsymbol{x} \geqq \mathbf{0}\}$ e
$C^1 = \{\boldsymbol{h} \in \mathbb{R}^n \ / \ \mathbf{A}\boldsymbol{h} = \mathbf{0}, \ \sum_{i=1}^{n} h_i = 1 \text{ e } \boldsymbol{h} \geqq \mathbf{0}\}$.

Então

$$\boldsymbol{x} \in X \iff \begin{cases} \boldsymbol{x} = & \sum_{j=1}^{p} \alpha_j \boldsymbol{v}^j \ + \ \sum_{i=0}^{q} \beta_i \boldsymbol{r}^i \\ & \\ \text{onde} & \{\boldsymbol{v}^i\}_{i=1}^{p} \equiv V(X) \\ & \{\boldsymbol{r}^j\}_{j=1}^{q} \equiv V(C^1) \\ & \sum_{j=1}^{p} \alpha_j = 1, \ \alpha_j \geqq 0, \ j \in \{1, 2, \ldots, p\}, \\ & \boldsymbol{r}^0 = \mathbf{0}, \ \beta_0 = 1, \ \beta_j \geqq 0, \ j \in \{1, 2, \ldots, q\} \end{cases}$$

Este resultado é básico para os desenvolvimentos que se seguem, mas não deve ser esquecido que o poliedro tem que estar na forma canônica

$$X = \{\boldsymbol{x} \in \mathbb{R}^n \ / \ \mathbf{A}\boldsymbol{x} = \boldsymbol{b} \ \text{ e } \ \boldsymbol{x} \geqq \mathbf{0}\}$$

Se por exemplo o poliedro fosse:

$X = \{\boldsymbol{x} \in \mathbb{R}^2 \ / \ 1 \geqq x_1 \geqq 0\} \not\equiv \emptyset$

este resultado não seria válido pois $V(X) \equiv \emptyset$.

Capítulo 3

O teorema fundamental de programação linear e algumas implicações em termos de algoritmos

3.1 O Teorema Fundamental de Programação Linear.

Tendo estudado a caracterização de poliedros, o passo seguinte é aplicar este conhecimento ao problema de programação linear. Como discutido anteriormente, sem perda de generalidade nos concentraremos no problema canônico de programação linear:

$$\text{(PLC)} \begin{cases} \max & \langle \boldsymbol{c}, \boldsymbol{x} \rangle \\ s.a & \mathbf{A}\boldsymbol{x} \; = \; \boldsymbol{b} \\ & \boldsymbol{x} \; \geqq \mathbf{0}. \end{cases}$$

Utilizando os resultados de caracterização de poliedros (canônicos), é claro que o (PLC) pode ser escrito como:

$$\text{(PLC)} \begin{cases} \max & \langle \boldsymbol{c}, \boldsymbol{x} \rangle \\ s.a & \boldsymbol{x} \in [V(X)] \; + \; C \\ \text{onde} & X \equiv \{\boldsymbol{x} \in \mathbb{R}^n / \; \mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geqq \mathbf{0}\} \;\; \text{e} \\ & C \equiv \{\boldsymbol{h} \in \mathbb{R}^n / \; \mathbf{A}\boldsymbol{h} = \mathbf{0}, \boldsymbol{h} \geqq \mathbf{0}\}. \end{cases}$$

O primeiro resultado que decorre trivialmente da caracterização acima é que

Fato 3.1 *O (PLC) é inviável* $\iff X \; \equiv \; \emptyset$
$\iff V(X) \; \equiv \; \emptyset$
$\iff$ *valor ótimo do (PLC)* $= -\infty$.

Este fato dispensa prova formal pois não é mais do que uma coletânea de definições equivalentes e do fato 2.5 ($X \equiv \emptyset \iff V(X) \equiv \emptyset$).

Portanto, se nos concentrarmos nos problemas viáveis, estaremos maximizando a função $\langle \boldsymbol{c}, \boldsymbol{x} \rangle$ para $X \in [V(X)] \; + \; C$, onde $V(X) \not\equiv \emptyset$ (PLC viável) e $C \not\equiv \emptyset$ (pois $\mathbf{0} \in C$).

Logo, o valor ótimo do (PLC) (supondo viabilidade) pode ser escrito como

$$\text{valor ótimo do (PLC)} = \sup_{\boldsymbol{x} \in X} \langle \boldsymbol{c}, \boldsymbol{x} \rangle = \sup_{\boldsymbol{x} \in [V(X)]} \langle \boldsymbol{c}, \boldsymbol{x} \rangle + \sup_{\boldsymbol{h} \in C} \langle \boldsymbol{c}, \boldsymbol{h} \rangle.$$

Em relação ao primeiro termo que compõe a expressão acima notemos que:

Fato 3.2 *Para todo* $\boldsymbol{c} \in \mathrm{I\!R}^n$ *e todo* $\boldsymbol{x} \in X$, $X \equiv \{\boldsymbol{x} \in \mathrm{I\!R}^n / \ \mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \ \geqq \mathbf{0}\} \ \not\equiv \ \emptyset$;

$$\sup_{\boldsymbol{x} \in [V(X)]} \langle \boldsymbol{c}, \boldsymbol{x} \rangle = \max_{\boldsymbol{v} \in V(X)} \langle \boldsymbol{c}, \boldsymbol{v} \rangle = \langle \boldsymbol{c}, \overline{\boldsymbol{v}} \rangle$$

onde $\overline{\boldsymbol{v}} \in argmax\{\langle \boldsymbol{c}, \boldsymbol{v} \rangle / \ \boldsymbol{v} \in V(X)\}$ *(i.e.,* $\overline{\boldsymbol{v}} \in V(X)$ *e* $\langle \boldsymbol{c}, \overline{\boldsymbol{v}} \rangle \ \geqq \langle \boldsymbol{c}, \boldsymbol{v} \rangle, \ \forall \boldsymbol{v} \in V(X)$*).*

Prova Como $X \not\equiv \emptyset$, temos que $V(X) \not\equiv \emptyset$ e assim $p := \#V(X)$, $p > 0$. Denotaremos os vértices de X por $V(X) \ = \ \{\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^p\}$ (do fato 2.4, $\#V(X)$ é finito).

Da definição de casco convexo,

$$\forall \boldsymbol{x} \in [V(X)], \ \exists \ \{\alpha_i\}_{i=1}^p \ : \ \sum_{i:=1}^p \alpha_i \ = \ 1, \quad \alpha_i \geqq 0 \ (i = 1, 2, \ldots, p) \text{ tais que } \boldsymbol{x} \ = \ \sum_{i:=1}^p \alpha_i \boldsymbol{v}^i.$$

Portanto, $\forall \boldsymbol{c} \in \mathrm{I\!R}^n$, $\forall \boldsymbol{x} \in [V(X)]$, teremos

$$\begin{aligned} \langle \boldsymbol{c}, \boldsymbol{x} \rangle \ &= \ \sum_{i:=1}^p \alpha_i \ \langle \boldsymbol{c}, \boldsymbol{v}^i \rangle \leqq \ \sum_{i:=1}^p \alpha_i \ \max_{j=1,2,\ldots,p} \langle \boldsymbol{c}, \boldsymbol{v}^j \rangle \\ &= \ \sum_{i=1}^p \alpha_i \langle \boldsymbol{c}, \overline{\boldsymbol{v}} \rangle = \langle \boldsymbol{c}, \overline{\boldsymbol{v}} \rangle \sum_{i:=1}^p \alpha_i = \langle \boldsymbol{c}, \overline{\boldsymbol{v}} \rangle. \end{aligned}$$

■

Quanto ao segundo termo da expressão apresentada para o valor ótimo do (PLC), $\sup_{\boldsymbol{h} \in C} \langle \boldsymbol{c}, \boldsymbol{h} \rangle$, temos:

Fato 3.3 *Qualquer que seja* $\boldsymbol{c} \in \mathrm{I\!R}^n$, *usando* $C := \{\boldsymbol{h} \in \mathrm{I\!R}^n \ / \ \mathbf{A}\boldsymbol{h} = \mathbf{0}, \ \boldsymbol{h} \ \geqq \mathbf{0}\}$, $\sup_{\boldsymbol{h} \in C} \ \langle \boldsymbol{c}, \boldsymbol{h} \rangle \ \in \ \{0, +\infty\}$. *Além disso,*

(i) $\langle \forall \boldsymbol{h} \in C, \ \ \langle \boldsymbol{c}, \boldsymbol{h} \rangle \leqq \mathbf{0} \rangle \iff \langle \sup_{\boldsymbol{h} \in C} \ \langle \boldsymbol{c}, \boldsymbol{h} \rangle = \max_{\boldsymbol{h} \in C} \ \langle \boldsymbol{c}, \boldsymbol{h} \rangle = 0 \rangle$ *e*

(ii) $\left\langle \exists \overline{\boldsymbol{h}} \in C \ \ \textit{tal que } \langle \boldsymbol{c}, \overline{\boldsymbol{h}} \rangle > 0 \ \right\rangle \iff \langle \sup_{\boldsymbol{h} \in C} \langle \boldsymbol{c}, \boldsymbol{h} \rangle \ = \ +\infty \rangle$.

Prova Se $\langle \forall \boldsymbol{h} \in C, \ \ \langle \boldsymbol{c}, \boldsymbol{h} \rangle \leqq 0 \rangle$, então $\sup_{\boldsymbol{h} \in C} \langle \boldsymbol{c}, \boldsymbol{h} \rangle \ \leqq 0$. Como $\mathbf{0} \in C$ e $\langle \boldsymbol{c}, \mathbf{0} \rangle = 0$, segue que $\sup_{\boldsymbol{h} \in C} \langle \boldsymbol{c}, \boldsymbol{h} \rangle \geqq \mathbf{0}$, e deste modo

$$\langle \forall \boldsymbol{h} \in C, \langle \boldsymbol{c}, \boldsymbol{h} \rangle \leqq 0 \rangle \iff \sup_{\boldsymbol{h} \in C} \langle \boldsymbol{c}, \boldsymbol{h} \rangle = \max_{\boldsymbol{h} \in C} \langle \boldsymbol{c}, \boldsymbol{h} \rangle = 0.$$

De forma análoga, é fácil ver que $\langle sup_{\boldsymbol{h} \in C} \langle \boldsymbol{c}, \boldsymbol{h} \rangle = max_{\boldsymbol{h} \in C} \langle \boldsymbol{c}, \boldsymbol{h} \rangle \rangle \Leftrightarrow \langle \forall \boldsymbol{h} \in C, \langle \boldsymbol{c}, \boldsymbol{h} \rangle \leqq 0 \rangle$, o que conclui a parte (i).

Caso exista $\overline{\boldsymbol{h}} \in C$, para o qual $\langle \boldsymbol{c}, \overline{\boldsymbol{h}} \rangle > 0$, pelo fato de C ser cone, é fácil ver que

$$L \ := \ \{\boldsymbol{h} \in \mathrm{I\!R}^n \ / \ \boldsymbol{h} = \lambda \overline{\boldsymbol{h}}, \ \lambda \geqq 0\} \subseteq C.$$

Deste modo, $\sup_{\boldsymbol{h} \in C} \langle \boldsymbol{c}, \boldsymbol{h} \rangle \ \geqq \ \sup_{\boldsymbol{h} \in L} \langle \boldsymbol{c}, \boldsymbol{h} \rangle \ \geqq \ \lim_{\lambda \to \infty} \lambda \langle \boldsymbol{c}, \overline{\boldsymbol{h}} \rangle = \langle \boldsymbol{c}, \overline{\boldsymbol{h}} \rangle \lim_{\lambda \to \infty} \lambda = +\infty$, e portanto

$$\sup_{\boldsymbol{h} \in C} \langle \boldsymbol{c}, \boldsymbol{h} \rangle \ = \ +\infty.$$

A prova se completa notando que $\left\langle \nexists \overline{\boldsymbol{h}} \in C \ : \ \ \langle \boldsymbol{c}, \overline{\boldsymbol{h}} \rangle > 0 \ \iff \ \forall \boldsymbol{h} \in C, \ \ \langle \boldsymbol{c}, \boldsymbol{h} \rangle \leqq 0 \right\rangle$. ■

Estes fatos formam a base do **teorema fundamental de programação linear** :

Teorema 3.1 *Dado um problema de programação linear na forma canônica, (PLC), velem as equivalências abaixo:*

(a) Valor ótimo do (PLC) $= -\infty$	$\Longleftrightarrow$	*(PLC) é inviável*
	$\Longleftrightarrow$	$V(X) \equiv \emptyset$;
(b) Valor ótimo do (PLC) $= +\infty$	$\Longleftrightarrow$	*(PLC) é ilimitado*
	$\Longleftrightarrow$	$X \not\equiv \emptyset$ *e* $\exists \overline{\boldsymbol{h}} \in C \ / \ \langle \boldsymbol{c}, \overline{\boldsymbol{h}} \rangle > 0$;
(c) Valor ótimo do (PLC) pertence aos reais	$\Longleftrightarrow$	*(PLC) tem solução ótima*
	$\Longleftrightarrow$	$\exists \boldsymbol{v} \in V(X)$ *que é solução ótima do (PLC).*

Prova O item (a) decorre diretamente do fato 3.1. Os itens (b) e (c) decorrem da identidade $\sup_{\boldsymbol{x}\in X}\langle \boldsymbol{c}, \boldsymbol{x}\rangle = \sup_{\boldsymbol{x}\in[V(X)]}\langle \boldsymbol{c}, \boldsymbol{x}\rangle + \sup_{\boldsymbol{h}\in C}\langle \boldsymbol{c}, \boldsymbol{h}\rangle$ e dos fatos 3.2 e 3.3. ∎

Este teorema admite um corolário, por vezes chamado na literatura de teorema fundamental:

Corolário 3.1 *Se o (PLC) possui solução ótima, então existe um vértice que é solução ótima.*

Prova A prova é trivial a partir da implicação $\langle$ (PLC) tem solução ótima $\Rightarrow$ valor ótimo é real $\rangle$ e da parte (c) do teorema 3.1. ∎

Do teorema 3.1 e do corolário acima podemos concluir uma propriedade de problemas de programação linear, "se o valor ótimo é real, então existe solução ótima". Tal não é o caso geral, como pode ser visto por "maximizar $(1 - e^{-x})$, para $x \geqq 0$", que tem valor ótimo igual a 1, mas não possui solução ótima.

Dos resultados acima, podemos observar a existência de duas condições necessárias para o (PLC) ser ilimitado (valor ótimo $= +\infty$) que são: i) X é ilimitado; e
ii) $\exists \overline{\boldsymbol{h}} \in C \ : \ \langle \boldsymbol{c}, \boldsymbol{h}\rangle > 0$.

Entretanto, é importante notar que estas condições não são suficientes para ilimitação do (PLC).

Exercício 3.1 *Prove que*

i) valor ótimo $= +\infty \implies \Big\langle C \not\equiv \{\boldsymbol{0}\}$ *e existe* $\overline{\boldsymbol{h}} \in C : \ \langle \boldsymbol{c}, \overline{\boldsymbol{h}}\rangle > 0\Big\rangle$;

ii) $\exists \overline{\boldsymbol{h}} \in C : \ \langle \boldsymbol{c}, \overline{\boldsymbol{h}}\rangle > 0 \not\Rightarrow$ *valor ótimo* $= +\infty$ *(a menos que acrescentemos a hipótese* $X \not\equiv \emptyset$*).*

Exercício 3.2 *Prove que*

i) Valor ótimo $= +\infty \implies X$ *ilimitado;*

ii) X *ilimitado* $\not\Longleftarrow$ *valor ótimo* $= +\infty$*;*

iii) valor ótimo $\in \mathbb{R} \not\Longleftarrow X$ *limitado.*

Utilizando a caracterização finita de poliedros, podemos reescrever o teorema 3.1 como:

Fato 3.4 *Dado o (PLC), sejam*

$V(X) = \{\boldsymbol{v}^i\}_{i=1}^p$ *e* *(por convenção:* $p = 0 \iff V(X) \equiv \emptyset$*)*

$V(C^1) = \{\boldsymbol{r}^j\}_{j=1}^q$ *(por convenção:* $q = 0 \Leftrightarrow V(C^1) \equiv \emptyset$*),*

onde $C^1 = \{\boldsymbol{h} \in \mathbb{R}^n / \ \mathbf{A}\boldsymbol{h} = \boldsymbol{0}, \ \ \boldsymbol{h} \geqq \boldsymbol{0}, \|\boldsymbol{h}\|_1 = 1\}$*, então:*

(a) (PLC) inviável $\iff V(X) \equiv \emptyset$*;*

(b) (PLC) ilimitado $\iff V(X) \not\equiv \emptyset$ *e* $\exists j \in \{1, 2, \ldots, q\} : \ \langle \boldsymbol{c}, \boldsymbol{r}^j\rangle > 0$*;*

(c) (PLC) tem solução ótima $\overline{\boldsymbol{x}} \iff \langle \exists \overline{\boldsymbol{v}} \in V(X), \ \langle \boldsymbol{c}, \overline{\boldsymbol{x}}\rangle = \langle \boldsymbol{c}, \overline{\boldsymbol{v}}\rangle = \max_{i=1,2,\ldots,p}\langle \boldsymbol{c}, \boldsymbol{v}^i\rangle$ *e* $\langle \boldsymbol{c}, \boldsymbol{r}^j\rangle \leqq 0, \forall j.\rangle$

Prova O item (a) segue do teorema 2.5 (página 15: $X \not\equiv \emptyset \Leftrightarrow \ V(X) \not\equiv \emptyset$). Os itens (b) e (c) seguem da seguinte equivalência

$$\exists \boldsymbol{h} \in C \ : \ \langle \boldsymbol{c}, \boldsymbol{h}\rangle \ > 0 \iff \langle C^1 \not\equiv \emptyset \ \text{ e } \ \exists \boldsymbol{r} \in V(C^1) : \langle \boldsymbol{c}, \boldsymbol{r}\rangle > 0\rangle.$$

∎

Exercício 3.3 *Seja X^* o conjunto de soluções ótimas do (PLC). Prove que*

$$X^* = [\{\boldsymbol{v}^i\}_{i\in I^*}] + C^*,$$

onde

$$I^* := \{i : \boldsymbol{v}^i \in V(X) \text{ e } \langle \boldsymbol{c}, \boldsymbol{v}^i\rangle = \text{ valor ótimo de (PLC) } \text{ e}$$

$$C^* = \{\boldsymbol{h} \in \mathbb{R}^n \ / \ \mathbf{A}\boldsymbol{h} = 0, \ \boldsymbol{h} \geqq 0, \ \ \langle \boldsymbol{c}, \boldsymbol{h}\rangle = 0 \ \}.$$

Exercício 3.4 *Construa um exemplo no qual $X^* \not\equiv \emptyset$ e $C^* \not\equiv \mathbf{0}$.*

3.2 Um Modelo de Algoritmo

O fato 3.4 motiva naturalmente um método finito de resolver o problema de programação linear (canônico), que é descrito abaixo:

Algoritmo 1 Algoritmo combinatório para resolver o (PLC)

Passo 1: *Calcule $V(X)$, o conjunto dos vértices de $X = \{\boldsymbol{x} \in \mathbb{R}^n / \ \mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \ \ \boldsymbol{x} \geqq \mathbf{0}\}$*
Se $V(X) \equiv \emptyset$ pare: o problema é inviável!

Passo 2: *Denote $V(X) = \{\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^p\}$ $(\not\equiv \emptyset)$*
Encontre $\overline{\boldsymbol{v}} \in argmax\{\langle \boldsymbol{c}, \boldsymbol{v}^i\rangle : i \in \{1, 2, \ldots, p\}\}$

Passo 3: *Calcule $V(C^1)$, o conjunto dos vértices de $C^1 = \{\boldsymbol{h} \in \mathbb{R}^n / \ \mathbf{A}\boldsymbol{h} = \mathbf{0}, \ \boldsymbol{h} \geqq \mathbf{0}, \|\boldsymbol{h}\|_1 = 1\}$*
Se $V(C^1) = \emptyset$, então pare: $\overline{\boldsymbol{v}}$ é solução do (PLC)! (isso decorre do fato 3.4)

Passo 4: *Seja $V(C^1) = \{\boldsymbol{r}^1, \boldsymbol{r}^2, \ldots, \boldsymbol{r}^q\}$ $(\not\equiv \emptyset)$*
Se $\exists \boldsymbol{r}^k \in V(C^1)$, tal que $\langle \boldsymbol{c}, \boldsymbol{r}^k\rangle > 0$, então pare: o problema é ilimitado!

> *Pois $L = \{\boldsymbol{x} \in \mathbb{R}^n \ / \ \boldsymbol{x} = \overline{\boldsymbol{v}} + \lambda \boldsymbol{r}^k, \ \ \lambda \geqq 0\}$ é uma semireta de ilimitação, isto é,*
>
> *$L \subseteq X$ e $\lim_{\lambda\to\infty} \langle \boldsymbol{c}, \overline{\boldsymbol{v}}\rangle + \lambda \boldsymbol{r}^k = +\infty$.*

Se $\langle \boldsymbol{c}, \boldsymbol{r}^j\rangle \leqq 0, \ \forall j \in \{1, \ldots, q\}$, pare: $\overline{\boldsymbol{v}}$ é solução do (PLC)!

Este algoritmo é finito se soubermos computar os vértices de um poliedro qualquer em um número finito de passos. Portanto, para poder discutir como implementar este tipo de algoritmo, temos que discutir como determinar vértices de um poliedro (X ou C^1).

Para gerar vértices é natural que nos apoiemos em uma caracterização algébrica de vértices, dada pelo teorema 2.2 (página 15):

$$\boldsymbol{x} \in V(X) \iff \begin{cases} (i)\ \boldsymbol{x} \in X, \text{ e} \\ (ii)\ \{\boldsymbol{a}^j\}_{j\in I(\boldsymbol{x})} \text{ é l.i. } (\Longrightarrow \#I(\boldsymbol{x}) \leqq m) \end{cases}$$

Esta caracterização sugere um método "inocente" de cômputo de elementos de $V(X)$, descrito a seguir.

Algoritmo 2 Método Combinatório "Inocente" para Cômputo de $V(X)$

Passo 1: *Seja $J = \{1, 2, \ldots, n\}$ $(m \leq n)$*

Passo 2: *Calcule: (combinação de n elementos, m a m)*
$\beta^0 := \{I \subset J / \ \#I \leqq m\}$
$\beta^1 := \{I \in \beta^0 / \ \{\boldsymbol{a}^j\}_{j\in I}$ *é um conjunto l.i.*$\}$
$\beta^2 := \{I \in \beta^1 / \ \boldsymbol{b} = \sum_{j\in I} \alpha_j \boldsymbol{a}^j \ \text{ e } \boldsymbol{\alpha} \geqq \mathbf{0}\}$

Passo 3: Para cada $I \in \beta^2$, faça

$$v_i = \alpha_i, \quad \textit{se } i \in I \quad e$$
$$v_j = 0, \quad \textit{se } j \notin I$$

(se o passo 3, $\beta^2 \equiv \emptyset$ não for executado sequer uma vez, então teremos $V(X) \equiv \emptyset$)

Em relação a este procedimento cabem as seguintes observações:

1: A lógica para gerar β^0 é fácil, porém muito "trabalhosa" no sentido que $\#\beta^0 = \sum_{i=0}^{m} \binom{n}{i}$.

2: Diferentes $I \in \beta^2$, podem gerar o mesmo vértice, mas cada $I \in \beta^2$ gera um único vértice pois o conjunto $\{\boldsymbol{a}^j\}_{j\in I}$ ser linearmente independente implica na unicidade dos $\{\alpha_j\}_{j\in I}$ tais que $\boldsymbol{b} = \sum_{j\in I} \alpha_j \boldsymbol{a}^j$.

Cabe a pergunta: Será necessário gerar todos os vértices para resolver o (PLC)? Isto é, será realmente necessária uma enumeração explícita dos vértices ou será possível uma enumeração implícita ?

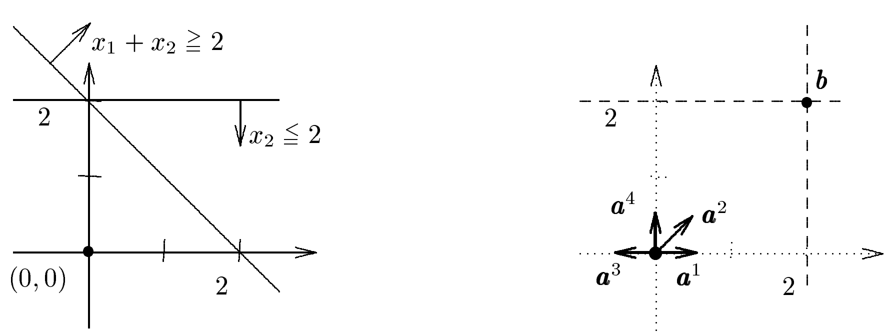
Para responder a estas questões, é conveniente explorar mais a caracterização algébrica de vértices. Iniciemos o estudo no caso em que a matriz $(\mathbf{A})$ associada ao poliedro (X) tem **característica plena** (posto $\mathbf{A} = m$), ou seja, as restrições $(\mathbf{A}\boldsymbol{x} = \boldsymbol{b})$ são linearmente independentes.

3.3 Vértices e Soluções Básicas

Consideremos o exemplo

Exemplo 3.3.1 *Seja*

$$\begin{aligned} Y &= \{\boldsymbol{x} \in \mathbb{R}^2 \ / \ x_1 + x_2 \geqq 2 \ \textit{e} \ x_2 \leqq 2, \ \boldsymbol{x} \geqq \mathbf{0}\} \\ &= \{\boldsymbol{x} \in \mathbb{R}^2 \ / \ \overline{\mathbf{A}}\boldsymbol{x} \leqq \begin{bmatrix} -2 \\ 2 \end{bmatrix} \ \boldsymbol{x} \geqq \mathbf{0}\} \quad \textit{onde} \quad \overline{\mathbf{A}} = \begin{bmatrix} -1 & -1 \\ 0 & 1 \end{bmatrix}. \end{aligned}$$



(a) Poliedro Y (na dimensão de $\boldsymbol{x}$) (b) Representação das colunas de $\mathbf{A}$ (na dimensão de $\boldsymbol{b}$)

Figura 3.1: *(a) Poliedro Y definido por dois semi-espaços no ortante positivo; (b) Correspondente canônico no espaço de $\boldsymbol{b}$.*

Da figura (a) acima, que representa o poliedro Y, é fácil ver que $V(Y) = \{\boldsymbol{y}^1, \boldsymbol{y}^2\}$ onde $\boldsymbol{y}^1 = (2,0)^t$, $\boldsymbol{y}^2 = (0,2)^t$.

Para discutir este caso, dentro de nosso escopo teórico, há a necessidade de recairmos em um poliedro com a forma canônica (representado, via colunas, na figura (b) acima):

$$\begin{aligned} X &= \{(\boldsymbol{x}, \boldsymbol{x}^r) \in \mathbb{R}^2 \times \mathbb{R}^2 / \begin{bmatrix} x_1 + x_2 - x_1^r = 2 \\ x_2 + x_2^r = 2 \end{bmatrix}, \ \boldsymbol{x} \geqq \mathbf{0}, \ \ \boldsymbol{x}^r \geqq \mathbf{0}\} \\ &= \{\boldsymbol{x} \in \mathbb{R}^4 / \ \mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \ \boldsymbol{x} \geqq \mathbf{0}\}, \quad \textit{onde } \mathbf{A} := \begin{bmatrix} 1 & 1 & -1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \ \textit{e} \ \boldsymbol{b} := \begin{bmatrix} 2 \\ 2 \end{bmatrix}. \end{aligned}$$

A relação entre $V(Y)$ e $V(X)$ é dada pelo seguinte fato:

Fato 3.5 *Seja* $Y = \{\boldsymbol{x} \in \mathbb{R}^n /\ \mathbf{M}\boldsymbol{x} \leqq \boldsymbol{b},\ \boldsymbol{x} \geqq \mathbf{0}\}$ *e* $X = \{(\boldsymbol{x}, \boldsymbol{x}^r) \in \mathbb{R}^n \times \mathbb{R}^m /\ \ \mathbf{M}\boldsymbol{x} + \boldsymbol{x}^r = \boldsymbol{b},\ \boldsymbol{x} \geqq \mathbf{0},\ \boldsymbol{x}^r \geqq \mathbf{0}\}$ *Então* $\overline{\boldsymbol{x}} \in V(Y)$ *e* $\overline{\boldsymbol{x}}^r = \boldsymbol{b} - \mathbf{M}\overline{\boldsymbol{x}} \Longleftrightarrow (\overline{\boldsymbol{x}}, \overline{\boldsymbol{x}}^r) \in V(X)$.

Prova Deixada a cargo do leitor. ∎

Observação: É importante notar que sem a restrição $\boldsymbol{x} \geqq \mathbf{0}$ no poliedro original, a implicação "$\Leftarrow$" não é válida. Um exemplo onde esta não vale é quando Y não tem sequer um vértice (p.e., $V(Y) := \{\boldsymbol{x} \in \mathbb{R}^2 /\ 0 \leqq x_1 \leqq 1\}$).

Portanto, usando o fato 3.5, segue que $V(X) = \{\boldsymbol{v}^1, \boldsymbol{v}^2\}$, onde $\boldsymbol{v}^1 = (2,0,0,2)$ e $\boldsymbol{v}^2 = (0,2,0,0)$.

Tendo em mente este resultado, vamos aplicar o método *inocente* para cômputo de $V(X)$ ao exemplo 3.3.1. Isto é, apliquemos o método da seção anterior para computar os vértices, no caso

$$\mathbf{A} = [\mathbf{M}|\mathbf{I}] = \begin{bmatrix} -1 & -1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \ \text{ e } \ \boldsymbol{b} = \begin{bmatrix} -2 \\ 2 \end{bmatrix},$$

com identificação $x_3 = x_1^r$ e $x_4 = x_2^r$. Então, $J = \{1,2,3,4\}$, com $m = 2$ e

$$\beta^0 = \{I \subset J \ /\ \#I \leqq 2\} = \{\emptyset, \{1\}, \{2\}, \{3\}, \{4\}, \{1,2\}, \{1,3\}, \{1,4\}, \{2,3\}, \{2,4\}, \{3,4\}\} = \{I^i\}_{i=0}^{10},$$
$$\#\beta^0 = \tbinom{4}{0} + \tbinom{4}{1} + \tbinom{4}{2} = 1 + 4 + 6 = 11 = \sum_{i=0}^{m} \tbinom{n}{i}.$$

Para cada $I \in \beta^0$, testamos sucessivamente se $I \in \beta^1$ e se $I \in \beta^2$. Caso $I \in \beta^2$, computamos o vértice associado:

(I^0) $I^0 = \emptyset :$ $I^0 \in \beta^1$ (trivial) e $I^0 \notin \beta^2$ pois $\boldsymbol{b} \neq \mathbf{0}$.

(I^1) $I^1 = \{1\}$: $I^1 \in \beta^1$ $(\boldsymbol{a}^1 \neq \mathbf{0})$ e $\nexists \alpha_1 : \alpha_1 \boldsymbol{a}^1 = \alpha_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$, portanto $I^1 \notin \beta^2$.

(I^2) $I^2 = \{2\}$: $I^2 \in \beta^2$ $(\boldsymbol{a}^2 \neq \mathbf{0})$ e $2\boldsymbol{a}^2 = 2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \boldsymbol{b}$, portanto $I^2 \in \beta^2$, e temos o vértice $\boldsymbol{y}^2 = (0, \alpha_2, 0, 0)^t = (0,2,0,0)^t$.

(I^3, I^4) Analogamente a (I^1), $I^j \notin \beta^2$, $j \in \{3,4\}$.

(I^5) $I^5 = \{1,2\}$: $I^5 \in \beta^1$ pois $\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\} = \{\boldsymbol{a}^1, \boldsymbol{a}^2\}$, é um conjunto l.i.

Mais ainda, $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ admite solução única $\alpha_1 = 0$ e $\alpha_2 = 2$ $(\alpha_1 \geqq 0, \alpha_2 \geqq 0)$, portanto $I^5 \in \beta^2$ e temos o vértice $\boldsymbol{y}^5 = (0,2,0,0)^t$.

(I^6) $I^6 = \{1,3\}$: $I^6 \notin \beta^1$ pois $\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix}\right\} = \{\boldsymbol{a}^1, \boldsymbol{a}^3\}$, é um conjunto l.d. .

(I^7) $I^7 = \{1,4\}$: $I^7 \in \beta^1$ pois $\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\} = \{\boldsymbol{a}^1, \boldsymbol{a}^4\}$, é um conjunto l.i.

Mais ainda, $(\boldsymbol{a}^1, \boldsymbol{a}^4) \begin{bmatrix} \alpha_1 \\ \alpha_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_4 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ admite solução única $\alpha_1 = 2$ e $\alpha_4 = 2$ $(\alpha_1 \geqq 0, \alpha_4 \geqq 0)$, portanto $I^7 \in \beta^2$ e temos o vértice $\boldsymbol{y}^7 = (2,0,0,2)^t$.

(I^8) $I^8 = \{2,3\}$: $I^8 \in \beta^1$ pois $\{\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}-1\\0\end{bmatrix}\} = \{\boldsymbol{a}^2, \boldsymbol{a}^3\}$, é um conjunto l.i.

Mais ainda, $(\boldsymbol{a}^2, \boldsymbol{a}^3)\begin{bmatrix}\alpha_2\\\alpha_3\end{bmatrix} = \begin{bmatrix}1 & -1\\1 & 0\end{bmatrix}\begin{bmatrix}\alpha_2\\\alpha_3\end{bmatrix} = \begin{bmatrix}2\\2\end{bmatrix}$ admite solução única $\alpha_2 = 2$ e $\alpha_3 = 0$ $(\alpha_2 \geqq 0, \alpha_3 \geqq 0)$, e temos o vértice $\boldsymbol{y}^8 = (0,2,0,0)^t$.

(I^9) $I^9 = \{2,4\}$, analogamente a I^8, $I^9 \in \beta^2$ e temos o vértice $\boldsymbol{y}^9 = (0,2,0,0)^t$.

(I^{10}) $I^{10} = \{3,4\}$: $I^{10} \in \beta^1$ pois $\{\begin{bmatrix}-1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}\} = \{\boldsymbol{a}^3, \boldsymbol{a}^4\}$, é um conjunto l.i.

Mais ainda, $(\boldsymbol{a}^3, \boldsymbol{a}^4)\begin{bmatrix}\alpha_3\\\alpha_4\end{bmatrix} = \begin{bmatrix}-1 & 0\\0 & 1\end{bmatrix}\begin{bmatrix}\alpha_3\\\alpha_4\end{bmatrix} = \begin{bmatrix}2\\2\end{bmatrix}$ admite solução única $\alpha_3 = -2 < 0$ e $\alpha_4 = 2$, portanto $I^{10} \notin \beta^2$

Pelo método *inocente* obtivemos $V(X) = \{\boldsymbol{y}^2, \boldsymbol{y}^5, \boldsymbol{y}^7, \boldsymbol{y}^8, \boldsymbol{y}^9\}$,com $\boldsymbol{y}^2 = \boldsymbol{y}^5 = \boldsymbol{y}^8 = \boldsymbol{y}^9 = (0,2,0,0)^t = \boldsymbol{v}^2$ e $\boldsymbol{y}^7 = (2,0,0,2)^t = \boldsymbol{v}^1$. Isto é, $V(X) = \{\boldsymbol{v}^1, \boldsymbol{v}^2\}$, como já haviamos determinado.

O primeiro ponto que o exemplo ilustra é que distintos $I \in \beta^2$ podem gerar o mesmo vértice. No caso I^2, I^5, I^8 e I^9 geram o vértice $\boldsymbol{v}^2$.

A razão pela qual isto ocorre, é que $I(\boldsymbol{v}^2) = \{2\} = I^2$, isto é, $\boldsymbol{b}$ pertence ao subespaço gerado pela coluna $\boldsymbol{a}^2$. Ao considerarmos os conjuntos $I^5 = \{1,2\}$, $I^8 = \{2,3\}$, e $I^9 = \{2,4\}$, usamos respectivamente as colunas $\boldsymbol{a}^1$, $\boldsymbol{a}^3$ e $\boldsymbol{a}^4$ para completar bases do $\mathbb{R}^2$. Em outros termos, como $\boldsymbol{b} = 2\boldsymbol{a}^2 = 2\boldsymbol{a}^2 + 0\boldsymbol{a}^1 = 2\boldsymbol{a}^2 + 0\boldsymbol{a}^3 = 2\boldsymbol{a}^2 + 0\boldsymbol{a}^4$, existem vários conjuntos em β^2 associados ao mesmo vértice.

Esta constatação serve de motivação para o seguinte fato:

Fato 3.6 *Seja* $X = \{\boldsymbol{x} \in \mathbb{R}^n \;/\; \mathbf{A}\boldsymbol{x} = \boldsymbol{b},\;\; \boldsymbol{x} \geqq \mathbf{0}\}$ *onde* $\mathbf{A}$ *é uma matriz com característica plena (posto* $\mathbf{A} = m$*). Então*

$$\boldsymbol{v} \in V(X) \quad \Leftrightarrow \quad \begin{cases} \exists I \subseteq \{1, 2 ... n\} \\ (i) \{\boldsymbol{a}^i\}_{i \in I} \text{ é uma base do } \mathbb{R}^m \\ (ii) \sum\limits_{i \in I} \alpha_i \boldsymbol{a}^i = \boldsymbol{b} \text{ tem como única solução } \begin{cases} \alpha_i = v_i \geqq 0 & (i \in I) \\ \alpha_i = v_i = 0 & (i \notin I) \end{cases} \end{cases}$$

Prova

$(\Longleftarrow)$ Trivial, pois (ii) implica que $\boldsymbol{v} \in X$ e que $I(\boldsymbol{v}) \subseteq I$. Assim sendo, (i) corresponde a caracterização de vértices por colunas l.i.

$(\Longrightarrow)$ (a) Se $\#I(\boldsymbol{v}) = m$, então trivialmente $I \equiv I(\boldsymbol{v})$.

(b) Se $\#I(\boldsymbol{v}) = p < m$, como $\mathbf{A}$ tem característica plena e $\{\boldsymbol{a}^j\}_{j \in I(\boldsymbol{v})}$ são l.i. , existem $\{j_i\}_{i=p+1}^m$ tais que $[\ \{\boldsymbol{a}^j\}_{j \in I(\boldsymbol{v})}, \{\boldsymbol{a}^{j_i}\}_{i=p+1}^m\ ]$ formam uma base do $\mathbb{R}^m$.

Logo $\sum\limits_{i \in I(\boldsymbol{v})} \alpha_i \boldsymbol{a}^i + \sum\limits_{i=p+1}^{m} \alpha_i \boldsymbol{a}^{j_i} = \boldsymbol{b}$ tem solução única $\overline{\boldsymbol{\alpha}}$. Como $\boldsymbol{v} \in V(X)$, segue que $\sum\limits_{j \in I(\boldsymbol{v})} \overline{\boldsymbol{\alpha}}_j \boldsymbol{a}^j = \boldsymbol{b}$, e daí temos

$$\begin{aligned} \overline{\boldsymbol{\alpha}}_i = v_i \geqq 0, & \qquad i \in I(\boldsymbol{v}) \\ \overline{\boldsymbol{\alpha}}_i = v_i = 0, & \qquad i \in \{p+1, p+2, \ldots, m\}. \end{aligned}$$

Portanto,

$$I = I(\boldsymbol{v}) \bigcup \{j_i \;/\; i = p+1, p+2, \ldots, m\},$$

é o conjunto desejado. ∎

A formulação usual do fato acima na literatura é calcada nas seguintes definições:

Definição 3.1 *O ponto* $\overline{\boldsymbol{x}}$ *é dito* **solução básica** *do sistema linear* $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$, *se existe* $I \subseteq \{1, 2, \ldots, n\}$ *tal que*
(i) $\{\boldsymbol{a}^i\}_{i \in I}$ *é uma base do* $\mathbb{R}^m$;
(ii) $I(\overline{\boldsymbol{x}}) \subseteq I$.

Definição 3.2 *O ponto* $\overline{\boldsymbol{x}}$ *é dito* **solução básica viável** *do sistema linear* $\{\mathbf{A}\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geqq \mathbf{0}\}$, *se* $\overline{\boldsymbol{x}}$ *é solução básica de* $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ *e* $\boldsymbol{x} \geqq \mathbf{0}$.

Por simplicidade falaremos em solução básica e solução básica viável, sem indicar os sistemas lineares associados, sempre que não haja ambiguidade. Assim sendo o fato 3.6 pode ser reescrito como:

Fato 3.7 *Seja* $X = \{\boldsymbol{x} \in \mathbb{R}^n \ / \ \mathbf{A}\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geqq \mathbf{0}\}$, *onde* $\mathbf{A}$ *é uma matriz com característica plena. Então*
$\boldsymbol{v} \in V(X) \iff \boldsymbol{v}$ *é solução básica viável (de* $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \quad \boldsymbol{x} \geqq \mathbf{0}$*).*

A caracterização de vértices por soluções básicas viáveis indica que a cada vértice podemos associar uma ou mais bases do $\mathbb{R}^m$ extraidas das colunas de $\mathbf{A}$. Isto implica que a enumeração de todos os conjuntos de p colunas ($p \leqq m$) é redundante, no caso da matriz $\mathbf{A}$ ter característica plena. Pois, se o vértice v for tal que $\#I(\boldsymbol{v}) < m$, ele será encontrado na enumeração de todos os conjuntos de m colunas l.i. que contiverem as colunas associadas a $I(\boldsymbol{v})$. Isto é retratado no exemplo, onde $I(\boldsymbol{v}^2) = \{2\}$ e $\boldsymbol{v}^2$ é "encontrado" ao enumerarmos $\{2, 1\}$, $\{2, 3\}$ e $\{2, 4\}$.

Portanto, se $\mathbf{A}$ tem característica plena, para encontrarmos os vértices basta enumerar os conjuntos de m colunas l.i. e verificar se a cada conjunto pode ser associado um ponto viável.

Esta verificação é simples pois m colunas l.i. formam uma base B do $\mathbb{R}^m$ e, portanto, b admite uma única representação nesta base (solução básica). Se todas as componentes de b, em relação a base B, são não negativas, então obtivemos um vértice (solução básica viável). Caso contrário, à esta base do $\mathbb{R}^m$ está associada uma solução básica, porém não a um vértice.

Com estas observações, podemos enunciar o seguinte método para encontrar o conjunto $V(X)$, quando $\mathbf{A}$ tiver característica plena.

3.3.1 Método para Cômputo de $V(X)$ para posto $\mathbf{A} = m$

Estamos sempre assumindo $n \geqq m$.

Algoritmo 3 Algoritmo Combinatório "melhorado"

Versão 1: análoga ao algoritmo 1
Passo 1: Seja $J = \{1, 2, \ldots, n\} \qquad (m \leq n)$
Passo 2: Calcule: (combinação n, m a m)
$M := \{I \subset J / \ \#I = m\}$
$CB := \{I \in M / \ \{\boldsymbol{a}^j\}_{j \in I}$ *é um conjunto l.i.*$\}$
$CBV := \{I \in CB / \ \boldsymbol{b} = \sum_{j \in I} \alpha_j \boldsymbol{a}^j \text{ e } \boldsymbol{\alpha} \geqq \mathbf{0}\}$
Passo 3: Para cada $I \in CBV$, faça (vértice $\boldsymbol{v}$)
$v_i = \alpha_i, \quad$ se $i \in I$ e
$v_j = 0, \quad$ se $j \notin I$

Versão 2: melhora na versão 1
Passo 1: Seja $J = \{1, 2, \ldots, n\} \qquad (m \leq n)$
Passo 2: Calcule: (combinação n, m a m)
$Perm := \{I \subset J : \#I = m\}$
Passo 3: Para cada $I \in Perm$, faça
se ($\{\boldsymbol{a}^i\}_{i \in I}$ é l.i.) então
encontre $\boldsymbol{\alpha} \in \mathbb{R}^m$ tal que $\sum_{i \in I} \alpha_i \boldsymbol{a}^i = \boldsymbol{b}$
se ($\boldsymbol{\alpha} \geq \mathbf{0}$) então defina
$v_i = \alpha_i \quad$ se $i \in I$ e
$v_j = 0 \quad$ se $j \notin I$

(se o passo 3 não for executado sequer uma vez, então teremos $V(X) \equiv \emptyset$*)*

Observe que $CB \neq \emptyset$ (existe alguma solução básica para $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$) se, e somente se, $\mathbf{A}$ tem característica plena. Em relação a este procedimento e ao método *inocente* apresentado anteriormente, cabem as seguintes comparações:

(i) Este procedimento, apesar de computacionalmente custoso, é superior ao inocente pois $\#M < \#\beta^0$, $\#CB \leqq \#\beta^1$ e $\#CBV \leqq \#\beta^2$.

(ii) A possibilidade de diferentes bases $\{A^i\}_{i \in I}$ gerarem o mesmo vértice é mantida, porém se houver duplicação, esta será menor que no método *inocente*.

Exercício 3.5 *Aplique o método para cômputo de V(X) para o exemplo 3.3.1 e compare com o método combinatório inocente.*

A existência de mais de uma base associada a um único vértice é parcialmente respondida por:

Fato 3.8 *Se $\overline{\boldsymbol{x}}$ é solução básica e $\#I(\overline{\boldsymbol{x}}) = m$, então existe e é única a base $\{\boldsymbol{a}^j\}_{j \in I}$ associada ao ponto $\overline{\boldsymbol{x}}$.*

Prova Trivial, pois $I = I(\overline{\boldsymbol{x}})$. ∎

Como no caso em que $\#I(\overline{\boldsymbol{x}}) < m$ podem existir várias bases associadas à mesma solução básica, dizemos que:

Definição 3.3 *Uma solução básica é dita degenerada (***solução básica degenerada** *), se $\#I(\overline{\boldsymbol{x}}) < m$, e uma solução básica viável degenerada é dita ser um* **vértice degenerado** .

Portanto uma condição suficiente para garantir que a cada vértice corresponde uma única base é supor que não existam vértices degenerados no poliedro X. Esta suposição é formalizada por:

Definição 3.4 *Dizemos que a* **hipótese de não degenerescência (HND)** *é válida para o poliedro $X = \{\boldsymbol{x} \in \mathbb{R}^n \ / \ \mathbf{A}\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geqq \mathbf{0}\}$, se para todo $\boldsymbol{v} \in V(X)$, $\#I(\boldsymbol{v}) = m$.*

É claro que a hipótese de não degenerescência (HND) implica que, se $X \not\equiv \emptyset$, então a característica de $\mathbf{A}$ é m (posto de $\mathbf{A}$). Mais ainda

Fato 3.9 *Seja $\boldsymbol{x} \in X := \{\boldsymbol{x} \in \mathbb{R}^n / \ \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geqq \mathbf{0}\}$, para o qual vale (HND). Então*

$$\boldsymbol{x} \in V(X) \iff \#I(\boldsymbol{x}) = m.$$

Prova

($\Rightarrow$) Trivial pela definição de (HND).

($\Leftarrow$) Suponha por contradição exista $\boldsymbol{x} \in X$, $\boldsymbol{x} \notin V(X)$ tal que $\#I(\boldsymbol{x}) = m$. Aplicando o método conceitual para obtenção de vértices (MCOV), na página 16, podemos obter $\boldsymbol{y} \in V(X)$ tal que $\#I(\boldsymbol{y}) < \#I(\boldsymbol{x}) = m$, que contraria a (HND). ∎

É fácil verificar que sob (HND) todo ponto do poliedro X tem m componentes não nulas:

Exercício 3.6 *Prove que sob (HND), $\boldsymbol{x} \in X \implies \#I(\boldsymbol{x}) \geqq m$.*

Sob a hipótese de não degenerescência, não há sentido em utilizar o método *inocente* para o cômputo de $V(X)$ (sob (HND), $\boldsymbol{x} \in X \implies \#I(\boldsymbol{x}) \geqq m$). Além disto, ao aplicarmos o método para cômputo de $V(X)$ (posto $\mathbf{A} = m$) cada vértice é gerado uma única vez. Isto é, a cada vértice está associado um único conjunto de índices $I \in CBV$ $(\{\boldsymbol{a}^j\}_{j\in I}$ forma uma base do $\mathbb{R}^m$ e $\boldsymbol{\alpha} \geqq \mathbf{0}$, para $\boldsymbol{b} = \sum_{j\in I} \alpha_j \boldsymbol{a}^j)$.

Uma observação interessante é que para toda solução básica $\overline{\boldsymbol{x}}$ cuja base é $\{\boldsymbol{a}^j\}_{j\in I}$ podemos obter $m!$ representações matriciais para o sistema $\sum_{j\in I} x_j \boldsymbol{a}^j = \boldsymbol{b}$. Estas várias representações estão associadas as diferentes ordenações de colunas e correspondentes variáveis.

Para podermos trabalhar dentro de um contexto matricial, doravante assumiremos uma ordenação quando analisarmos qualquer subconjunto de m-índices das colunas $\boldsymbol{a}^i$ que formam uma base. Ou seja, um conjunto I de m índices será sempre tratado através de uma **m-upla ordenada** (**lista**) $B = (j_1, j_2, \ldots, j_m)$. Para simplificar notação utilizaremos estas listas como se fossem conjuntos, isto é, usaremos "$j \in B$" para significar $j \in \{j_1, j_2, \ldots, j_m\}$.

Esta idéia de ordenação não é nova, pois sabemos que quaisquer m vetores l.i. no $\mathbb{R}^m$ formam uma base. Mas para podermos exprimir um vetor em coordenadas associadas a esta base, devemos fixar uma ordenação da base dada. Por exemplo o vetor $(2,1)^t$ pode ser expresso em relação a base $\boldsymbol{e}^1 = (1,0)^t$, $\boldsymbol{e}^2 = (0,1)^t$ de duas formas diferentes:

$$(2,1)^t = 2 \times (1^{\underline{o}} \text{ vetor base }) \;+\; 1 \times (2^{\underline{o}} \text{ vetor base}), \;\; \text{se } B = (\boldsymbol{e}^1, \boldsymbol{e}^2) \text{ e}$$

$$(2,1)^t = 1 \times (1^{\underline{o}} \text{ vetor base }) \;+\; 2 \times (2^{\underline{o}} \text{ vetor base}), \;\; \text{se } B = (\boldsymbol{e}^2, \boldsymbol{e}^1).$$

Fato 3.10 *Dado um conjunto I, $I = \{j_1, j_2, \ldots, j_m\}$, as seguintes afirmações são equivalentes sobre a lista B associada:*

(a) $B \in CB$, isto é, $\{\boldsymbol{a}^j\}_{j\in B}$ é um conjunto l.i. e $\#B = m$;

(b) $B = (j_1, j_2, \ldots, j_m)$ está associado à base ordenada do $\mathbb{R}^m$ $(\boldsymbol{a}^{j_1}, \ldots, \boldsymbol{a}^{j_m})$;

(c) a matriz $\mathbf{A}_B = [\boldsymbol{a}^{j_1}| \cdots |\boldsymbol{a}^{j_m}]$, onde a i-ésima coluna é $\boldsymbol{a}^{j_i}$, é uma matriz (quadrada) não singular.

Prova Omitida por ser elementar. ■

Podemos então reescrever o método para cômputo de $V(X)$, no caso de A ter característica plena sob forma matricial, cujas vantagens computacionais serão exploradas adiante.

3.3.2 Método Matricial para Cômputo de $V(X)$ para posto $\mathbf{A} = m$

Algoritmo 4 Algoritmo Matricial para Cômputo de $V(X)$

Passo 1: Seja $J = \{1, 2, \ldots, n\}$ $\quad (m \leq n)$

Passo 2: Calcule $M = \{I_1, I_2, \ldots, I_Q\} = \{I \subseteq J \;/\; \#I = m\}$

Passo 3: Para $i \in \{1, 2, \ldots, Q\}$, *enumere* $I_i = (j_1^i, j_1^i, \ldots, j_m^i)$
 Seja $B := I_i = (j_1, j_2, \ldots, j_m)$ $\quad$ *(para simplificar notação)*
 Se existe $\mathbf{A}_B^{-1} = [\boldsymbol{a}^{j_1}|\boldsymbol{a}^{j_2}| \cdots |\boldsymbol{a}^{j_m}]^{-1}$, *então*
  $\boldsymbol{b}' := \mathbf{A}_B^{-1}\boldsymbol{b}$
  Se $\boldsymbol{b}' \geqq \mathbf{0}$, *então* I_i *é base viável*
   / calcule o vértice* $\boldsymbol{v}$ *associado a* I_i *fazendo: */*
   $v_{j_k} := b'_k$, $k \in \{1, 2, \ldots, m\}$ $(j_k \in I_i)$, $\quad$ *e*
   $v_j := 0$, $\quad j \in \{1, 2, \ldots, n\} \setminus I_i$.
  Senão: não há vértice associado a I_i

Desta forma o cálculo dos vértices de $V(X)$ recaiu em um estudo de soluções de sistemas lineares, isto é, de inversão de matrizes (ou ao menos, diagonalização). No caso em que $det(\mathbf{A}_B) \neq 0$, a m-upla $B = (j_1, j_2, \ldots, j_m)$

está associada a uma base $(\boldsymbol{a}^{j_1}, \boldsymbol{a}^{j_2}, \ldots, \boldsymbol{a}^{j_m})$ e $\mathbf{A}_B$ é denominada **matriz de base** B . Mais importante do que $\mathbf{A}_B$ é sua inversa e a ela nos referiremos com frequência.

Definição 3.5 *Dada uma m-upla $B = (j_1, j_2, \ldots, j_m)$ (lista com m componentes), tal que $(\boldsymbol{a}^{j_1}, \boldsymbol{a}^{j_2}, ..., \boldsymbol{a}^{j_m})$ formam uma base ordenada do $\mathbb{R}^m$, dizemos que $\mathbf{A}_B^{-1} = [\boldsymbol{a}^{j_1} | \boldsymbol{a}^{j_2} | \cdots | \boldsymbol{a}^{j_m}]^{-1}$ é a* **inversa da matriz associada a base ordenada cujos índices são dados por** B *, ou simplesmente, $\mathbf{A}_B^{-1}$ é a* **inversa de base** .

A razão pela qual destacamos a inversa de uma base é que a correspondente solução básica $\overline{\boldsymbol{x}}$ é obtida pelo cômputo de $\mathbf{A}_B^{-1}\boldsymbol{b}$ $(= \overline{\boldsymbol{x}})$. Mas precisamente, se $B = (j_1, \ldots, j_m)$ e $\boldsymbol{b}' := \mathbf{A}_B^{-1}\boldsymbol{b}$, temos que

$$\begin{array}{lll} \text{se } k \in B, \text{ isto é, } k = j_i \text{ para algum } i \in \{1, 2, \ldots, m\}, & \text{então} & \overline{x}_k = b'_i \ \ (= \overline{x}_{j_i}\ ), \ \text{e} \\ \text{se } k \notin B, & \text{então} & \overline{x}_k = 0. \end{array}$$

Exemplo 3.3.2 *Se considerarmos* $\mathbf{A} = \begin{bmatrix} -1 & -1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$ *e* $b = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$, *então* $B = (4, 1)$ *resulta em uma base. Calculando $\boldsymbol{b}' = \mathbf{A}_B^{-1}\boldsymbol{b}$, teremos*

$$\boldsymbol{b}' = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} -2 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} -2 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix},$$

e daí podemos obter a solução básica $\boldsymbol{x}$, fazendo

$$\begin{array}{ll} x_1 = x_{j_2} = b'_2 = 2 & (1 \in B) \\ x_2 = 0 & (2 \notin B) \\ x_3 = 0 & (3 \notin B) \\ x_4 = x_{j_1} = b'_1 = 3 & (4 \in B). \end{array}$$

Neste caso (posto $\mathbf{A} = m$), é fácil ver que o cômputo de vértices é feito através da detecção de bases, cálculo das soluções básicas utilizando as inversas de base e teste de viabilidade (não negatividade).

O algoritmo 1 para resolução do (PLC) continua aplicável se substituirmos a palavra vértice por "base viável" ou por "solução básica viável", com ressalva que a um mesmo vértice podem corresponder várias soluções básicas viáveis (para vértice degenerado).

3.4 Modelos de Algoritmo Usando a Caracterização Matricial de Vértices

No algoritmo 1, para resolver o (PLC), geramos todos os elementos de $V(X)$ sucessivamente (usando um dos algoritmos combinatórios 2, 3.1 ou 3.2). Portanto, se ao gerar o vértice $\overline{\boldsymbol{v}}$ pudessemos determinar se este é o melhor vértice, obteriamos um algoritmo que só no pior todos os vértices do poliedro.

Para tal precisamos de um teste de otimalidade que verifique se o vértice gerado é o melhor dentro todos os vértices. Isto é, dado o vértice $\overline{\boldsymbol{v}}$, queremos saber se

$$\overline{\boldsymbol{v}} \ \in \ argmax \ \{\langle \boldsymbol{c}, \boldsymbol{v} \rangle / \ \boldsymbol{v} \in V(X)\}.$$

De posse deste teste poderíamos enunciar o algoritmo abaixo:

Algoritmo 5 Algoritmo de enumeração de vértices

Passo 1: /* *Teste de viabilidade do problema* */
encontre um vértice $\boldsymbol{v}^1 \ \in \ V(X)$;

se $V(X) \equiv \emptyset$, *então:* <u>*problema inviável, PARE*</u>
faça $i \longleftarrow 1$;

Passo 2: /* *Teste de otimalidade do vertice e geração de novo vértice* */
se $\boldsymbol{v}^i \in argmax\{\langle \boldsymbol{c},\boldsymbol{v}\rangle / \; v \in V(X)\}$, *então faça:* $\boldsymbol{v}^* = \boldsymbol{v}^i$ *e vá para o passo* 3
senão
se ainda existir vértice não pesquisado, então
encontre $\boldsymbol{v}^{i+1} \in V(X)$ *não computado;*
faça $i \longleftarrow i+1$ *e volte ao passo* 2

Passo 3: /* *Cone coincide com a origem?* */
encontre $\boldsymbol{r}^i \in V(C^1)$;
se $V(C^1) \equiv \emptyset$, *então:* <u>$\boldsymbol{v}^*$ *é uma solução ótima, PARE*</u>
faça $i \longleftarrow 1$

Passo 4: /* *Semireta de ilimitação* */
se $\langle \boldsymbol{c},\boldsymbol{r}^i\rangle > 0$, *então:* $S = \{\boldsymbol{x} \in \mathrm{IR}^n / \; \boldsymbol{x} = \boldsymbol{v}^* + \lambda \boldsymbol{r}^i, \; \lambda \geqq 0\}$ <u>*é uma semireta de ilimitação: PARE*</u>

Passo 5: /* *Gera novo raio extremal* */
/* $\langle \boldsymbol{c},\boldsymbol{r}^i\rangle \leqq 0$, *pois passou pelo passo* 4 */
se $\boldsymbol{r}^i \in argmax\{\langle \boldsymbol{c},\boldsymbol{r}\rangle / \; \boldsymbol{r} \in V(C^1)\}$, *então:* <u>$\boldsymbol{v}^*$ *é uma solucão ótima do (PLC), PARE*</u>
senão
se ainda existir vértice não pesquisado, então
encontre $\boldsymbol{r}^{i+1} \in V(C^1)$ *não computado;*
faça $i \longleftarrow i+1$ *e volte ao passo* 4

Para este algoritmo, existem três questões em aberto:

(i) Como testar otimalidade?

(ii) Como achar outro vértice ainda não computado?

(iii) Será que a enumeração de bases até encontrar a primeira base viável é o melhor método para gerar o vértice inicial?

Para responder a primeira das perguntas (teste de otimalidade), vamos utilizar mais detalhadamente a caracterização matricial associada a um vértice, em particular, utilizaremos a inversa de base a ele associada. E, para tal, novamente introduziremos notações adicionais:

Notação Dado um vértice $\boldsymbol{v}$ e uma base $B = (j_1, j_2, \ldots, j_m)$ a ele associada (isto é, $I(\boldsymbol{v}) = \{j_1, j_2, \ldots, j_m\}$) denotaremos

(a) $\boldsymbol{x}_B \in \mathrm{IR}^m$ por $(\boldsymbol{x}_B)_i = \boldsymbol{x}_{j_i}$ ou $\boldsymbol{x}_B = (x_{j_1}, x_{j_2}, \ldots, x_{j_m})^t$

(b) $N = (j_{m+1}, j_{m+2}, \ldots, j_n)$ uma lista com (n-m) elementos, cujas componentes são os índices $j \notin I(\boldsymbol{v})$ (isto é, indices fora da base)

(c) $\mathbf{A}_N = \tilde{\mathbf{A}} = [\boldsymbol{a}^{j_{m+1}} | \boldsymbol{a}^{j_{m+2}} | \ldots | \boldsymbol{a}^{j_n}]$

(d) $\boldsymbol{x}_N = \tilde{\boldsymbol{x}} = (x_{j_{m+1}}, x_{j_{m+2}}, \ldots, x_{j_n})^t$ ou $\tilde{x}_k = x_{j_{m+k}}$.

Chamaremos as variáveis correspondentes aos índices da base B $(x_{j_i}, \; i \in \{1, 2, \ldots, m\}$, componentes de $\boldsymbol{x}_B)$ **variáveis na base** ou **variáveis básicas** enquanto que as correspondentes aos demais índices, N $(x_{j_i}, \; i \in \{m+1, \ldots, n\}$, componentes de $\tilde{\boldsymbol{x}})$ chamaremos de **variáveis fora da base** ou **variáveis não básicas** . A mesma denominação é dada às respectivas colunas de $\mathbf{A}$. Respeitando a ordenação dos elementos da m-upla B, construimos a matriz de base $\mathbf{A}_B$ (definição 3.5) e, analogamente, respeitando a ordenação da

$(n-m)$-upla N, construimos a matriz associada às variáveis não básicas $\mathbf{A}_N = \tilde{\mathbf{A}}$. Assim, com a conveniente reordenação das variáveis (e correspondentes colunas da matriz dos coeficientes) do sistema linear $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ temos

$$\mathbf{A}\boldsymbol{x} = \boldsymbol{b} \iff [\mathbf{A}_B|\tilde{\mathbf{A}}]\begin{bmatrix}\boldsymbol{x}_B\\ \tilde{\boldsymbol{x}}\end{bmatrix} = \boldsymbol{b} \iff \mathbf{A}_B\boldsymbol{x}_B + \tilde{\mathbf{A}}\tilde{\boldsymbol{x}} = \boldsymbol{b}.$$

Portanto,

$$\mathbf{A}\boldsymbol{x} = \boldsymbol{b} \iff \boldsymbol{x}_B = \mathbf{A}_B^{-1}\boldsymbol{b} - \mathbf{A}_B^{-1}\tilde{\mathbf{A}}\tilde{\boldsymbol{x}},$$

relação esta que nos permite escrever as variáveis básicas em função das variáveis não básicas, e que será utilizada repetidas vezes no que segue.

Exemplo 3.4.1 *Seja* $X = \left\{\boldsymbol{x} \in \mathbb{R}^5 \;/\; \left\langle \begin{array}{llll} x_1 + x_2 - x_3 & & & = 4 \\ x_1 + 3x_2 & -x_4 & & = 6 \\ x_1 + x_2 & & +x_5 & = 10 \end{array} \right\rangle \;\; e \;\; \boldsymbol{x} \geqq \mathbf{0} \right\}$

O ponto $\overline{\boldsymbol{x}} = (3,1,0,0,6)^t$ *é um vértice de* X, *pois* $\overline{\boldsymbol{x}} \in X$ *e* $\{\boldsymbol{a}^1, \boldsymbol{a}^2, \boldsymbol{a}^5\}$ *são l.i. Esta última verificação é fácil, pois adotando* $B = (5,1,2)$ *ou equivalentemente* $\boldsymbol{x}_B = (x_5, x_1, x_2)^t$, *temos*

$$\mathbf{A}_B = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 1 \end{bmatrix}, \quad \det \mathbf{A}_B = 2 \neq 0, \quad e \quad \mathbf{A}_B^{-1} = \begin{bmatrix} -1.0 & 0.0 & 1.0 \\ 1.5 & -0.5 & 0.0 \\ -0.5 & 0.5 & 0.0 \end{bmatrix}.$$

Tomando $N = (4,3)$[1], *temos*

$$\tilde{\boldsymbol{x}} = \begin{bmatrix} x_4 \\ x_3 \end{bmatrix}, \quad \tilde{\mathbf{A}} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \\ 0 & 0 \end{bmatrix}.$$

A aplicação do método matricial para cômputo de vértices corresponde a fixar as variáveis não básicas como nulas e então resolver o sistema linear $\mathbf{A}_B\boldsymbol{x}_B = \boldsymbol{b}$, *para obter as variáveis básicas. Deste modo, o vértice* $\overline{\boldsymbol{x}}$ *é obtido por*

$$\tilde{\overline{\boldsymbol{x}}} = \begin{bmatrix} \overline{x}_4 \\ \overline{x}_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\overline{\boldsymbol{x}}_B = \begin{bmatrix} \overline{x}_5 \\ \overline{x}_1 \\ \overline{x}_2 \end{bmatrix} = \mathbf{A}_B^{-1}\boldsymbol{b} = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 4 \\ 6 \\ 10 \end{bmatrix} = \begin{bmatrix} -1.0 & 0.0 & 1.0 \\ 1.5 & -0.5 & 0.0 \\ -0.5 & 0.5 & 0.0 \end{bmatrix} \begin{bmatrix} 4 \\ 6 \\ 10 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \\ 1 \end{bmatrix}$$

Utilizando as ordenações induzidas por B *e* N, *temos*

$$\begin{bmatrix} 1 & 1 & -1 & 0 & 0 \\ 1 & 3 & 0 & -1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \\ 10 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} z_5 \\ z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ -1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z_4 \\ z_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \\ 10 \end{bmatrix}$$

$$\Leftrightarrow \begin{bmatrix} z_5 \\ z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 1 \\ 1.5 & -0.5 & 0 \\ -0.5 & 0.5 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 6 \\ 10 \end{bmatrix} - \begin{bmatrix} -1 & 0 & 1 \\ 1.5 & -0.5 & 0 \\ -0.5 & 0.5 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ -1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z_4 \\ z_3 \end{bmatrix}$$

$$\Leftrightarrow \begin{bmatrix} z_5 \\ z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ 0.5 & -1.5 \\ -0.5 & 0.5 \end{bmatrix} \begin{bmatrix} z_4 \\ z_3 \end{bmatrix}$$

ou seja, podemos escrever as variáveis básicas em função das não básicas como:

[1] No exemplo poderiamos tomar $N = (3,4)$. Em geral existem (n-m)! ordenações para N.

$$\begin{cases} z_5 = & 6 - z_4 \\ z_1 = & 3 - \frac{1}{2}z_4 + \frac{3}{2}z_3 \\ z_2 = & 1 + \frac{1}{2}z_4 - \frac{1}{2}z_3 \end{cases}$$

Claramente o vértice $\overline{\boldsymbol{x}}$ é obtido fixando $z_3 = z_4 = 0$. □

Utilizando o fato que podemos exprimir as variáveis básicas em função das não básicas, lembrando que pontos viáveis do (PLC) tem componentes não negativas e que este problema é de maximização, segue

Teorema 3.2 *(***Condição suficiente de otimalidade para o (PLC)** *)*

Seja $\boldsymbol{v}$ um vértice de X, conjunto de pontos viáveis do (PLC). Então $\boldsymbol{v}$ é solução ótima do (PLC) se existe uma base B a ele associada tal que:

$$\tilde{\boldsymbol{\gamma}} = \tilde{\boldsymbol{c}} - \tilde{\mathbf{A}}^t(\mathbf{A}_B^{-1})^t\boldsymbol{c}_B \leqq \mathbf{0}.$$

Prova Como $\boldsymbol{v}$ é viável em (PLC), basta provar que:

$$\langle \boldsymbol{c}, \boldsymbol{x}\rangle \leqq \langle \boldsymbol{c}, \boldsymbol{v}\rangle, \quad \forall \boldsymbol{x} \in X := \{\boldsymbol{x} \in \mathbb{R}^n / \; \boldsymbol{Ax} = \boldsymbol{b},\; \boldsymbol{x} \geqq \mathbf{0}\}.$$

Utilizando a base B associada a $\boldsymbol{v}$ (que existe por hipótese) podemos escrever

$$\langle \boldsymbol{c}, \boldsymbol{x}\rangle = \langle \boldsymbol{c}_B, \boldsymbol{x}_B\rangle + \langle \tilde{\boldsymbol{c}}, \tilde{\boldsymbol{x}}\rangle.$$

Como para todo $\boldsymbol{x} \in X$ tem-se que $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$, então

$$\begin{aligned} \langle \boldsymbol{c}, \boldsymbol{x}\rangle &= \langle \boldsymbol{c}_B, \mathbf{A}_B^{-1}\boldsymbol{b} - \mathbf{A}_B^{-1}\tilde{\mathbf{A}}\tilde{\boldsymbol{x}}\rangle + \langle \tilde{\boldsymbol{c}}, \tilde{\boldsymbol{x}}\rangle = \langle \boldsymbol{c}_B, \mathbf{A}_B^{-1}\boldsymbol{b}\rangle + \langle \boldsymbol{c}_B, -(\boldsymbol{A}_B^{-1}\tilde{\mathbf{A}})\tilde{\boldsymbol{x}}\rangle + \langle \tilde{\boldsymbol{c}}, \tilde{\boldsymbol{x}}\rangle \\ &= \langle \boldsymbol{c}_B, \mathbf{A}_B^{-1}\boldsymbol{b}\rangle + \langle \tilde{\boldsymbol{c}} - (\mathbf{A}_B^{-1}\tilde{\mathbf{A}})^t\boldsymbol{c}_B, \tilde{\boldsymbol{x}}\rangle = \langle \boldsymbol{c}_B, \mathbf{A}_B^{-1}\boldsymbol{b}\rangle + \langle \tilde{\boldsymbol{\gamma}}, \tilde{\boldsymbol{x}}\rangle, \end{aligned}$$

lembrando que $\tilde{\boldsymbol{\gamma}} = \tilde{\boldsymbol{c}} - \tilde{\mathbf{A}}^t(\mathbf{A}_B^{-1})^t\boldsymbol{c}_B$.

O vértice $\boldsymbol{v}$ é obtido anulando as variáveis não básicas, portanto

$$\langle \boldsymbol{c}, \boldsymbol{v}\rangle = \langle \boldsymbol{c}_B, \mathbf{A}_B^{-1}\boldsymbol{b}\rangle.$$

Como, por hipótese, $\tilde{\boldsymbol{\gamma}} \leqq \mathbf{0}$ e para todo $\boldsymbol{x} \in X$, segue daí que

$$\langle \boldsymbol{c}, \boldsymbol{x}\rangle = \langle \boldsymbol{c}_B, \mathbf{A}_B^{-1}\boldsymbol{b}\rangle + \langle \tilde{\boldsymbol{\gamma}}, \tilde{\boldsymbol{x}}\rangle \leqq \langle \boldsymbol{c}_B, \mathbf{A}_B^{-1}\boldsymbol{b}\rangle = \langle \boldsymbol{c}, \boldsymbol{v}\rangle, \quad \forall \boldsymbol{x} \in X.$$ ■

É necessário notar que o teorema 3.2 fornece apenas uma condição suficiente de otimalidade. Portanto, se um vértice encontrado produz $\tilde{\boldsymbol{\gamma}} \leqq \mathbf{0}$, esta é uma solução ótima do (PLC) e pode-se interromper os demais passos do algoritmo.

Outra característica importante do teste de otimalidade é que caso não seja verificado, ele indica como encontrar um ponto "melhor" (com valor de função objetivo maior) que o vértice testado. Observando a prova, é fácil ver que se pudermos tornar positiva uma variável não básica x_k, com $\gamma_k > 0$, geraremos um novo ponto $\boldsymbol{x}'$ tal que $\langle \boldsymbol{c}, \boldsymbol{x}'\rangle > \langle \boldsymbol{c}, \boldsymbol{v}\rangle$.

Exercício 3.7 *Seja $B = (j_1, j_2, \ldots, j_m)$ uma base associada a um vértice $\boldsymbol{v}$. Se existe $(k, \lambda) \in N \times \mathbb{R}$, tal que λ é estritamente positivo e*

(i) $\gamma_k = c_k - (\mathbf{A}_B^{-1}\boldsymbol{a}^k)^t\boldsymbol{c}_B > \mathbf{0}$

(ii) $\mathbf{A}_B^{-1}\boldsymbol{b} - \lambda\mathbf{A}_B^{-1}\boldsymbol{a}^k \geqq \mathbf{0}$

então o ponto $\boldsymbol{x}'$ definido por

$$\begin{cases} x'_{j_i} &= (\mathbf{A}_B^{-1}\boldsymbol{b} - \lambda \mathbf{A}_B^{-1}\boldsymbol{a}^k)_i, \quad i \in \{1,2,\ldots,m\} \\ x'_k &= \lambda \\ x'_l &= 0, \quad \textit{se } l \notin B \textit{ e } l \neq k \end{cases}$$

é um ponto viável do (PLC) tal que $\langle \boldsymbol{c}, \boldsymbol{x}' \rangle > \langle \boldsymbol{c}, \boldsymbol{v} \rangle$. *Observe que se* $\lambda = 0$, *então* $\boldsymbol{x}' = \boldsymbol{v}$.

Dos algoritmos anteriores notamos que, fixado um vértice $\boldsymbol{x}$ (e portanto $X \neq \emptyset$) podem ocorrer três possibilidades: $\boldsymbol{v}$ é vértice ótimo; existe outro vértice "melhor" que $\boldsymbol{v}$; existe um semireta de ilimitação.

O teorema 3.2 nos fornece ferramentas para resolver o primeiro caso: uma condição suficiente de otimalidade para $\boldsymbol{v}$. E o exercício anterior, nos dá um critério para encontrar um vértice "melhor" quando a condição suficiente não for satisfeita e tivermos algumas hipóteses adicionais.

Resta determinarmos quando, a partir de um dado vértice, podemos identificar ilimitação, o que é estabelecido no fato abaixo, e em caso contrário, como gerar um vértice melhor.

Fato 3.11 *Dado* $\boldsymbol{v} \in V(X)$ *com base associada* $B = (j_1, j_2, \ldots, j_m)$, *se existir* $k \notin B$ *tal que:*

$$\begin{cases} \gamma_k = c_k - (\mathbf{A}_B^{-1}\boldsymbol{a}^k)^t \boldsymbol{c}_B > 0 \\ \boldsymbol{\pi} = \mathbf{A}_B^{-1}\boldsymbol{a}^k \leqq \mathbf{0}, \end{cases}$$

então existe uma semireta de ilimitação SI para o (PLC), dada por:

$$SI = \{\boldsymbol{x} \in \mathbb{R}^n / \ \boldsymbol{x} = \boldsymbol{v} + \lambda \overline{\boldsymbol{h}}, \quad \lambda \geqq 0\}$$

onde $\overline{\boldsymbol{h}} := \boldsymbol{e}^k - \sum_{j_i \in B} \pi_i \boldsymbol{a}^{j_i}$, *ou seja,* $\begin{cases} \overline{h}_{j_i} = -\pi_i, & i \in \{1,2,\ldots,m\} \ (e \ j_i \in B) \\ \overline{h}_k = 1 \\ \overline{h}_j = 0, & j \notin B, \ j \neq k. \end{cases}$

Prova Para provar que conjunto SI acima descrito, uma semireta com origem em $\boldsymbol{v}$ e direção $\overline{\boldsymbol{h}}$ é uma semireta de ilimitação para o (PLC), basta verificar que: $\overline{\boldsymbol{h}} \in C$ e $\langle \boldsymbol{c}, \overline{\boldsymbol{h}} \rangle > 0$.

Para provar que $\overline{\boldsymbol{h}} \in C$, basta notar que

$$\mathbf{A}\overline{\boldsymbol{h}} = \mathbf{A}_B \overline{\boldsymbol{h}}_B + \mathbf{A}_N \overline{\boldsymbol{h}}_N = \mathbf{A}_B \overline{\boldsymbol{h}}_B + \overline{h}_k \boldsymbol{a}^k = \mathbf{A}_B(-\mathbf{A}_B^{-1}\boldsymbol{a}^k) + \boldsymbol{a}^k = \mathbf{0},$$

e que $\boldsymbol{\pi} \leqq \mathbf{0}$ implica que $\overline{\boldsymbol{h}} \geqq \mathbf{0}$ (logo $\overline{\boldsymbol{h}} \in C$).

Para mostrar que $\langle \boldsymbol{c}, \overline{\boldsymbol{h}} \rangle > 0$, também usaremos a decomposição $\boldsymbol{x} = (\boldsymbol{x}_B, \boldsymbol{x}_N)^t$,

$$\begin{aligned} \langle \boldsymbol{c}, \overline{\boldsymbol{h}} \rangle &= \langle \boldsymbol{c}_B, \overline{\boldsymbol{h}}_B \rangle + \langle \tilde{\boldsymbol{c}}, \tilde{\overline{\boldsymbol{h}}} \rangle = \langle \boldsymbol{c}_B, -\boldsymbol{\pi} \rangle \ + \ c_k = c_k - \langle c_B, \mathbf{A}_B^{-1}\boldsymbol{a}^k \rangle = c_k - \langle (\mathbf{A}_B^{-1})^t c_B, \boldsymbol{a}^k \rangle \\ &= \gamma_k > 0. \end{aligned}$$

■

Ainda falta uma situação a ser analisada: quando $\boldsymbol{\gamma} \ngeqq \mathbf{0}$ e $\boldsymbol{\pi} \ngeqq \mathbf{0}$, situação que indica a existência de outra base.

Fato 3.12 *Dado* $\boldsymbol{v} \in V(X)$ *com base associada* $B = (j_1, \ldots, j_m)$, *se existir* $k \notin B$ *tal que:*

$$\begin{cases} \gamma_k = c_k - (\mathbf{A}_B^{-1}\boldsymbol{a}^k)^t c_B \ > \ 0 \\ \boldsymbol{\pi} = \mathbf{A}_B^{-1}\boldsymbol{a}^k \ \ngeqq \ 0, \end{cases}$$

então existe uma base viável B' *diferente de* B *em um único índice, tal que:*

(i) ou B' *define um vértice* v': $\langle \boldsymbol{c}, \boldsymbol{v}' \rangle > \langle \boldsymbol{c}, \boldsymbol{v} \rangle$;

(ii) ou B' *é outra base associada ao vértice* $\boldsymbol{v}$ *(neste caso* $\boldsymbol{v}$ *é vértice degenerado).*

Prova Definindo $\overline{\boldsymbol{h}}$ como no fato anterior, isto é, $\overline{\boldsymbol{h}} := \boldsymbol{e}^k - \sum_{j_i \in B} \pi_i \boldsymbol{a}^{j_i}$, temos que, como $\mathbf{A}\overline{\boldsymbol{h}} = \mathbf{0}$ e $\boldsymbol{b} = \mathbf{A}\boldsymbol{v} = \mathbf{A}_B \boldsymbol{v}_B + \mathbf{A}_N \boldsymbol{v}_N = \mathbf{A}_B \boldsymbol{v}_B$ $(\boldsymbol{v}_B = \mathbf{A}_B^{-1}\boldsymbol{b} \geqq \mathbf{0})$. Definindo $\boldsymbol{x}(\lambda) := \boldsymbol{v} + \lambda \overline{\boldsymbol{h}}$, para $\lambda \geqq 0$, temos que:

$$\begin{aligned}
\boldsymbol{x}(\lambda) = \boldsymbol{v} + \lambda\overline{\boldsymbol{h}} \in X \quad &\Longleftrightarrow \boldsymbol{x}(\lambda) = \boldsymbol{x}_B(\lambda) + \boldsymbol{x}_N(\lambda) \geqq \mathbf{0} \\
&\Longleftrightarrow \boldsymbol{v}_B + \lambda\overline{\boldsymbol{h}}_B + \boldsymbol{v}_N + \lambda\overline{\boldsymbol{h}}_N = \mathbf{A}_B^{-1}\boldsymbol{b} - \lambda\boldsymbol{\pi}_B + \mathbf{0} + \lambda\boldsymbol{e}^k = \mathbf{A}_B^{-1}\boldsymbol{b} - \lambda\boldsymbol{\pi} + \lambda\boldsymbol{e}^k \geqq \mathbf{0} \\
&\Longleftrightarrow \lambda \in [0, \overline{\lambda}], \quad \text{onde } \overline{\lambda} = \min_{\pi_i > 0} \frac{(\mathbf{A}_B^{-1}b)_i}{\pi_i} = \min_{\pi_i>0} \frac{b'_i}{\pi_i} \geqq 0 \quad (\lambda\boldsymbol{e}^k \geq 0,\ \forall \lambda > 0).
\end{aligned}$$

Sejam l tal que $\frac{b'_l}{\pi_l} = \overline{\lambda}$ e $B' := (j'_1, \ldots, j'_m)$ $(= B \setminus \{j_l\} \cup \{k\})$, ou seja, $\begin{cases} j'_i = j_i, & \text{se } i \neq l \\ j'_l = k. \end{cases}$

Claramente $I(\boldsymbol{x}(\overline{\lambda})) \subseteq B'$. Portanto, se B' formar uma base, esta será viável. Suponhamos por contradição, que $\{\boldsymbol{a}^{j'_i}\}_{i=1}^m = \{\boldsymbol{a}^j\}_{j\in B}$ seja l.d. Como $\{\boldsymbol{a}^{j_i}\}_{i=1}^m$ são l.i. e $\{\boldsymbol{a}^{j'_i}\}_{i=1}^m = \{\boldsymbol{a}^{j_i}\}_{j_i\in B, i\neq l} \cup \{\boldsymbol{a}^k\}$ l.d., então $\boldsymbol{a}^k$ pode ser escrito como combinação linear de $\{\boldsymbol{a}^{j_i}\}_{j_i \in B, i\neq l}$. Logo, existem $\{\alpha_i /\ i \in \{1, 2, \ldots, m\}, i \neq l\}$, não todos nulos, tais que

$$\boldsymbol{a}^{j'_l} = \boldsymbol{a}^k = \sum_{i=1, i\neq l}^m \alpha_i \boldsymbol{a}^{j_i} = \sum_{i=1}^m \alpha_i \boldsymbol{a}^{j_i}, \quad \text{com} \quad \alpha_l = 0.$$

Isto é, existe $\boldsymbol{\alpha} \in \mathbb{R}^m$, para o qual $\mathbf{A}_B\boldsymbol{\alpha} = \boldsymbol{a}^k$, com $\alpha_l = 0$. Como $\mathbf{A}_B$ é não singular, $\boldsymbol{\alpha} = \mathbf{A}_B^{-1}\boldsymbol{a}^k = \boldsymbol{\pi}$, com $\alpha_l = 0$. Portanto $\pi_l = 0$, o que resultaria em uma contradição, daí $\{\boldsymbol{a}^{j'_i}\}_{i=1}^m$ é l.i.

Assim $\boldsymbol{x}(\overline{\lambda})$ é uma solução básica viável do (PLC), ou seja, $\boldsymbol{x}(\overline{\lambda})$ é um vértice de X. Em resumo:

- Se $\overline{\lambda} = 0$, então $\boldsymbol{x}(\overline{\lambda}) = \boldsymbol{v}$ e B' é uma nova base associada ao vértice $\boldsymbol{v}$;
- Se $\overline{\lambda} > 0$, $\boldsymbol{x}(\overline{\lambda})$ é um vértice tal que
  $\langle \boldsymbol{c}, \boldsymbol{x}(\overline{\lambda})\rangle \ = \ \langle \boldsymbol{c}, \boldsymbol{v}\rangle + \lambda_k\overline{\lambda} > \langle \boldsymbol{c}, \boldsymbol{v}\rangle.$ ∎

Índice Remissivo