MAC-315 - Programação Linear

Prof. Marcelo Queiroz

http://www.ime.usp.br/~mqz

Notas de Aula[1]

2 Geometria e Programação Linear

2.1 Poliedros e conjuntos convexos

Alguns conceitos geométricos importantes no estudo de programação linear são introduzidos a seguir.

Definição 2.1 *$P \subset \mathbb{R}^n$ é um* **poliedro** *se existem $A \in \mathbb{R}^{m\times n}$ e $b \in \mathbb{R}^m$ tais que $P = \{x \in \mathbb{R}^n \mid Ax \geq b\}$.*

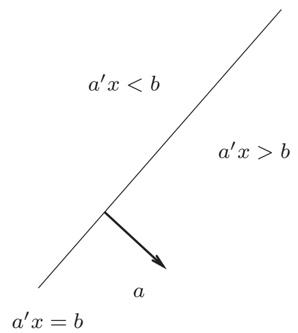
Como já vimos, qualquer problema de programação linear pode ser escrito na *forma geral*, onde o conjunto de pontos viáveis é expresso por $Ax \geq b$; assim o conjunto viável de qualquer PL é um poliedro. O conjunto $\{x \in \mathbb{R}^n \mid Ax = b,\ x \geq 0\}$ define um tipo especial de poliedro que chamaremos de *poliedro na forma canônica*.

Definição 2.2 *Um conjunto $S \subset \mathbb{R}^n$ é* **limitado** *se existe $K \in \mathbb{R}_+$ tal que $|x_i| < K,\ \forall x \in S$.*

Exercício 2.1 *Mostre que $S \subset \mathbb{R}^n$ é limitado se, e somente se, existe $K \in \mathbb{R}_+$ tal que $\|x\| < K,\ \forall x \in S$.*

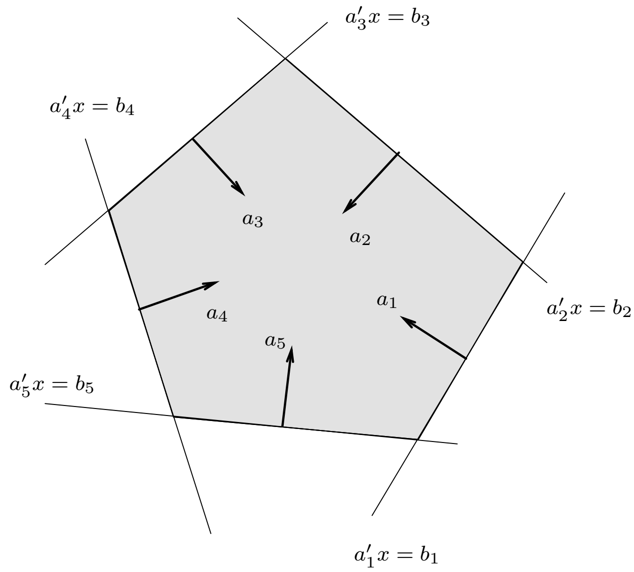
Definição 2.3 *Um* **hiperplano** *é um conjunto da forma $\{x \in \mathbb{R}^n \mid a'x = b\}$, e um* **semi-espaço** *é um conjunto da forma $\{x \in \mathbb{R}^n \mid a'x \geq b\}$, onde $a \in \mathbb{R}^n$ e $b \in \mathbb{R}$.*

Exercício 2.2 *Mostre que a é ortogonal a qualquer direção interna ao hiperplano $\{x \mid a'x = b\}$.*

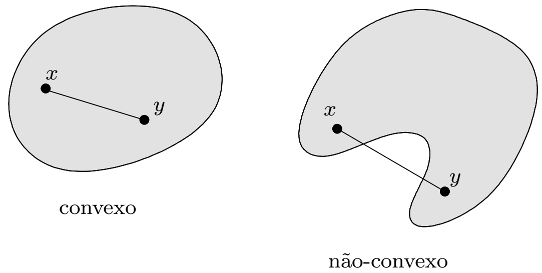


Note que hiperplanos e semi-espaços são casos particulares de poliedros, e que todo poliedro é uma intersecção finita de semi-espaços.

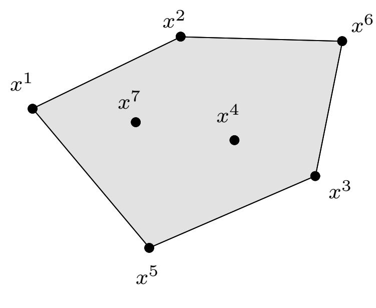
[1] Baseadas no livro de Bertsimas & Tsitsiklis: *Introduction to Linear Optimization*.



Definição 2.4 *$S \subset \mathbb{R}^n$ é* **convexo** *se $\forall x, y \in S$, $\forall \lambda \in [0,1]$, $\lambda x + (1-\lambda)y \in S$.*



Definição 2.5 *Sejam $x^1, \ldots, x^k \in \mathbb{R}^n$ e $\lambda_1, \ldots, \lambda_k \in \mathbb{R}_+$ tais que $\lambda_1 + \cdots + \lambda_k = 1$. O vetor $\sum_{i=1}^k \lambda_i x^i$ é dito uma* **combinação convexa** *dos vetores $x^1, \ldots, x^k$. O* **casco convexo** *dos vetores $x^1, \ldots, x^k$ é o conjunto de todas as combinações convexas destes vetores.*



Teorema 2.1

1. *A intersecção de conjuntos convexos é convexa.*

2. *Todo poliedro é convexo.*

3. *Um conjunto convexo é fechado por combinações convexas.*

4. *O casco convexo de um conjunto finito de vetores é convexo.*

Prova.

1. Sejam $x, y \in S = \bigcap_{i \in I} S_i \subset \mathbb{R}^n$ e $\lambda \in [0,1]$. Então $x, y \in S_i$, $\forall i \in I$ e, como S_i é convexo, $\lambda x + (1-\lambda)y \in S_i$, $\forall i \in I$, ou seja, $\lambda x + (1-\lambda)y \in S$.

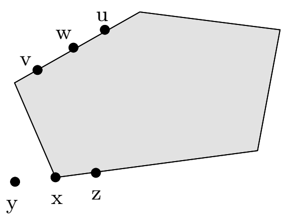
2. Vamos mostrar que todo semi-espaço é convexo: seja $S = \{x \in \mathbb{R}^n \mid a'x \geq b\}$ um semi-espaço, $x, y \in S$ e $\lambda \in [0,1]$. Então $a'(\lambda x + (1-\lambda)y) = \lambda a'x + (1-\lambda)a'y \geq \lambda b + (1-\lambda)b = b$ (usando $a'x \geq b, a'y \geq b, \lambda, (1-\lambda) \geq 0$) e assim $\lambda x + (1-\lambda)y \in S$; logo S é convexo. Pela parte (1) toda intersecção finita de semi-espaços (ou seja, todo poliedro) é convexa(o).

3. A afirmação é que se S é convexo, $x^1, \ldots, x^k \in S$ e $\lambda_1, \ldots, \lambda_k \in \mathbb{R}_+$ satisfazem $\lambda_1 + \cdots + \lambda_k = 1$, então $\sum_{i=1}^k \lambda_i x^i \in S$. O caso $k = 2$ é a própria definição de convexidade; considere então que o resultado vale para k e vamos provar que vale também para $k+1$, completando a prova por indução. Sejam $x^1, \ldots, x^{k+1} \in S$ e $\lambda_1, \ldots, \lambda_{k+1} \in \mathbb{R}_+$ tais que $\lambda_1 + \cdots + \lambda_{k+1} = 1$. Suponha que $\lambda_{k+1} \neq 1$ (senão teríamos $\sum_{i=1}^{k+1} \lambda_i x^i = x^{k+1} \in S$); então
$$\sum_{i=1}^{k+1} \lambda_i x^i = \lambda_{k+1} x^{k+1} + (1-\lambda_{k+1}) \sum_{i=1}^{k} \frac{\lambda_i}{1-\lambda_{k+1}} x^i.$$
Como $\frac{\lambda_i}{1-\lambda_{k+1}} \geq 0$ e $\sum_{i=1}^k \frac{\lambda_i}{1-\lambda_{k+1}} = \frac{\sum_{i=1}^k \lambda_i}{1-\lambda_{k+1}} = \frac{1-\lambda_{k+1}}{1-\lambda_{k+1}} = 1$, temos que $y = \sum_{i=1}^k \frac{\lambda_i}{1-\lambda_{k+1}} x^i \in S$, e por convexidade, $\sum_{i=1}^{k+1} \lambda_i x^i = \lambda_{k+1} x^{k+1} + (1-\lambda_{k+1}) y \in S$.

4. Seja S o casco convexo de $x^1, \ldots, x^k$ e $y, z \in S$. Então $y = \sum_{i=1}^k \mu_i x^i$ e $z = \sum_{i=1}^k \nu_i x^i$, com $\mu, \nu \geq 0$ e $\sum_{i=1}^k \mu_i = \sum_{i=1}^k \nu_i = 1$. Seja $\lambda \in [0,1]$; então $\lambda y + (1-\lambda)z = \lambda(\sum_{i=1}^k \mu_i x^i) + (1-\lambda)(\sum_{i=1}^k \nu_i x^i) = \sum_{i=1}^k (\lambda\mu_i + (1-\lambda)\nu_i) x^i$. Como $\lambda\mu_i + (1-\lambda)\nu_i \geq 0$, $\forall i$ e $\sum_{i=1}^k \lambda\mu_i + (1-\lambda)\nu_i = \lambda \sum_{i=1}^k \mu_i + (1-\lambda)\sum_{i=1}^k \nu_i = \lambda + (1-\lambda) = 1$, segue que $\lambda y + (1-\lambda)z \in S$. Logo S é convexo.

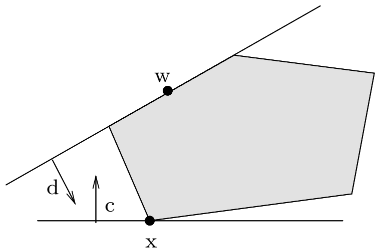
2.2 Pontos extremos, vértices e soluções básicas viáveis

Definição 2.6 *Seja P um poliedro. $x \in P$ é um* **ponto extremo** *de P se não existem $y, z \in P$, $y \neq z$ e $\lambda \in (0,1)$ tais que $x = \lambda y + (1-\lambda)z$.*



Exercício 2.3 *Mostre que a condição* $\left[\nexists y, z \in P,\ y \neq z \ t.q.\ x = \frac{1}{2}(y+z)\right]$ *é equivalente à da definição acima.*

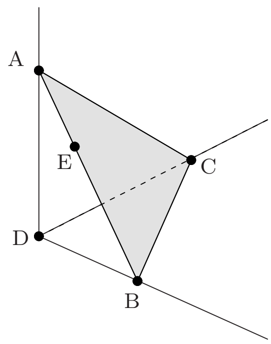
Definição 2.7 *Seja* $P \subset \mathbb{R}^n$ *um poliedro.* $x \in P$ *é um* **vértice** *de* P *se existe* $c \in \mathbb{R}^n$ *t.q.* $c'x < c'y,\ \forall y \in P \backslash \{x\}$.



Procuramos traduzir os conceitos expressos pelas definições 2.6 e 2.7 para um contexto algébrico, que utilize a representação do poliedro e que permita uma verificação computacional. Para isso considere um poliedro $P \subset \mathbb{R}^n$ descrito pelas igualdades e desigualdades a seguir:

$$\begin{aligned} a_i'x &\geq b_i, \quad i \in M_1 \\ a_i'x &\leq b_i, \quad i \in M_2 \\ a_i'x &= b_i, \quad i \in M_3, \end{aligned}$$

onde M_1, M_2 e M_3 são conjuntos (finitos) de índices.
Definição 2.8 *Se* x^* *satisfaz* $a_i'x^* = b_i$ *para* $i \in \bigcup_{k=1}^{3} M_k$ *dizemos que esta restrição é* **ativa** *em* x^*.



$P = \{x \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 1,\ x \geq 0\}$

Note que podemos associar a cada ponto do poliedro um conjunto de restrições ativas naquele ponto. Estamos interessados em considerar conjuntos de restrições ativas que estão associados a um único ponto do poliedro. Esta unicidade está associada à independência linear dos vetores que definem as restrições, como veremos a seguir.

Teorema 2.2 *Seja* $x^* \in \mathbb{R}^n$ *e* $I = \{i \in \bigcup_{k=1}^{3} M_k \mid a_i' x^* = b_i\}$. *Então são equivalentes:*

1. *Existem* n *vetores linearmente independentes dentre os* $a_i,\ i \in I$;

2. *O espaço gerado por* $\{a_i \mid i \in I\}$ *é o* $\mathbb{R}^n$;

3. *O sistema de equações* $a_i' x = b_i,\ i \in I$ *tem solução única.*

Prova.

- $(1 \Rightarrow 2)$ Como existem n vetores a_i linearmente independentes, o espaço gerado por eles tem dimensão $\geq n$. Como este espaço gerado é subconjunto do $\mathbb{R}^n$ e o único subespaço do $\mathbb{R}^n$ de dimensão n é o próprio $\mathbb{R}^n$, segue a implicação.

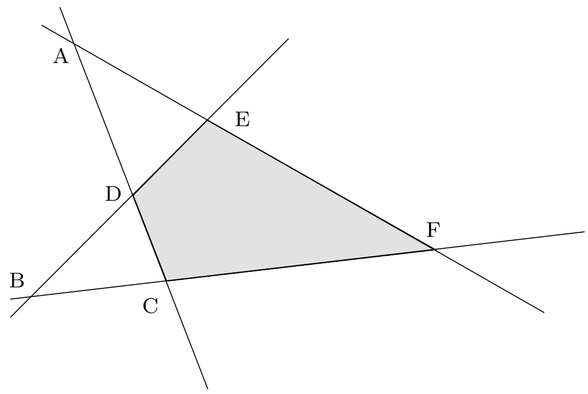
- $(2 \Rightarrow 3)$ Suponha por contradição que exista outra solução $y^* \neq x^*$ para o sistema $a_i' x = b_i,\ i \in I$. Então $x^* - y^*$ é ortogonal a todos os $a_i,\ i \in I$, e portanto não pode ser gerado por eles. Isso contradiz (2).

- $(3 \Rightarrow 1)$ Suponha por contradição que não existam n vetores linearmente independentes dentre os $a_i,\ i \in I$; então existe $d \neq 0$ ortogonal a todos os $a_i,\ i \in I$ e portanto se x^* é solução de $a_i' x = b_i,\ i \in I$, $x^* + d$ também é solução deste sistema, contrariando (3).

Frequentemente diremos que certas restrições são *linearmente independentes* querendo dizer que os vetores a_i que as definem são linearmente independentes. Note que se x^* é solução do sistema $a_i' x = b_i,\ i \in I$ não necessariamente é verdade que x^* seja viável; outras restrições (com índices fora do conjunto I) poderão ser violadas. Isso nos leva à seguinte definição.

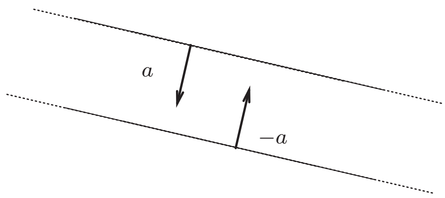
Definição 2.9 *Seja* P *um poliedro e* $x^* \in \mathbb{R}^n$. x^* *é dito* **solução básica** *se todas as restrições de igualdade são satisfeitas e além disso dentre as restrições ativas existem* n *linearmente independentes. Uma* **solução básica viável** *é uma solução básica que satisfaz todas as restrições, inclusive as inativas.*

Esta distinção entre igualdades e desigualdades faz com que a definição acima seja *dependente* da representação do poliedro: se trocássemos todas as igualdades por pares de desigualdades, novas soluções básicas (inviáveis) surgiriam.



Soluções básicas (A e B são inviáveis)

Observe que se um poliedro possui menos restrições do que variáveis ($m < n$), então o número de restrições ativas em qualquer ponto é também menor do que n e portanto não existem soluções básicas. Isso também ocorre quando $m \geq n$ mas existem menos do que n restrições linearmente independentes:



O teorema a seguir relaciona as definições anteriores.

Teorema 2.3 *Seja P um poliedro e $x^* \in P$. São equivalentes:*

1. *x^* é um vértice;*
2. *x^* é um ponto extremo;*
3. *x^* é uma solução básica viável.*

Prova.

Sem perda de generalidade considere que P é descrito pelas desigualdades $a_i'x \geq b_i$ e igualdades $a_i'x = b_i$.

- ($1 \Rightarrow 2$) Suponha que $x^* \in P$ é vértice, ou seja, que existe $c \in \mathbb{R}^n$ tal que $c'x^* < c'y$, $\forall y \in P \backslash \{x^*\}$. Para quaisquer $y, z \in P$, $y \neq z$ e $\lambda \in (0, 1)$ temos que $c'(\lambda y + (1 - \lambda)z) = \lambda c'y + (1 - \lambda)c'z > \lambda c'x^* + (1 - \lambda)c'x^* = c'x^*$ (a desigualdade se deve a ou $x^* \neq y$ ou $x^* \neq z$), assim $\lambda y + (1 - \lambda)z \neq x^*$; ou seja, x não pode ser expresso como combinação convexa de outros pontos de P (definição de ponto extremo)

- $(2 \Rightarrow 3)$ Suponha por contradição que $x^* \in P$ não seja uma solução básica viável; vamos mostrar que x^* não pode ser ponto extremo. Seja $I = \{i \mid a_i' x^* = b_i\}$. Como x^* não é solução básica viável, existem menos de n dentre os a_i, $i \in I$ linearmente independentes, e portanto existe um $d \neq 0$ tal que $a_i' d = 0$, $\forall i \in I$. Como todas as restrições inativas (com índices fora de I) satisfazem $a_i' x^* > b_i$, existe um $\epsilon > 0$ tal que o intervalo $[x^* - \varepsilon d, x^* + \varepsilon d] \subset P$ e, em particular, $x^* = \frac{1}{2}(x^* - \varepsilon d) + \frac{1}{2}(x^* + \varepsilon d)$, com $x^* - \varepsilon d \neq x^* + \varepsilon d$. Isso contradiz a extremalidade de x^*.
- $(3 \Rightarrow 1)$ Seja $c = \sum_{i \in I} a_i$, onde $I = \{i \mid a_i' x^* = b_i\}$. Então para qualquer $y \in P$,

  $$c' y = \sum_{i \in I} a_i' y \geq \sum_{i \in I} b_i = \sum_{i \in I} a_i' x^* = c' x^*.$$

  Além disso, x^* é a única solução do sistema $a_i' x = b_i$, $i \in I$, ou seja, $\forall y \in P \backslash \{x^*\}$ vale $c' y > c' x^*$, e portanto x^* é vértice.

Note que a definição de vértice e de ponto extremo independe da representação do poliedro; pelo teorema anterior concluimos que a definição de solução básica viável também não depende da representação, muito embora a definição de solução básica dependa. Note também que a definição algébrica (solução básica viável) pode ser implementada computacionalmente: pode-se testar se um conjunto de vetores é linearmente independente utilizando o método da triangularização. Outra propriedade importante é:

Corolário 2.1 *Dado um número finito de desigualdades e igualdades lineares, o número de soluções básicas (e consequentemente de vértices) do poliedro correspondente é finito.*

Prova.

> Basta notar que o número de subconjuntos contendo n restrições linearmente independentes é finito, e cada um está associado a no máximo uma solução básica x^* (no sentido de corresponder exatamente ao conjunto de restrições ativas em x^*).

Apesar de finito, o número de vértices de um poliedro pode ser muito alto mesmo que a descrição do poliedro seja "pequena": o hipercubo unitário em $\mathbb{R}^n$ é descrito por $2n$ desigualdades ($x_i \geq 0$ e $x_i \leq 1$, $i = 1, \ldots, n$) e no entanto todos os 2^n pontos da forma $(x_1, \ldots, x_n)'$ com $x_i \in \{0, 1\}$ são vértices do hipercubo.

Outro conceito importante associado aos vértices é o de vizinhança ou adjacência:

Definição *Duas soluções básicas são ditas* **adjacentes** *se existem* $n-1$ *restrições linearmente independentes que sejam ativas em ambas as soluções. Se duas soluções básicas adjacentes são viáveis, o segmento de reta que as une é chamado de* **aresta** *do poliedro.*

Exercício 2.4 *Mostre que todos os pontos da forma* $(x_1, \ldots, x_n)'$ *com* $x_i \in \{0, 1\}$ *são vértices do hipercubo unitário em* $\mathbb{R}^n$*. Mostre que dois vértices* $(x_1, \ldots, x_n)'$ *e* $(y_1, \ldots, y_n)'$ *são adjacentes se, e somente se, existe um* $j \in \{1, \ldots, n\}$ *tal que* $x_i = y_i,\ \forall i \neq j$ *e* $x_j = 1 - y_j$*.*

2.3 Poliedros na forma canônica

Vamos especializar as definições e resultados da seção anterior para poliedros da forma $P = \{x \in \mathbb{R}^n \mid Ax = b,\ x \geq 0\}$; estes resultados serão muito importantes para o desenvolvimento do método simplex.

Frequentemente faremos a hipótese fundamental de que a matriz $A \in \mathbb{R}^{m \times n}$ possui linhas linearmente independentes e, em particular, que $m \leq n$. Veremos mais tarde que isso não acarreta perda de generalidade na discussão, pois se $P \neq \emptyset$ então linhas linearmente dependentes de A correspondem a restrições supérfluas e que podem ser removidas.

Note que toda solução básica de P precisa satisfazer as restrições $Ax = b$ por definição; isso fornece m restrições linearmente independentes, de acordo com a hipótese fundamental sobre A. Para obter um vértice precisamos obter mais $n - m$ restrições ativas, ou seja, variáveis $x_i = 0$ (tais que a restrição $x_i \geq 0$ fique ativa). A necessidade de que o conjunto de restrições ativas resultante seja linearmente independente nos fornece a seguinte caracterização:

Teorema 2.4 *Considere* $P = \{x \in \mathbb{R}^n \mid Ax = b,\ x \geq 0\}$ *e que* $A \in \mathbb{R}^{m \times n}$ *possui linhas linearmente independentes. Então* $x \in \mathbb{R}^n$ *é solução básica se e somente se* $Ax = b$ *e existem índices* $b_1, \ldots, b_m$ *tais que as colunas* $A^{b_1}, \ldots, A^{b_m}$ *são linearmente independentes e* $x_i = 0,\ \forall i \notin \{b_1, \ldots, b_m\}$*.*

Prova.

Inicialmente suponha que $x \in \mathbb{R}^n$ satisfaz $Ax = b$ e existem índices $b_1, \ldots, b_m$ tais que as colunas $A^{b_1}, \ldots, A^{b_m}$ são linearmente independentes e $x_i = 0,\ \forall i \notin \{b_1, \ldots, b_m\}$. O sistema linear $\sum_{i=1}^m A^{b_i} x_{b_i} = b$ possui solução única, pois a matriz $\left[A^{b_1} \cdots A^{b_m}\right]$ é inversível. Então o sistema $Ax = b,\ x_i = 0,\ \forall i \notin \{b_1, \ldots, b_m\}$ possui solução única, visto que substituindo $x_i = 0$ em $Ax = b$ obtemos $Ax = \sum_{i=1}^{n} A^i x_i = \sum_{i=1}^{m} A^{b_i} x_{b_i} = b$. Logo x^* é solução básica.

Por outro lado, considere que x^* é solução básica e seja $I = \{i \mid x_i \neq 0\} = \{b_1, \ldots, b_k\}$. Por definição, x^* é a única solução do sistema $Ax = b,\ x_i = 0,\ \forall i \notin I$, e o sistema equivalente $\sum_{i=1}^{k} A^{b_i} x_{b_i} = b$ possui solução única. Então as colunas $A^{b_1}, \ldots A^{b_k}$ são linearmente independentes; em particular $k \leq m$. Como posto$(A) = \dim(\text{col}(A)) = \dim(\text{lin}(A)) = m$, podemos obter $m-k$ colunas $A^{b_{k+1}}, \ldots, A^{b_m}$ de tal forma que $A^{b_1}, \ldots, A^{b_m}$ são linearmente independentes. Além disso $\forall i \notin \{b_1, \ldots, b_m\}$ temos que $i \notin \{b_1, \ldots, b_k\} = I$, e portanto $x_i = 0$.

Exercício 2.5 *Mostre que* $x^* \in P = \{x \in \mathbb{R}^n \mid Ax = b,\ x \geq 0\}$ *é um vértice de* P *se e somente se as colunas* $\{A^i \mid x_i^* \neq 0\}$ *são linearmente independentes. Dica: separe em 2 casos; 1 - A possui linhas linearmente independentes e 2 - A possui linhas linearmente dependentes.*

Com o resultado acima podemos construir todos os vértices de um poliedro usando o algoritmo abaixo.

Algoritmo para construir vértices

1. Escolha m índices $I = \{b_1, \ldots, b_m\}$ tais que as colunas $A^{b_1}, \ldots, A^{b_m}$ sejam linearmente independentes;

2. Faça $x_i = 0,\ \forall i \notin I$;

3. Resolva o sistema $\sum_{i \in I} A^i x_i = b$;

4. Teste se a solução x^* satisfaz $x^* \geq 0$.

Veja que no passo 4 a viabilidade da solução se reduz ao teste $x^* \geq 0$, pois o fato dela satisfazer $\sum_{i \in I} A^i x_i^* = b$ e $x_i^* = 0,\ \forall i \notin I$ garantem que $Ax^* = b$. Se x^* é uma solução básica associada aos índices $\{b_1, \ldots, b_m\}$, as variáveis $x_{b_1}, \ldots, x_{b_m}$ são chamadas **variáveis básicas** e as restantes são chamadas **não-básicas**. As colunas $A^{b_1}, \ldots, A^{b_m}$ são chamadas **colunas básicas** e como são linearmente independentes formam uma **base** do $\mathbb{R}^m$. Chamamos de **matriz básica** à matriz

$$B = \left[A^{b_1} \cdots A^{b_m}\right]$$

e o correspondente vetor básico é $x_B = (x_{b_1}, \ldots, x_{b_m})'$. O valor das variáveis básicas é obtido resolvendo-se o sistema $Bx_B = b$, cuja única solução é $x_B = B^{-1}b$.

Exemplo 2.1 *Seja o poliedro* $P = \{x \in \mathbb{R}^7 \mid Ax = b,\ x \geq 0\}$ *onde*

$$A = \begin{bmatrix} 1 & 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & 1 & 6 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 8 \\ 12 \\ 4 \\ 6 \end{bmatrix}.$$

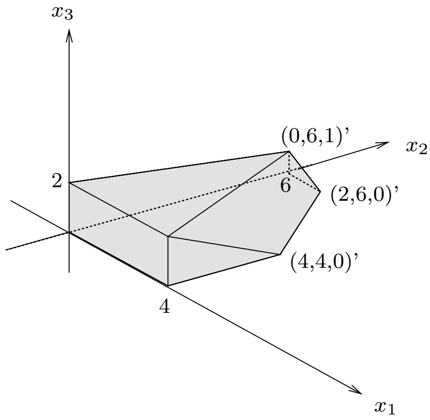
Note que a matriz A *possui uma submatriz identidade, cujas colunas são linearmente independentes. Considerando as colunas* A^4, A^5, A^6, A^7 *como básicas temos a solução associada* $x = (0, 0, 0, 8, 12, 4, 16)$ *que é um vértice (por ser não-negativa). Tomando as colunas* A^3, A^5, A^6, A^7 *e resolvendo o sistema correspondente teremos a solução* $x = (0, 0, 4, 0, -12, 4, 6)$ *que é uma solução básica inviável.*

Suponha que a matriz possuisse uma coluna A^8 *idêntica à* A^7*; então os conjuntos* $\{A^3, A^5, A^6, A^7\}$ *e* $\{A^3, A^5, A^6, A^8\}$ *seriam idênticos. Porém as duas bases (associadas aos conjuntos de índices* $\{3, 5, 6, 7\}$ *e* $\{3, 5, 6, 8\}$*) são distintas. Note que as soluções associadas seriam* $x = (0, 0, 4, 0, -12, 4, 6, 0)$ *e* $y = (0, 0, 4, 0, -12, 4, 0, 6)$*, que são distintas.*

Por outro lado, ainda no problema original, considere a solução $x = (4, 0, 2, 0, 0, 0, 6)$*. Verifique que ela está associada a quatro bases distintas, formadas pelos conjuntos* $\{A^1, A^2, A^3, A^7\}$*,* $\{A^1, A^3, A^4, A^7\}$*,* $\{A^1, A^3, A^5, A^7\}$ *e* $\{A^1, A^3, A^6, A^7\}$*. Repare ainda que esta solução está associada a 8 restrições ativas, de onde se podem extrair também quatro subconjuntos de 7 restrições linearmente independentes (garantindo solução única do sistema de equações associado).*

O poliedro acima pode ser visto como uma representação alternativa para o poliedro em $\mathbb{R}^3$ *abaixo, e toda a discussão acima (traduzindo a idéia de "base" para "subconjuntos de 3 restrições ativas") pode ser adaptada para este contexto com a ajuda do desenho que segue.*

$$Q = \left\{ x \in \mathbb{R}^3 \;\middle|\; \begin{array}{ccccccc} x_1 & + & x_2 & + & 2x_3 & \leq & 8 \\ & & x_2 & + & 6x_3 & \leq & 12 \\ x_1 & & & & & \leq & 4 \\ & & x_2 & & & \leq & 6 \\ & & x_1, x_2, x_3 \geq 0 & & & & \end{array} \right\}.$$



Correspondência entre bases e soluções básicas

Como vimos anteriormente, soluções básicas distintas correspondem a bases distintas, pois uma base está associada a um sistema de equações que possui solução única. No entanto duas bases diferentes podem corresponder à mesma solução básica (que satisfaz os dois sistemas de equações); além do exemplo que já vimos, considere o caso extremo em que $b = 0$ e todas as bases estão associadas à mesma solução viável ($x = 0$). A consideração deste fenômeno é muito importante do ponto de vista computacional, e está associado ao conceito de *degenerescência* que veremos na próxima seção.

Soluções básicas adjacentes e bases adjacentes

Lembremos que duas soluções básicas distintas são ditas adjacentes se elas possuem $n - 1$ restrições ativas linearmente independentes em comum. Em problemas na forma canônica, dizemos que duas bases são adjacentes se elas possuem $m - 1$ colunas em comum (diferem em apenas uma coluna).

Exercício 2.6 *Prove que duas soluções básicas adjacentes tem sempre associadas duas bases adjacentes. Prove que se duas bases adjacentes produzem soluções básicas distintas, então estas soluções são adjacentes.*

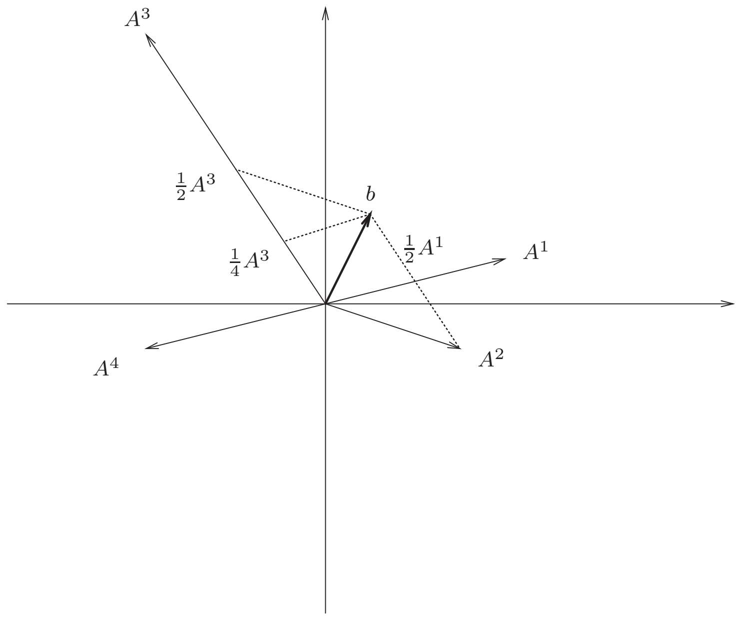
Exemplo 2.2 *No exemplo 2.1 as bases* $\{A^4, A^5, A^6, A^7\}$ *e* $\{A^3, A^5, A^6, A^7\}$ *são adjacentes pois diferem em apenas uma coluna. As soluções básicas correspondentes* $x = (0, 0, 0, 8, 12, 4, 6)'$ *e* $y = (0, 0, 4, 0, -12, 4, 6)'$ *também são adjacentes: o problema está em* $\mathbb{R}^7$ *e elas possuem 6 restrições ativas linearmente independentes em comum (*$x_1 \geq 0, x_2 \geq 0$ *e as 4 restrições de igualdade). No desenho em* $\mathbb{R}^3$ *estas soluções correspondem a* $x = (0, 0, 0)'$ *e* $y = (0, 0, 4)'$ *(esta última é inviável e não aparece no desenho). Verifique as colunas básicas e restrições ativas em comum para outros pares de soluções adjacentes daquele desenho.*

Para ter uma intuição geométrica de soluções básicas podemos utilizar uma outra interpretação do problema, que corresponde a encontrar coeficientes x_i que mostrem que b é combinação linear dos vetores A^i ($b = \sum_{i=1}^{n} A^i x_i = Ax$). Numa solução básica isso é atingido utilizando apenas m vetores linearmente independentes; para a solução básica ser viável, todos os coeficientes x_i têm que ser positivos. Lembrando de nossos problemas-exemplo do capítulo 1, as colunas de A estão normalmente associadas ao modo como as variáveis de decisão se relacionam com as exigências do problema, e b representa um atendimento ideal (ou mínimo ou máximo) destas exigências. Assim A^i representa como o computador i utiliza as matérias-primas UCP, memória, leitor de discos, e b representa as quantidades máximas disponíveis destas matérias primas (problemas da DEC e problema de produção); A^i representa quanto o alimento i fornece de cada nutriente e b a quantidade de cada nutriente numa dieta balanceada (ou ideal) (problema da dieta); A^i representa a rotina semanal de trabalho dos enfermeiros que começam no dia i e b é o quadro funcional mínimo do hospital ao longo da semana (problema do plantão). Representando as colunas A^i e o vetor b graficamente teremos um esquema dos recursos e exigências do problema e poderemos obter soluções básicas utilizando m (ou menos) recursos.

Exemplo *Considere o poliedro* $P = \{x \in \mathbb{R}^4 \mid Ax = b,\ x \geq 0\}$ *onde*

$$A = \begin{bmatrix} 4 & 3 & -4 & -4 \\ 1 & -1 & 6 & -1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

As colunas A^i *e o vetor* b *podem ser representados como segue:*



Note que A^1 *e* A^2 *são linearmente independentes, mas a solução básica correspondente é inviável pois será necessária uma quantidade* x_2 *negativa para escrever* $b = x_1A^1 + x_2A^2$*; também serão inviáveis as bases* *associadas a* $\{A^2, A^4\}$ *e* $\{A^3, A^4\}$*. Os vetores* A^1 *e* A^3 *são linearmente independentes e produzem uma solução básica viável* $x = (\frac{1}{2}, 0, \frac{1}{4}, 0)'$*; também será uma solução básica viável* $x = (0, 1, \frac{1}{2}, 0)'$ *associada a* $\{A^2, A^3\}$*.* A^1 *e* A^4 *não formam uma base pois são linearmente dependentes.*

A hipótese posto$(A) = \dim($**linha**$(A)) = m$

Havíamos mencionado que a hipótese de que A possui linhas linearmente independentes não restringia em nada a discussão desta seção. Como exemplo, imagine que uma das linhas de A seja combinação linear das outras, digamos $A_m = \sum_{i<m} \alpha_i A_i$. Então existem duas possibilidades:

1. $b_m = \sum_{i<m} \alpha_i b_i$: neste caso qualquer solução x que satisfaça $A_i x = b_i$ para $i = 1, \ldots, m-1$ irá satisfazer

$$A_m x = \left(\sum_{i<m} \alpha_i A_i\right) x = \sum_{i<m} \alpha_i A_i x = \sum_{i<m} \alpha_i b_i = b_m,$$

ou seja, a restrição $A_m x = b_m$ é desnecessária;

2. $b_m \neq \sum_{i<m} \alpha_i b_i$: pelo mesmo argumento concluiremos que qualquer x que satisfaça $A_i x = b_i$ para $i = 1, \ldots, m-1$ NÃO irá satisfazer $A_m x = b_m$, e portanto o poliedro correspondente é vazio.

Restringindo a discussão para poliedros não-vazios, vemos então que restrições que se escrevem como combinação linear das demais podem ser descartadas. Isso justifica o seguinte teorema:

Teorema 2.5 *Seja* $P = \{x \in \mathbb{R}^n \mid Ax = b,\ x \geq 0\} \neq \emptyset$*, onde* $A \in \mathbb{R}^{m\times n}$*. Suponha que posto*$(A) = k < m$ *e que as linhas* $A_1, \ldots, A_k$ *sejam linearmente independentes. Então* $P = \{x \in \mathbb{R}^n \mid A_i x = b_i,\ i = 1, \ldots, k,\ x \geq 0\}$*.*

Prova.

Como posto$(A) = k$ e $A_1, \ldots, A_k$ são linearmente independentes, estas linhas formam uma base do espaço-linha associado a A. Assim as linhas $A_j,\ j = k+1, \ldots, m$ são geradas (como combinação linear) pelas linhas $A_1, \ldots, A_k$. Utilizando o argumento acima, sempre caimos no caso 1 (pois o poliedro é não-vazio), e com isso provamos que $\{x \in \mathbb{R}^n \mid A_i x = b_i,\ i = 1, \ldots, k,\ x \geq 0\} \subset P$. A inclusão contrária é imediata, pois qualquer x que satisfaz $Ax = b$ satisfaz $A_i x = b_i,\ i = 1, \ldots, k, k+1, \ldots, m$.

Deste teorema concluimos que qualquer poliedro não-vazio na forma canônica pode ser re-escrito como $P = \{x \in \mathbb{R}^n \mid Dx = f,\ x \geq 0\}$, onde $D \in \mathbb{R}^{k \times n}$ e $f \in \mathbb{R}^k$ podem ser escritos como

$$D = \left[\begin{array}{c} A_{i_1} \\ \hline A_{i_2} \\ \hline \vdots \\ \hline A_{i_k} \end{array}\right] \quad \text{e} \quad f = \begin{bmatrix} b_{i_1} \\ b_{i_2} \\ \vdots \\ b_{i_k} \end{bmatrix}$$

e as linhas de D são linearmente independentes.

Exemplo 2.3 *Considere o poliedro definido por*

$$\begin{array}{rcrcrcl} 2x_1 & + & x_2 & + & x_3 & = & 2 \\ x_1 & + & x_2 & & & = & 1 \\ x_1 & & & + & x_3 & = & 1 \\ & & \multicolumn{3}{c}{x_1, x_2, x_3 \geq 0.} & & \end{array}$$

A primeira restrição corresponde à soma das outras duas, que por sua vez são linearmente independentes. Assim o poliedro acima é equivalente ao definido por

$$\begin{array}{rcrcrcl} x_1 & + & x_2 & & & = & 1 \\ x_1 & & & + & x_3 & = & 1 \\ & & \multicolumn{3}{c}{x_1, x_2, x_3 \geq 0.} & & \end{array}$$

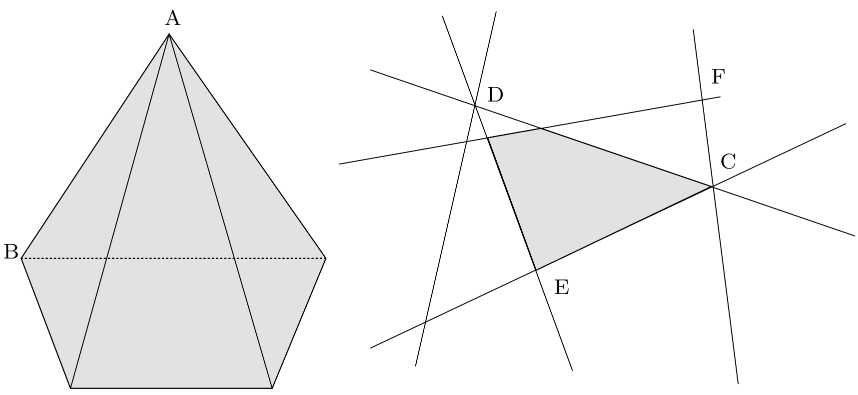
2.4 Degenerescência

O fenômeno conhecido como degenerescência está associado a termos em uma solução básica mais do que o número mínimo necessário de restrições ativas:

Definição 2.10 *Uma solução básica $x \in \mathbb{R}^n$ é* **degenerada** *se mais do que n restrições são ativas em x.*

Assim, se o poliedro está em $\mathbb{R}^2$ uma solução básica degenerada está na intersecção de 3 ou mais retas; em $\mathbb{R}^3$ a intersecção de 4 ou mais planos define uma solução básica degenerada. No caso de poliedros canônicos, pode-se traduzir a definição acima como:

Definição 2.11 *Seja $P = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$ com $A \in \mathbb{R}^{m \times n}$ e x uma solução básica. Dizemos que x é* **degenerada** *se x possui mais de $n - m$ componentes nulas.*



Soluções básicas viáveis não-degeneradas: B,E
Soluções básicas viáveis degeneradas: A,C
Solução básica (inviável) não-degenerada: F
Solução básica (inviável) degenerada: D

Exemplo *Voltando ao exemplo 2.1 da página 24, podemos observar que a solução básica* $x = (2, 6, 0)'$ *da formulação em* $\mathbb{R}^3$ *é não-degenerada, pois possui exatamente 3 restrições ativas:* $x_1 + x_2 + 2x_3 = 8$, $x_2 = 6$ *e* $x_3 = 0$*; a solução correspondente* $x = (2, 6, 0, 0, 6, 2, 0)'$ *em* $\mathbb{R}^7$ *também é não-degenerada, pois possui exatamente* $n - m = 7 - 4 = 3$ *componentes nulas. Por outro lado, a solução* $x = (4, 0, 2)'$ *em* $\mathbb{R}^3$ *é degenerada, pois são ativas 4 restrições:* $x_1 + x_2 + 2x_3 = 8$, $x_2 + 6x_3 = 12$, $x_1 = 4$ *e* $x_2 = 0$*; respectivamente, a solução* $x = (4, 0, 2, 0, 0, 0, 6)'$ *em* $\mathbb{R}^7$ *também é degenerada, pois possui 4 componentes nulas.*

Degenerescência não é uma propriedade exclusivamente geométrica

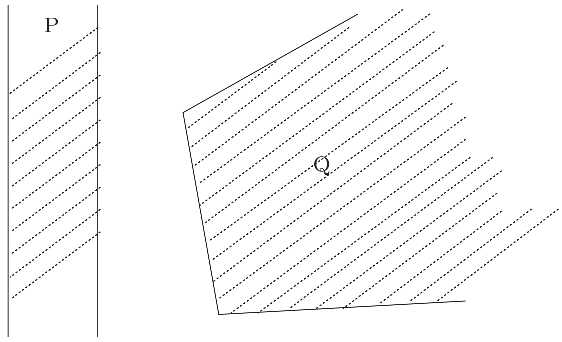
Note que a propriedade de degenerescência está ligada à representação do poliedro. Por exemplo, no poliedro canônico $P = \{x \in \mathbb{R}^3 \mid x_1 - x_2 = 0,\ x_1 + x_2 + 2x_3 = 2,\ x \geq 0\}$, a solução $(1, 1, 0)'$ é não-degenerada (pois possui apenas uma componente nula) enquanto a solução $(0, 0, 1)'$ é degenerada (possui duas componentes nulas); mas o mesmo conjunto $P \subset \mathbb{R}^3$ pode ser descrito como $\{x \in \mathbb{R}^3 \mid x_1 - x_2 = 0,\ x_1 + x_2 + 2x_3 = 2,\ x_1 + x_2 \geq 0,\ x_1 \leq 1,\ x_2 \leq 1\}$ (poliedro não-canônico) e nessa representação a solução $(1, 1, 0)'$ é degenerada e $(0, 0, 1)'$ é não-degenerada. (verifique!)

Outro exemplo: seja x^* uma solução básica viável não-degenerada de um poliedro canônico $P = \{x \in \mathbb{R}^n \mid Ax = b,\ x \geq 0\}$, onde $A \in \mathbb{R}^{m \times n}$; como x^* é básica e não-degenerada, x^* possui exatamente $n - m$ zeros. Mas P pode ser colocado na forma geral $P = \{x \in \mathbb{R}^n \mid Ax \geq b,\ -Ax \geq -b,\ x \geq 0\}$; nesta nova representação x^* continuará possuindo $n - m$ zeros, e além destes mais $2m$ restrições ativas ($Ax \geq b$ e $-Ax \geq -b$), num total de $n + m$ restrições ativas, o que mostra que x^* é degenerada em relação a esta representação.

2.5 Existência de vértices

Note que nem todo poliedro possui vértices. Por exemplo, um semi-espaço em $\mathbb{R}^2$ não possui vértices; se $A \in \mathbb{R}^{m\times n}$ e $m < n$ então o poliedro $P = \{x \in \mathbb{R}^n \mid Ax \geq b\}$ não possui solução básica viável.

Definição 2.12 *Dizemos que um poliedro* $P \subset \mathbb{R}^n$ **contém uma reta** *se existem* $x \in P$ *e* $d \in \mathbb{R}^n \backslash \{0\}$ *tais que* $x + \lambda d \in P$ *para qualquer* $\lambda \in \mathbb{R}$.



P contém uma reta e não possui vértices
Q possui vértices e não contém uma reta

O teorema a seguir mostra que é a propriedade acima caracteriza a (não) existência de vértices.

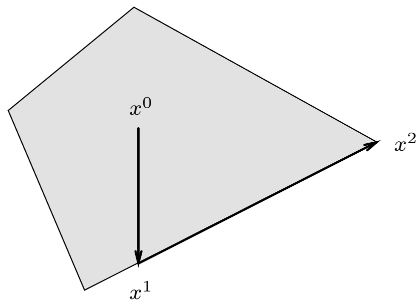
Teorema 2.6 *Seja* $P = \{x \in \mathbb{R}^n \mid Ax \geq b\} \neq \emptyset$. *São equivalentes:*

1. *P possui (pelo menos) um vértice;*
2. *A possui n linhas linearmente independentes;*
3. *P não contém uma reta.*

Prova.

- ($1 \Longrightarrow 2$) Pelo teorema 2.3, se $x \in P$ é vértice, então x é solução básica viável e, em particular, possui n restrições ativas linearmente independentes.
- ($2 \Longrightarrow 3$) Suponha por contradição que P contém uma reta, ou seja, que existem $x \in P$ e $d \in \mathbb{R}^n \backslash \{0\}$ tais que $A(x + \lambda d) \geq b$ para qualquer $\lambda \in \mathbb{R}$. Afirmamos que $Ad = 0$: se $A_i d$ não fosse 0 existiriam valores de λ que violam $A_i(x + \lambda d) \geq b_i$ (por exemplo, $\lambda = \frac{b_i - A_i x - 1}{A_i d}$). Então d é ortogonal a todas as linhas da matriz, e como existem n dentre elas linearmente independentes temos $d = 0$, uma contradição.
- ($3 \Longrightarrow 1$) Seja $x^0 \in P$ e $I = \{i \mid A_i x^0 = b_i\}$. Se $\{A_i \mid i \in I\}$ contém n linhas linearmente independentes, x^0 é um vértice e temos o resultado. Do contrário,

existe um vetor $d \in \mathbb{R}^n \backslash \{0\}$ ortogonal ao espaço gerado pelos A_i, $i \in I$. Por hipótese, a reta $x^0 + \lambda d$, $\lambda \in \mathbb{R}$ não está contida em P, e portanto existe um $\lambda^* \neq 0$ tal que $x^0 + \lambda^* d \in P$ e alguma restrição $A_j x \geq b_j$ inativa em x^0 ($A_j x^0 > b_j$) torna-se ativa em $x^0 + \lambda^* d$ (ou seja, satisfaz $A_j(x^0 + \lambda^* d) = b_j$). Para verificar esta afirmação, escolha j tal que $\frac{|b_j - A_j x^0|}{|A_j d|} = \min \left\{ \frac{|b_k - A_k x^0|}{|A_k d|} \mid A_k d \neq 0 \right\}$ e faça $\lambda^* = \frac{b_j - A_j x^0}{A_j d}$. Note que A_j não pode ser gerado por $\{A_i \mid i \in I\}$, visto que $A_j d = \frac{1}{\lambda^*}(b_j - A_j x^0) \neq 0$. Além disso, todas as restrições ativas em x^0 permanecem ativas em $x^0 + \lambda^* d$, pois para $i \in I$ temos $A_i(x^0 + \lambda^* d) = A_i x^0 + \lambda^* A_i d = A_i x^0 + \lambda^* \cdot 0 = b_i$.



Com isso temos um ponto $x^1 = x^0 + \lambda^* d \in P$ que possui pelo menos uma restrição ativa a mais do que x^0, linearmente independente em relação às anteriores. Se x^1 possui n restrições ativas l.i., x^1 é vértice e temos o resultado. Do contrário, podemos repetir o argumento acima e construir $x^2 \in P$ com pelo menos uma restrição ativa a mais do que x^1, l.i. em relação às anteriores. No máximo depois de n iterações, o método descrito acima produzirá um vértice de P, o que mostra o resultado.

Um corolário direto do teorema acima é:

Corolário 2.2 *Todo poliedro limitado não-vazio possui (pelo menos) um vértice. Todo poliedro canônico não-vazio possui (pelo menos um vértice.*

Prova.

Para a primeira afirmação: um poliedro limitado não contém uma reta. Para a segunda: todo poliedro canônico está contido em $\{x \in \mathbb{R}^n \mid x \geq 0\}$, e este último conjunto não contém uma reta.

2.6 Otimalidade de vértices

Teorema 2.7 *Seja o problema de programação linear*

$$(PL) \begin{cases} \min & c'x \\ s.a & x \in P, \end{cases}$$

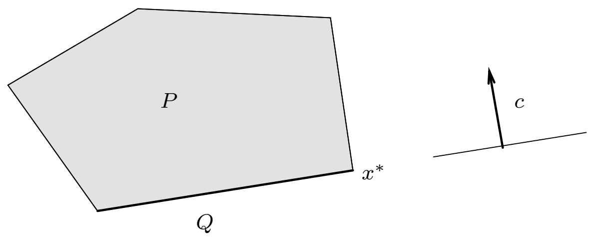
e suponha que este problema admite uma solução ótima e que P possui pelo menos um vértice. Então existe um vértice de P que é solução ótima de (PL).

Prova.

Seja Q o conjunto de soluções ótimas de (PL), ou seja,

$$Q = \{x \in P \mid c'x = v\},$$

onde $v = \min\{c'x \mid x \in P\} \in \mathbb{R}$ e $Q \neq \emptyset$ (pela hipótese de existência da solução ótima). Então Q também é um poliedro, não vazio, e Q não pode conter uma reta, visto que $Q \subset P$. Pelo teorema 2.6 Q possui um vértice x^*; vamos mostrar que x^* também é vértice de P, o que juntamente com $c'x^* = v$ concluirá a demonstração.



Suponha por contradição que x^* não é vertice de P. Então existem $y, z \in P$, $y \neq z$ tais que $x^* = \frac{1}{2}y + \frac{1}{2}z$. Como y e z são viáveis, $c'y \geq v$ e $c'z \geq v$; além disso $c'x^* = \frac{1}{2}c'y + \frac{1}{2}c'z = v$, o que só é possível com $c'y = c'z = v$. Mas isso mostra que $y, z \in Q$, e portanto x^* não seria vértice de Q, uma contradição. Logo x^* é vértice de P e é solução ótima, o que conclui a prova.

Uma das hipóteses do teorema 2.7 é que exista uma solução ótima, para então concluir que existe um vértice ótimo. O próximo resultado completa o anterior, ainda no caso em que P possui vértices (não contém retas): ele afirma que se não houver um vértice ótimo então o problema é ilimitado (não possui solução ótima).

Teorema 2.8 *Seja o problema de programação linear*

$$(PL) \begin{cases} \min & c'x \\ s.a & x \in P, \end{cases}$$

e suponha que P possui pelo menos um vértice. Então ou o valor ótimo de (PL) é $-\infty$, ou existe um vértice de P que é solução ótima de (PL).

Prova.

O núcleo da demonstração é a seguinte propriedade: a menos que o problema seja ilimitado, a partir de qualquer ponto viável podemos obter um vértice com o valor da função objetivo melhor do que o ponto inicial; como o número de vértices é finito, o melhor deles será a solução ótima do problema.

Considere $P = \{x \in \mathbb{R}^n \mid Ax \geq b\}$ (poliedro na forma geral) e seja $x \in P$ qualquer. Se x é vértice, tomamos $v = x$. Do contrário, podemos escolher $d \in \mathbb{R}^n \backslash \{0\}$ tal que $A_i d = 0$, para $i \in I = \{i \mid A_i x = b_i\}$, e podemos supor sem perda de generalidade que $c'd \leq 0$ (do contrário, usaríamos $-d$).

Agora existem duas possibilidades: $c'd < 0$ ou $c'd = 0$.

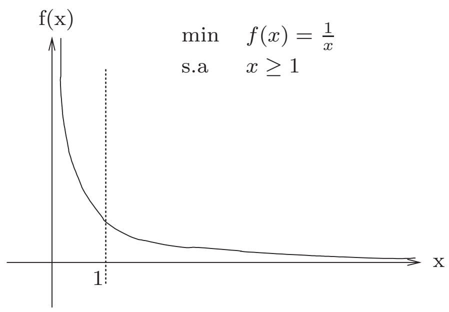
No primeiro caso, ou a semi-reta $\{x + \lambda d \mid \lambda \geq 0\}$ está contida em P, assumindo valores $c'(x + \lambda d) = \overbrace{c'x}^{\text{cte}} + \lambda \overbrace{c'd}^{<0}$ e portanto o valor ótimo do problema é $-\infty$, ou existe um valor $\lambda^* > 0$ tal que pelo menos mais uma restrição A^j é ativa em $x^1 = x + \lambda^* d \in P$ e não é combinação linear de $\{A^i\}_{i \in I}$ (como na demonstração do teorema 2.6) e além disso $c'x^1 = c'x + \lambda^* c'd < c'x$. Basta escolher $\lambda^* = \frac{b_j - A_j x^0}{A_j d} = \min \left\{ \frac{b_k - A_k x^0}{A_k d} \mid A_k d < 0 \right\}$.

No segundo caso ($c'd = 0$), como a reta $\{x + \lambda d \mid \lambda \in \mathbb{R}\}$ não está contida em P, existe um $\lambda^* \in \mathbb{R}$ tal que pelo menos mais uma restrição A^j é ativa em $x^1 = x + \lambda^* d \in P$ e não é combinação linear de $\{A^i\}_{i \in I}$, e além disso $c'x^1 = c'x + \lambda c'd = c'x$.

Repetindo o argumento no máximo n vezes, ou teremos uma direção de ilimitação da função objetivo, ou teremos um vértice v^x tal que $c'v^x \leq c'x$.

Como o número de vértices é finito (corolário 2.1), existe um vértice v^* tal que $c'v^* \leq c'v$ para qualquer outro vértice v. Então se o problema não é ilimitado, $c'v^* \leq c'v^x \leq c'x$, $\forall x \in P$, o que mostra que v^* é solução ótima do problema.

Um exemplo de programação não-linear mostra que a existência de valor ótimo real não implica na existência de solução ótima em geral:



Esta é uma propriedade forte que vale para o caso de programação linear. Com o auxílio dos teoremas anteriores, podemos generalizar o resultado e caracterizar todas as possibilidades de um problema de programação linear:

Corolário 2.3 *Seja o problema de programação linear*

$$(PL)\begin{cases} \min & c'x \\ s.a & x \in P, \end{cases}$$

Então exatamente uma das seguintes afirmações é verdadeira sobre o problema (PL):

1. $P = \emptyset$ *e o valor ótimo é* $+\infty$*;*
2. *Existe uma direção de ilimitação da função objetivo, e o valor ótimo é* $-\infty$*;*
3. *Existe um vértice ótimo de uma reformulação canônica (PLC) equivalente ao (PL).*

Exercício 2.7 *Demonstre o teorema acima, formalizando o seguinte argumento e citando os teoremas utilizados: qualquer problema de programação linear (PL) pode ser transformado em um problema na forma canônica (PLC) que é equivalente no sentido do teorema 1.1; em particular, existe uma função injetora do conjunto viável de (PLC) para o conjunto viável de (PL) que preserva os valores das funções objetivo de (PL) e (PLC). Mostre que as conclusões do teorema 2.8 para o (PLC) podem ser transportadas para o (PL).*

Exercícios sugeridos para o capítulo 2: 2.1, 2.3-2.10, 2.12-2.17, 2.19 e 2.22.