Capítulo 1

Transformações entre problemas lineares

Versão 1: 25 de setembro de 2003 19:17

1.1 Poliedros e sistemas lineares

Nossa maior preocupação nesta seção é providenciar uma interpretação gráfica de poliedros e sistemas lineares.

Nossa primeira discussão foi sobre as diferenças em interpretar a matriz $\mathbf{A}$ do poliedro canônico $X := \{\boldsymbol{x} : \mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geqq \mathbb{0}\}$. É possível olhar como um vetor (de dimensão n) de produtos escalares com as linhas de $\mathbf{A}$, e "inclinação" definida por $\boldsymbol{b}$. Mas também é possível examinar $\boldsymbol{b}$ como combinação a coeficientes das colunas de $\mathbf{A}$ (e portanto de dimensão m).

Assim, se $X := \{\boldsymbol{x} \in \mathbb{R}^n : \mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \leqq \mathbb{0}\}$, $\mathbf{A} \in \mathbb{R}^{m \times n}$, podemos:

- *Por linhas*: espaço $\mathbb{R}^n$

  Olhar X como o conjunto dos pontos do $\mathbb{R}^n$ nos m hiperplanos $\boldsymbol{a}_i\boldsymbol{x} = b_i$ $(i \in \{1, 2, \ldots, n\})$ interceptado pelo ortante positivo.

$$\mathbf{A}\boldsymbol{x} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \vdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_1 \\ \hline \boldsymbol{a}_2 \\ \hline \cdots \\ \hline \boldsymbol{a}_m \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

- *Por coluna*: espaço $\mathbb{R}^m$

  Olhar X como os multiplicadores das semi-retas (cones) $\lambda\boldsymbol{a}^i$ $(\lambda \in \mathbb{R}_+)$ que geram $\boldsymbol{v}$.

$$\mathbf{A}\boldsymbol{x} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \vdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\boldsymbol{a}^1 + x_2\boldsymbol{a}^2 + \cdots + x_n\boldsymbol{a}^n = \boldsymbol{b}$$

Para ilustrar a diferença entre as representações, vamos examinar um exemplo no $\mathbb{R}^2$. Seja X definido por

$$\begin{aligned} ax_1 + bx_2 &= 1 \\ cx_1 + dx_2 &= 0. \end{aligned}$$

Assim, a matriz $\mathbf{A}$ ficaria na forma

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

e o sistema de igualdade assume a forma matricial

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Se olharmos para essa matriz da maneira "usual", por linhas, verificamos que ela representa um sistema de equações. Como estamos em $\mathbb{R}^2$, esse sistema de equações representa a intersecção de dois hiperplanos, caracterizando um poliedro. Mas se olharmos a matriz como colunas, verificamos o seguinte:

$$x_1 \begin{pmatrix} a \\ c \end{pmatrix} + x_2 \begin{pmatrix} b \\ d \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Sendo então uma combinação linear, e mais que isso: como $x_i \geqq 0,\ \forall i \in \{1,2\}$, podemos olhar cada uma das parcelas como geradores de cones.

Generalizando, nessa interpretação de matriz por coluna, para qualquer poliedro canônico na forma $X := \{\boldsymbol{x} : \mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geqq \mathbb{0}\}$, deve-se notar que $x_1\boldsymbol{a}^1 + \cdots + x_n\boldsymbol{a}^n$ gera um cone de dimensão igual a quantidade de colunos l.i. em $\mathbf{A}$. Logo, para que um vetor $\boldsymbol{x}$ pertença ao poliedro canônico X, é necessário e suficiente que $\boldsymbol{b}$ esteja dentro do cone gerado pelos vetores $\{\boldsymbol{a}^1, \boldsymbol{a}^2, \ldots, \boldsymbol{a}^n\}$.

Assim, sendo $I(\boldsymbol{x}) := \{i \in \{1,2,\ldots,n\} : x_i \neq 0\}$, fixando a matrix $\mathbf{A}_{m\times n}$ e um vetor $\boldsymbol{x}$ em X, teremos

$$\boldsymbol{b} = \mathbf{A}\boldsymbol{x} = \sum_{i\in\{1,2,\ldots,n\}} x_i\boldsymbol{a}^i = \sum_{i\in I(x)} x_i\boldsymbol{a}^i + \sum_{i\notin I(x)} x_i\boldsymbol{a}^i.$$

Como $x_i = 0$, para todo $i \in \{1,2,\ldots,n\} \setminus I(x)$ $(\equiv i \notin I(x))$, então $\sum_{i\notin I(x)} x_i\boldsymbol{a}^i = 0$ e portanto

$$\boldsymbol{b} = \mathbf{A}\boldsymbol{x} = \sum_{i\in I(x)} x_i\boldsymbol{a}^i.$$

Usando uma idéia parecida com a decomposição acima, podemos separar as colunas de $\mathbf{A}$ segundo qualquer subconjunto de $\{1,2,\ldots,n\}$. Assim, tomando qualquer $B \subseteq \{1,2,\ldots,n\}$, podemos escrever o sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ na forma separada

$$\boldsymbol{b} = \mathbf{A}_B\boldsymbol{x}_B + \mathbf{A}_N\boldsymbol{x}_N = \mathbf{A}_B\boldsymbol{x}_B.$$

Mais que isso, se tivermos um subconjunto $I \subseteq \{1,2,\ldots,n\}$ tal que $\#I < m$ e $\{\boldsymbol{a}^i\}_{i\in I}$ é l.i., podemos computar um $\boldsymbol{x}$ que resolve o sistema do seguinte modo: sejam[1]

$B := I(x) \cup J$
$N := \{1,2,\ldots,n\} \setminus (I(x) \cup J)$
$J := \{i_1, i_2, \ldots, i_k\} : \{\boldsymbol{a}^i\}_{i\in B}$ seja l.i. e $k + \#I(x) = m$ (ou seja, $\{\boldsymbol{a}^i\}_{i\in B}$ é base do $\mathbb{R}^m$)
$\overline{\mathbf{A}} := \mathbf{A}_B^{-1}$,

basta definir $\boldsymbol{x}$ como

$$\boldsymbol{x}_B = \overline{\mathbf{A}}\boldsymbol{b} \text{ e } \boldsymbol{x}_N = \mathbb{0}, \text{ isto é, } x_i = \begin{cases} \overline{\boldsymbol{a}}^i\boldsymbol{b}, & i \in B \\ 0, & i \in N. \end{cases}$$

Sugestão de aprendizado: Para quem não ficou satisfeito com as equivalências acima, sugerimos que "coloque a mão na massa". Por exemplo, pegue o exemplo inicial, e substitua a, b, c, d por $1, 1, -1, 1$. Trace a intersecção dos hiperplanos em $\mathbb{R}^2$. Verá que isto resulta em um único ponto, que é um poliedro unitário (o ponto encontrado é um vértice do poliedro - como veremos futuramente). Tente interpretar a matriz como coluna, trace os cones e teste numericamente o último resultado.

[1] Note que, trabalhando com colunas de $\mathbf{A}$ estamos considerando o espaço onde o vetor lado-direito $\boldsymbol{b}$ se encontra, o $\mathbb{R}^m$.

1.1.1 Transformação de poliedros genéricos em poliedros canônicos

Um resultado bastante útil de PL é que qualquer poliedro (intersecção finita de hiperplanos e semi-espaços) pode ser transformada na forma canônica. A técnica de transformação se resume a dois casos, abaixo expostos.

- Transformar restrições de desigualdade em restrições de igualdade introduzindo uma **variável de folga** não negativa:

$$\langle \boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}_i\boldsymbol{x} \leqq b_i \rangle \iff \langle (\boldsymbol{x}; x_i^r) \in \mathbb{R}^n \times \mathbb{R}_+ : \boldsymbol{a}_i\boldsymbol{x} + x_i^r = b_i \rangle$$
$$\langle \boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{a}_i\boldsymbol{x} \geqq b_i \rangle \iff \langle (\boldsymbol{x}; x_i^r) \in \mathbb{R}^n \times \mathbb{R}_+ : \boldsymbol{a}_i\boldsymbol{x} - x_i^r = b_i \rangle$$

- Transformar variáveis livres de sinal[2] em variáveis não negativas, trocando cada uma por um par de **variáveis artificiais** : sejam $a^+ := max(a, 0)$ e $a^- := -min(x, 0)$ (então $a = a^+ - a^-$),

$$\left\langle \begin{array}{l} x_i \in \mathbb{R} \text{ e } x_j \in \Gamma, j \in \{1, 2, \ldots, n\} \backslash \{i\} \\ x_1\boldsymbol{a}^1 + \cdots + x_i\boldsymbol{a}^i + \cdots + x_n\boldsymbol{a}^n x_2 = \boldsymbol{b} \end{array} \right\rangle \iff \left\langle \begin{array}{l} (x_i^+, x_i^-) \in \mathbb{R}_+ \times \mathbb{R}_+ \text{ e } x_j \in \Gamma, j \in \{1, 2, \ldots, n\} \backslash \{i\} \\ x_1\boldsymbol{a}^1 + \cdots + x_i^+\boldsymbol{a}^i - x_i^-\boldsymbol{a}^i + \cdots + x_n\boldsymbol{a}^n x_2 = \boldsymbol{b} \end{array} \right\rangle$$

Para ilustrar a técnica, consideremos um poliedro na forma $\overline{\mathbf{A}}\overline{\boldsymbol{x}} \leqq \boldsymbol{b}$. Queremos transformá-lo em um poliedro canônico ($X := \{\boldsymbol{x} : \mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geqq \mathbb{0}\}$). A primeira coisa a fazer, é criar uma variável de folga para produzir a identidade ($= \boldsymbol{b}$),

$$x_i^r = b_i - \overline{\boldsymbol{a}}_i\,\overline{\boldsymbol{x}}.$$

Note que $x_i^r \geqq 0$, desta forma, acertamos o problema da desigualdade, pois

$$\overline{\boldsymbol{a}}_i\,\overline{\boldsymbol{x}} + x_i^r = b_i, \quad \text{com} \quad x_i \geqq 0,$$

introduzindo uma nova variável (que ao menos está necessariamente no ortante positivo como desejamos).

A seguir, devemos eliminar as variáveis livres de sinal: para cada $\overline{x}_i \in \mathbb{R}$ devemos associar dois diferentes pares de variáveis artificiais, x_i^+ e x_i^- ($x_i^+ \geqq 0$ e $x_i^- \geqq 0$). Desta forma, aplicando a técnica de criação de variáveis artificiais: definindo

$$x_i^+ := max(\overline{x}_i, 0) \quad \text{e} \quad x_i^- := -min(\overline{x}_i, 0),$$

teremos $\overline{x}_i = x_i^+ - x_i^-$.

Para um melhor entendimento das duas técnicas expostas, é conveniente aplicá-las a um exemplo. Considere o poliedro $Y = \{\boldsymbol{y} \in \mathbb{R}^2 : 0 \leqq y_1 + y_2 \leqq 1\}$. Devemos inicialmente "eliminar" as desigualdade: $y_1 + y_2 \geqq 0$ e $y_1 + y_2 \leqq 1$. Aplicando a primeira técnica de transformação, obtemos as variáveis artificiais,

$$x_1 := b_1 - \boldsymbol{a}_1\boldsymbol{y} = 1 - y_1 - y_2 \implies \langle \boldsymbol{a}_1\boldsymbol{y} + x_1 = b_1,\ x_1 \geqq 0 \iff \boldsymbol{a}_1\boldsymbol{y} \leqq b_1 \rangle, \quad \text{e}$$
$$x_2 := -b_2 + \boldsymbol{a}_2\boldsymbol{y} = y_1 + y_2, \implies \langle \boldsymbol{a}_2\boldsymbol{y} - x_2 = b_2,\ x_2 \geqq 0 \iff \boldsymbol{a}_2\boldsymbol{y} \geqq b_2 \rangle.$$

Como y_1 e y_2 são livres de sinal, é necessário definirmos variáveis artificiais correspondentes para cada uma delas (no ortante positivo, $\boldsymbol{z} \geqq \mathbb{0}$),

$$\langle x_3 := y_1^+ \text{ e } x_4 := y_1^- \rangle \implies y_1 = x_3 - x_4$$
$$\langle x_5 := y_2^+ \text{ e } x_6 := y_2^- \rangle \implies y_2 = x_5 - x_6$$

Deste modo, definindo

$$\overline{\mathbf{A}} := \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}, \ \mathbf{A} := \begin{bmatrix} 1 & 0 & 1 & -1 & 1 & -1 \\ 0 & 1 & -1 & 1 & -1 & 1 \end{bmatrix} \ \text{e} \ \boldsymbol{b} := \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

[2] Com sinal invertido é simples, basta trocar $x_i \leqq 0$ por $\overline{x}_i \geqq 0$, fazendo $\overline{x}_i = -x_i$.

$$
\left.\begin{aligned}
Y &= \left\{ \boldsymbol{y} \in \mathbb{R}^2 : \begin{bmatrix} y_1 + y_2 \leqq 1 \\ -y_1 - y_2 \leqq 0 \end{bmatrix} \right\} \\
&= \left\{ \boldsymbol{y} \in \mathbb{R}^2 : \overline{\mathbf{A}}\boldsymbol{y} \leqq \boldsymbol{b} \right\}
\end{aligned}\right\}
\longleftrightarrow
\left\{\begin{aligned}
X &= \left\{ \boldsymbol{x} \in \mathbb{R}^6_+ : \begin{bmatrix} x_1 + x_3 - x_4 + x_5 - x_6 = 1 \\ x_2 - x_3 + x_4 - x_5 + x_6 = 0 \end{bmatrix} \right\} \\
&= \left\{ \boldsymbol{x} \in \mathbb{R}^6_+ : \mathbf{A}\boldsymbol{x} = \boldsymbol{b} \right\}
\end{aligned}\right.
$$

Com estas definições, gere pontos em cada um dos poliedros e, aplicando as regras, obtenha o ponto correspondente no outro.