MCTA017 – Programação Matemática

EPR202 – Métodos de Otimização Aplicados à Eng. de Produção

Aula 6

Geometria e Estrutura de Conjuntos Poliédricos (Parte II)

Prof. Dr. Aritanan Gruber

Centro de Matemática, Computação e Cognição
Universidade Federal do ABC

Quadrimestre Suplementar 2020

Agenda
(em três partes)

Vamos desenvolver, em paralelo, as teorias algébrica de sistemas de inequações lineares e geométrica de poliedros: uma informa, motiva e ilustra a outra, tornando a inter-relação rica e poderosa.

Especificamente, vamos:

- estudar representações interna e externa de poliedros e cones;
- estudar a estrutura de poliedros e ligá-la a conceitos vetoriais lineares e afins;
- introduzir um algoritmo iterativo de busca que soluciona problemas viáveis;
- apresentar provas construtivas de resultados clássicos fundamentais como Farkas e Carathéodory;
- provar Dualidade Forte e de Folgas Complementares;
- refinar o algoritmo acima no método Simplex.

Outline

Dualidade Forte e Folgas Complementares

Estruturas de Cones e Poliedros (I)

Dualidade Fraca

Sejam $\mathcal{P} = \{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\, \boldsymbol{x} \geq \boldsymbol{0}\}$ e $\mathcal{Q} = \{\boldsymbol{y} : \boldsymbol{y}^\top \boldsymbol{A} \leq \boldsymbol{c}^\top\}$ poliedros e considere os PLs primal e dual:

$$(P) : \ \min \{\boldsymbol{c}^\top \boldsymbol{x} : \boldsymbol{x} \in \mathcal{P}\} \qquad \text{e} \qquad (D) : \ \max \{\boldsymbol{y}^\top \boldsymbol{b} : \boldsymbol{y} \in \mathcal{Q}\}.$$

Provamos o Teorema Fraco de Dualidade, que estabelece que

$$\inf \{\boldsymbol{c}^\top \boldsymbol{x} : \boldsymbol{x} \in \mathcal{P}\} \geq \sup \{\boldsymbol{y}^\top \boldsymbol{b} : \boldsymbol{y} \in \mathcal{Q}\}.$$

Se $\mathcal{P} \neq \emptyset \neq \mathcal{Q}$, o $\inf$ acima é substituído por $\min$ e o $\sup$, por $\max$:

$$\min \{\boldsymbol{c}^\top \boldsymbol{x} : \boldsymbol{x} \in \mathcal{P}\} \geq \max \{\boldsymbol{y}^\top \boldsymbol{b} : \boldsymbol{y} \in \mathcal{Q}\}.$$

Se ao menos um dos poliedros for não vazio, obtemos as relações entre ilimitação e inviabilidade ($\inf \emptyset = \infty$ e $\sup \emptyset = -\infty$):

(P) é ilimitado $\implies$ (D) é inviável;

(D) é ilimitado $\implies$ (P) é inviável.

Dualidade Fraca

A Dualidade Fraca ainda fornece uma condição suficiente para otimalidade: Se $\boldsymbol{x}_* \in \mathcal{P}$ é primal-viável, $\boldsymbol{y}_* \in \mathcal{Q}$ é dual-viável e $\boldsymbol{c}^\top \boldsymbol{x}_* = \boldsymbol{y}_*^\top \boldsymbol{b}$, então $\boldsymbol{x}_*$ é primal-ótima e $\boldsymbol{y}_*$ é dual-ótima.

Nosso próximo objetivo é mostrar que tal condição também é suficiente, isto é, se $\boldsymbol{x}_* \in \mathcal{P}$ é primal-ótima e $\boldsymbol{y}_* \in \mathcal{Q}$ é dual-ótima, então $\boldsymbol{c}^\top \boldsymbol{x}_* = \boldsymbol{y}_*^\top \boldsymbol{b}$.

Se a suficiência (Dualidade Fraca) é trivial, a necessidade (Dualidade Forte) requer um pouco mais de trabalho: vamos apresentar uma prova baseada no Lema de Farkas I — que provamos via Algoritmo Fundamental ou Fourier-Motzkin. Logo, nossa prova de Dualidade Forte será, de certa forma, construtiva.

Posteriormente, ilustraremos como uma prova de Dualidade Forte pode ser obtida à partir do Simplex (que não é muito diferente do Algoritmo Fundamental, como já mencionamos).

Aplicação: Teorema de Sperner

Antes, no entanto, vamos apresentar uma aplicação interessante de PL que só requer Dualidade Fraca.

Para $n \geq 1$ inteiro, $[n] := \{1, 2, \ldots, n\}$ e $2^{[n]}$ denota o *conjunto potência* de $[n]$; isto é, $2^{[n]} := \{S \subseteq [n]\}$, o conjunto de todos os subconjuntos de $[n]$.

Exemplo: se $n = 3$,

$$2^{[3]} = \big\{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\big\}.$$

$\mathcal{A} \subseteq 2^{[n]}$ é uma *anti-cadeia* se para todos $X, Y \in \mathcal{A}$ distintos, nem $X \subseteq Y$ nem $Y \subseteq X$.

Exemplo: $\big\{\{1\}, \{2\}, \{3\}\big\}$, $\big\{\{1,2\}, \{3\}\big\}$ e $\big\{\{1,2,3\}\big\}$ são anti-cadeias de $2^{[3]}$, mas $\big\{\{2\}, \{3\}, \{1,3\}\big\}$ não.

Aplicação: Teorema de Sperner

Uma anti-cadeia $\mathcal{A}$ de $2^{[n]}$ é:

- *maximal* se $\mathcal{A} = \mathcal{B}$ para toda anti-cadeia $\mathcal{B}$ de $2^{[n]}$ tq $\mathcal{A} \subseteq \mathcal{B}$;
- *máxima* se $|\mathcal{A}| \geq |\mathcal{B}|$, para toda anti-cadeia $\mathcal{B}$ de $2^{[n]}$.

Claramente, toda anti-cadeia máxima é também maximal, mas a recíproca, em geral, não é verdadeira.

Exemplo: $\big\{\{1\}, \{2\}, \{3\}\big\}$ é máxima para $2^{[3]}$; $\big\{\{1,2\}, \{3\}\big\}$ é maximal e não máxima; $\big\{\{1\}, \{2\}\big\}$ não é maximal nem máxima.

$\mathcal{C} \subseteq 2^{[n]}$ é uma *cadeia* se para todos $X, Y \in \mathcal{C}$, $X \subseteq Y$ ou $Y \subseteq X$.

Exemplo: $\big\{\emptyset, \{1\}, \{1,2\}, \{1,2,3\}\big\}$ e $\big\{\{2\}, \{1,2,3\}\big\}$ são cadeias de $2^{[3]}$, mas $\big\{\{1\}, \{2\}\big\}$ e $\big\{\emptyset, \{1,2\}, \{2,3\}, \{1,2,3\}\big\}$ não.

Aplicação: Teorema de Sperner

Uma cadeia $\mathcal{C}$ de $2^{[n]}$ é:

- *maximal* se $\mathcal{C} = \mathcal{D}$ para toda cadeia $\mathcal{D}$ de $2^{[n]}$ tq $\mathcal{C} \subseteq \mathcal{D}$;
- *máxima* se $|\mathcal{C}| \geq |\mathcal{D}|$, para toda cadeia $\mathcal{D}$ de $2^{[n]}$.

Claramente, toda cadeia máxima é também maximal, mas a recíproca, em geral, não é verdadeira. (Em $2^{[n]}$ a recíproca vale.)

Fato

$|\mathcal{A} \cap \mathcal{C}| \leq 1$ *para toda anti-cadeia* $\mathcal{A}$ *e toda cadeia* $\mathcal{C}$.

Não é difícil perceber que toda cadeia máxima de $2^{[n]}$ tem $n + 1$ elementos e que existem $n!$ delas.

Com relação a anti-cadeias, é fácil acharmos uma de $\binom{n}{\lfloor n/2 \rfloor}$ elementos, mas não é claro se não existe uma maior.

O Teorema de Sperner afirma que não. Consequentemente, se n é ímpar, existe uma única anti-cadeia máxima; se n é par, existem duas.

Aplicação: Teorema de Sperner

Teorema (Sperner)

Uma anti-cadeia máxima de $2^{[n]}$ tem $\binom{n}{\lfloor n/2 \rfloor}$ elementos.

Prova. Seja $\mathcal{A}$ uma anti-cadeia máxima de $2^{[n]}$ que sabemos existir. Defina variáveis (primais) $x_J = 1$ se $J \in \mathcal{A}$, e $x_J = 0$ em caso contrário, para cada $J \subseteq [n]$.

Considere o Programa Inteiro (PI):

$$
\begin{array}{lll}
\max & \displaystyle\sum_{J \subseteq [n]} x_J & \\
\text{s.t.} & \displaystyle\sum_{J \in \mathcal{C}} x_J \leq 1 & (\forall \text{ cadeia maximal } \mathcal{C} \subseteq 2^{[n]}) \\
& x_J \in \{0, 1\} & (\forall\, J \subseteq [n]) \qquad (\diamond)
\end{array}
$$

Com a restrição $(\diamond)$, uma solução ótima para o PI equivale a uma anti-cadeia máxima.

Aplicação: Teorema de Sperner

Considere o relaxamento linear (PL):

$$\begin{aligned}
\max \quad & \sum_{J \subseteq [n]} x_J \\
\text{s.t.} \quad & \sum_{J \in \mathcal{C}} x_J \leq 1 \qquad (\forall \, \text{cadeia maximal } \mathcal{C} \subseteq 2^{[n]}) \\
& x_J \geq 0 \qquad (\forall \, J \subseteq [n])
\end{aligned}$$

Claramente, $\max$ (PI) $\leq$ $\max$ (PL).

Nota: Observe que tanto o PI quanto o PL têm 2^n variáveis (binárias no PI, contínuas no PL) e $n!$ restrições! $\implies$ os programas são gigantes para valores moderados de n $\implies$ não é possível utilizar solvers! Neste caso, buscamos uma solução analítica.

Para obter o dual do PL, defina variáveis (duais) $y_{\mathcal{C}}$ para cada cadeia maximal $\mathcal{C}$.

Aplicação: Teorema de Sperner

PL Dual (DL):

$$\begin{aligned} \min \quad & \sum \{ y_{\mathcal{C}} : \mathcal{C} \text{ cadeia maximal de } 2^{[n]} \} \\ \text{s.t.} \quad & \sum_{\mathcal{C} \ni J} y_{\mathcal{C}} \geq 1 \qquad (\forall\, J \subseteq [n]) \\ & \quad\; y_{\mathcal{C}} \geq 0 \qquad (\forall \text{ cadeia maximal } \mathcal{C} \subseteq 2^{[n]}) \end{aligned}$$

Procuramos, agora, pela melhor solução dual-viável em que todas as variáveis $y_{\mathcal{C}}$ são iguais, i.é, têm o mesmo valor y. Determinemos y.

Precisamos, para todo $J \subseteq [n]$, que

$$y \cdot \big(\# \text{ de cadeias maximais contendo } J\big) \geq 1.$$

Ou seja, para todo $j \in \{0, 1, \ldots, n\}$, precisamos (por que?) que

$$y \cdot j!(n-j)! \geq 1.$$

Aplicação: Teorema de Sperner

Fato: $j!(n-j)!$ atinge o valor mínimo quando $j = \left\lfloor \frac{n}{2} \right\rfloor$.

[Prove usando a desigualdade de Stirling: $n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$.]

Temos, então, que

$$y = \frac{1}{\lfloor n/2 \rfloor!\big(n - \lfloor n/2 \rfloor\big)!}$$

é dual-viável e o valor associado à função objetivo de (DL) é

$$n!\, y = \frac{n!}{\lfloor n/2 \rfloor!\big(n - \lfloor n/2 \rfloor\big)!} = \binom{n}{\lfloor n/2 \rfloor}.$$

Pelo Teorema Fraco da Dualidade e pelo desenvolvimento acima, temos que

$$\tbinom{n}{\lfloor n/2 \rfloor} \leq \max \text{ (PI)} \leq \max \text{ (PL)} \leq \min \text{ (DL)} \leq \tbinom{n}{\lfloor n/2 \rfloor}.$$

$\square$

Nota: Cadeias e Anti-cadeias

Cadeias e anti-cadeias são conceitos naturais em teorias de Reticulados e Conjuntos Parcialmente Ordenados — estas últimas focam na noção de Ordem, generalizam muitos problemas e domínios e têm muitas aplicações.

Em otimização linear, podemos citar a ocorrência de cadeias e anti-cadeias em teorias de emparelhamentos, em funções submodulares, em problemas de empacotamento, cobertura e orientação, e em soluções obtidas via Programação Dinâmica discreta (e, até certo ponto, estocástica).

Em particular, os programas (PL) e (DL) são exemplos simplórios de problemas de empacotamento e cobertura, respectivamente. Outros exemplos virão com o tempo.

Teorema Forte de Dualidade

Teorema (von Newmann: Dualidade Forte)

Sejam $\boldsymbol{A}$ *uma matriz e* $\boldsymbol{b}, \boldsymbol{c}$ *vetores. Temos que*

$$\min\left\{\boldsymbol{c}^\top \boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\, \boldsymbol{x} \geq \boldsymbol{0}\right\} = \max\left\{\boldsymbol{y}^\top \boldsymbol{b} : \boldsymbol{y}^\top \boldsymbol{A} \leq \boldsymbol{c}^\top\right\}$$

quando ambos os conjuntos são não vazios.

Em outras palavras, se os PLs da esquerda (primal) e da direita (dual) forem ambos viáveis, as soluções ótimas do primal e do dual possuem o mesmo valor.

Prova. Sejam $\mathcal{P} := \left\{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\, \boldsymbol{x} \geq \boldsymbol{0}\right\}$ e $\mathcal{Q} := \left\{\boldsymbol{y} : \boldsymbol{y}^\top \boldsymbol{A} \leq \boldsymbol{c}^\top\right\}$ não vazios. É imediato que $\boldsymbol{c}^\top \boldsymbol{x} \geq \boldsymbol{y}^\top \boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}^\top \boldsymbol{b}$ para todos $\boldsymbol{x} \in \mathcal{P}$ e $\boldsymbol{y} \in \mathcal{Q}$, o que resulta em

$$\min\left\{\boldsymbol{c}^\top \boldsymbol{x} : \boldsymbol{x} \in \mathcal{P}\right\} \geq \max\left\{\boldsymbol{y}^\top \boldsymbol{b} : \boldsymbol{y} \in \mathcal{Q}\right\}. \quad \text{(Dualidade Fraca)}$$

Resta mostrar que existem $\boldsymbol{x}_* \in \mathcal{P}$ e $\boldsymbol{y}_* \in \mathcal{Q}$ tais que $\boldsymbol{c}^\top \boldsymbol{x}_* \leq \boldsymbol{y}_*^\top \boldsymbol{b}$.

Teorema Forte de Dualidade

Considere o sistema $\boldsymbol{A}' \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix} \leq \boldsymbol{b}'$:

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{0} \\ -\boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{A}^\top \\ -\boldsymbol{c}^\top & \boldsymbol{b}^\top \end{pmatrix} \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix} \leq \begin{pmatrix} \boldsymbol{b} \\ -\boldsymbol{b} \\ \boldsymbol{c} \\ 0 \end{pmatrix}. \qquad \begin{array}{l} \leftarrow \boldsymbol{p} \\ \leftarrow \boldsymbol{q} \\ \leftarrow \boldsymbol{r} \\ \leftarrow \mu \end{array}$$

Pelo Lema de Farkas I, a existência de uma solução $\begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix}$ com $\boldsymbol{x} \geq 0$ para o sistema acima equivale à veracidade da afirmação: para todos $\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{r} \geq \boldsymbol{0}$ e $\mu \geq 0$, se

$$\boldsymbol{p}^\top \boldsymbol{A} - \boldsymbol{q}^\top \boldsymbol{A} - \mu \boldsymbol{c}^\top \geq \boldsymbol{0} \quad \text{e} \quad \boldsymbol{r}^\top \boldsymbol{A}^\top + \mu \boldsymbol{b}^\top = \boldsymbol{0}, \quad (1)$$

então

$$\boldsymbol{p}^\top \boldsymbol{b} - \boldsymbol{q}^\top \boldsymbol{b} + \boldsymbol{r}^\top \boldsymbol{c} \geq 0. \quad (2)$$

Teorema Forte de Dualidade

Sejam $\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{r}, \mu$ não negativos satisfazendo (1). Caso $\mu = 0$, sejam $\boldsymbol{x}_0 \in \mathcal{P}$ e $\boldsymbol{y}_0 \in \mathcal{Q}$. Temos que $(\boldsymbol{p}^\top - \boldsymbol{q}^\top)\boldsymbol{A}\boldsymbol{x}_0 \geq 0$ e $\boldsymbol{y}_0^\top \boldsymbol{A}\boldsymbol{r} = 0$. Logo,

$$\boldsymbol{p}^\top \boldsymbol{A}\boldsymbol{x}_0 - \boldsymbol{q}^\top \boldsymbol{A}\boldsymbol{x}_0 + \boldsymbol{r}^\top \big(\boldsymbol{y}_0^\top \boldsymbol{A}\big)^\top \geq 0,$$

o que implica em

$$\boldsymbol{p}^\top \boldsymbol{b} - \boldsymbol{q}^\top \boldsymbol{b} + \boldsymbol{r}^\top \boldsymbol{c} \geq 0.$$

Caso $\mu > 0$, por substituição de (1) em (2), temos que

$$\begin{aligned}\boldsymbol{p}^\top \boldsymbol{b} - \boldsymbol{q}^\top \boldsymbol{b} + \boldsymbol{r}^\top \boldsymbol{c} &= \frac{\mu}{\mu}(\boldsymbol{p}^\top - \boldsymbol{q}^\top)\boldsymbol{b} + \frac{\mu}{\mu}\boldsymbol{c}^\top \boldsymbol{r} \\ &\geq \mu^{-1}(\boldsymbol{q}^\top - \boldsymbol{p}^\top)\boldsymbol{A}\boldsymbol{r} - \mu^{-1}(\boldsymbol{q}^\top - \boldsymbol{p}^\top)\boldsymbol{A}\boldsymbol{r} = 0.\end{aligned}$$

Portanto, segue que $\min\left\{\boldsymbol{c}^\top \boldsymbol{x} : \boldsymbol{x} \in \mathcal{P}\right\} = \max\left\{\boldsymbol{y}^\top \boldsymbol{b} : \boldsymbol{y} \in \mathcal{Q}\right\}$. $\square$

Teorema Forte de Dualidade

Não é difícil perceber que uma prova semelhante pode ser obtida para quando o PL primal estiver na forma canônica; ou em qualquer outra forma. Mais ainda: como todo primal pode ser reduzido à forma padrão (e todo dual, à forma canônica a menos de não negatividade), a prova acima pode ser considerada "definitiva."

É conveniente ter os pares primal-dual em mente:

(1) $\min\left\{\boldsymbol{c}^\top\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\, \boldsymbol{x} \geq \boldsymbol{0}\right\} = \max\left\{\boldsymbol{y}^\top\boldsymbol{b} : \boldsymbol{y}^\top\boldsymbol{A} \leq \boldsymbol{c}^\top\right\}$

(2) $\min\left\{\boldsymbol{c}^\top\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b},\, \boldsymbol{x} \geq \boldsymbol{0}\right\} = \max\left\{\boldsymbol{y}^\top\boldsymbol{b} : \boldsymbol{y}^\top\boldsymbol{A} \geq \boldsymbol{c}^\top,\, \boldsymbol{y} \leq \boldsymbol{0}\right\}$

(3) $\min\left\{\boldsymbol{c}^\top\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b},\, \boldsymbol{x} \geq \boldsymbol{0}\right\} = \max\left\{\boldsymbol{y}^\top\boldsymbol{b} : \boldsymbol{y}^\top\boldsymbol{A} \leq \boldsymbol{c}^\top,\, \boldsymbol{y} \geq \boldsymbol{0}\right\}$

(4) $\min\left\{\boldsymbol{c}^\top\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}\right\} = \max\left\{\boldsymbol{y}^\top\boldsymbol{b} : \boldsymbol{y}^\top\boldsymbol{A} = \boldsymbol{c}^\top\right\}$

(5) $\min\left\{\boldsymbol{c}^\top\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}\right\} = \max\left\{\boldsymbol{y}^\top\boldsymbol{b} : \boldsymbol{y}^\top\boldsymbol{A} = \boldsymbol{c}^\top,\, \boldsymbol{y} \leq \boldsymbol{0}\right\}$

(6) $\min\left\{\boldsymbol{c}^\top\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}\right\} = \max\left\{\boldsymbol{y}^\top\boldsymbol{b} : \boldsymbol{y}^\top\boldsymbol{A} = \boldsymbol{c}^\top,\, \boldsymbol{y} \geq \boldsymbol{0}\right\}$

Teorema Forte de Dualidade

A título de curiosidade, prove a versão para as formulações genéricas.

Para matrizes $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{D}, \boldsymbol{E}, \boldsymbol{F}, \boldsymbol{G}, \boldsymbol{H}, \boldsymbol{K}$ e vetores $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}, \boldsymbol{e}, \boldsymbol{f}$, desde que os dois PLs sejam viáveis, eles têm o mesmo valor ótimo.

$$
\begin{array}{rl|rl}
\min\limits_{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}} & \boldsymbol{a}^\top\boldsymbol{x} + \boldsymbol{b}^\top\boldsymbol{y} + \boldsymbol{c}^\top\boldsymbol{z} & \max\limits_{\boldsymbol{p},\boldsymbol{q},\boldsymbol{r}} & \boldsymbol{p}^\top\boldsymbol{d} + \boldsymbol{q}^\top\boldsymbol{e} + \boldsymbol{r}^\top\boldsymbol{f} \\
\text{s.t.} & \boldsymbol{A}\boldsymbol{x} + \boldsymbol{D}\boldsymbol{y} + \boldsymbol{G}\boldsymbol{z} \leq \boldsymbol{d} & \text{s.t.} & \boldsymbol{p}^\top\boldsymbol{A} + \boldsymbol{q}^\top\boldsymbol{B} + \boldsymbol{r}^\top\boldsymbol{C} \geq \boldsymbol{a} \\
& \boldsymbol{B}\boldsymbol{x} + \boldsymbol{E}\boldsymbol{y} + \boldsymbol{H}\boldsymbol{z} = \boldsymbol{e} & & \boldsymbol{p}^\top\boldsymbol{D} + \boldsymbol{q}^\top\boldsymbol{E} + \boldsymbol{r}^\top\boldsymbol{F} = \boldsymbol{b} \\
& \boldsymbol{C}\boldsymbol{x} + \boldsymbol{F}\boldsymbol{y} + \boldsymbol{K}\boldsymbol{z} \geq \boldsymbol{f} & & \boldsymbol{p}^\top\boldsymbol{G} + \boldsymbol{q}^\top\boldsymbol{H} + \boldsymbol{r}^\top\boldsymbol{K} \leq \boldsymbol{c} \\
& \boldsymbol{x} \geq \boldsymbol{0} & & \boldsymbol{p} \geq \boldsymbol{0} \\
& \boldsymbol{z} \leq \boldsymbol{0} & & \boldsymbol{r} \leq \boldsymbol{0}
\end{array}
$$

Teorema Forte de Dualidade

$$
\begin{array}{cc|ccc|cc}
 & & \boldsymbol{x} \geq \mathbf{0} & \boldsymbol{y} & \boldsymbol{z} \leq \mathbf{0} & & \\
\hline
\mathbf{0} \leq & \boldsymbol{p} & \boldsymbol{A} & \boldsymbol{D} & \boldsymbol{G} & \leq & \boldsymbol{d} \\
 & \boldsymbol{q} & \boldsymbol{B} & \boldsymbol{E} & \boldsymbol{H} & = & \boldsymbol{e} \\
\mathbf{0} \geq & \boldsymbol{r} & \boldsymbol{C} & \boldsymbol{F} & \boldsymbol{K} & \geq & \boldsymbol{f} \\
\hline
 & & \geq \boldsymbol{a} & = \boldsymbol{b} & \leq \boldsymbol{c} & &
\end{array}
$$

$$\min \boldsymbol{a}^\top \boldsymbol{x} + \boldsymbol{b}^\top \boldsymbol{y} + \boldsymbol{c}^\top \boldsymbol{z} = \max \boldsymbol{p}^\top \boldsymbol{d} + \boldsymbol{q}^\top \boldsymbol{e} + \boldsymbol{r}^\top \boldsymbol{f}$$

Teorema Forte de Dualidade

Em espaços de dimensão finita (nosso caso) vale que:

Proposição (exercício)

O dual de um dual é o primal.

O Teorema Forte de Dualidade requer que tanto o primal quanto o dual tenham soluções viáveis. No caso de um deles ser inviável e o outro ilimitado, ainda é possível provarmos um resultado ligeiramente mais fraco:

$$\inf\left\{\boldsymbol{c}^\top\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\, \boldsymbol{x} \geq \boldsymbol{0}\right\} = \sup\left\{\boldsymbol{y}^\top\boldsymbol{b} : \boldsymbol{y}^\top\boldsymbol{A} \leq \boldsymbol{c}^\top\right\}.$$

Claramente, a relação acima é inválida caso ambos os PLs sejam inviáveis.

Nota: Lembre-se que o $\inf$ de um PL inviável é ∞ e de um PL ilimitado é $-\infty$; para $\sup$, inverta os sinais.

Interpretações de Dualidade

Uma interpretação Geométrica de pares primal-dual e folgas complementares pode ser obtida em Bertsimas-Tsitsiklis, pág. 153–155. O conceito é interessante e será coberto de outra forma posteriormente.

Variáveis duais ótimas podem ser interpretadas como Custos Marginais e têm relativa importância econômica (dependendo do problema), apesar da simplicidade conceitual. Vamos postergar este assunto para depois de cobrirmos o algoritmo Simplex e seu uso no `gurobi`.

Enquanto isso, você pode consultar sobre custos marginais em Bertsimas-Tsitsiklis, pág. 155–156.

Folgas Complementares

Considere o par primal-dual viável e não ilimitado

$$\min\left\{\boldsymbol{c}^\top\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b},\, \boldsymbol{x} \ge \boldsymbol{0}\right\} = \max\left\{\boldsymbol{y}^\top\boldsymbol{b} : \boldsymbol{y}^\top\boldsymbol{A} \le \boldsymbol{c}^\top,\, \boldsymbol{y} \ge \boldsymbol{0}\right\} \quad (\divideontimes)$$

e sejam $\boldsymbol{x}$ e $\boldsymbol{y}$ soluções primal- e dual-viáveis.

Corolário (Folgas Complementares)

$\boldsymbol{x}$ e $\boldsymbol{y}$ são soluções primal- e dual-ótimas se e somente se

(P) $y_i\left(b_i - \boldsymbol{a}_i^\top\boldsymbol{x}\right) = \boldsymbol{0}$ *para todo i, e* *[folgas primais]*

(D) $\left(\boldsymbol{y}^\top\boldsymbol{A}_j - c_j\right)x_j = \boldsymbol{0}$ *para todo j.* *[folgas duais]*

Em sendo o caso, $\boldsymbol{x}$ e $\boldsymbol{y}$ são complementares com relação a $(\divideontimes)$.

[Em notação vetorial, $\boldsymbol{y} \odot \left(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}\right) = \boldsymbol{0}$ e $\left(\boldsymbol{y}^\top\boldsymbol{A} - \boldsymbol{c}^\top\right) \odot \boldsymbol{x} = \boldsymbol{0}$.]

Folgas Complementares

Prova. Sejam $\boldsymbol{s} := \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{0}$ e $\boldsymbol{t} := \boldsymbol{y}^\top \boldsymbol{A} - \boldsymbol{c}^\top \geq \boldsymbol{0}$ as folgas primal e dual para $\boldsymbol{x}$ e $\boldsymbol{y}$ viáveis. Temos que

$$\boldsymbol{y}^\top \boldsymbol{s} + \boldsymbol{t}^\top \boldsymbol{x} = \boldsymbol{y}^\top \boldsymbol{b} - \boldsymbol{y}^\top \boldsymbol{A}\boldsymbol{x} + \boldsymbol{y}^\top \boldsymbol{A}\boldsymbol{x} - \boldsymbol{c}^\top \boldsymbol{x} = \boldsymbol{y}^\top \boldsymbol{b} - \boldsymbol{c}^\top \boldsymbol{x}.$$

Se $\boldsymbol{x}$ e $\boldsymbol{y}$ são complementares, $\boldsymbol{y}^\top \boldsymbol{s} = \boldsymbol{t}^\top \boldsymbol{x} = 0$ e assim, $\boldsymbol{y}^\top \boldsymbol{b} = \boldsymbol{c}^\top \boldsymbol{x}$. Pelo Teorema Fraco de Dualidade, $\boldsymbol{x}$ e $\boldsymbol{y}$ são soluções ótimas.

Se $\boldsymbol{x}$ e $\boldsymbol{y}$ são soluções ótimas, o Teorema Forte de Dualidade implica que $\boldsymbol{y}^\top \boldsymbol{b} = \boldsymbol{c}^\top \boldsymbol{x}$ e, consequentemente, que $\boldsymbol{y}^\top \boldsymbol{s} + \boldsymbol{t}^\top \boldsymbol{x} = 0$. Como $\boldsymbol{x} \geq \boldsymbol{0}$ e $\boldsymbol{y} \geq \boldsymbol{0}$, temos que $\boldsymbol{x}$ e $\boldsymbol{y}$ são complementares. $\square$

Nota: No caso do primal estar na forma padrão

$$\min \left\{ \boldsymbol{c}^\top \boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\, \boldsymbol{x} \geq \boldsymbol{0} \right\} = \max \left\{ \boldsymbol{y}^\top \boldsymbol{b} : \boldsymbol{y}^\top \boldsymbol{A} \leq \boldsymbol{c}^\top \right\},$$

as folgas primais são trivialmente satisfeitas; restam só as duais.

Folgas Complementares

Em palavras, é necessário e suficiente para soluções $\boldsymbol{x}$ primal- e $\boldsymbol{y}$ dual-viáveis serem ótimas que:

- a restrição primal correspondente a uma variável dual não nula em $\boldsymbol{y}$ seja satisfeita com igualdade;
- a restrição dual correspondente a uma variável primal não nula em $\boldsymbol{x}$ seja satisfeita com igualdade.

De forma mais geral, para cada desigualdade $\boldsymbol{a}_{i:}^\top \boldsymbol{x} \leq b_i$ do primal, uma e somente uma condição é válida: ou existe $\boldsymbol{x}^*$ primal-ótimo tal que $\boldsymbol{a}_{i:}^\top \boldsymbol{x}^* < b_i$, ou existe $\boldsymbol{y}^*$ dual-ótimo tal que $y_i^* > 0$. Afirmações semelhantes podem ser feitas para as desigualdades do dual.

Caso uma variável primal (dual) seja em $\boldsymbol{x}$ ($\boldsymbol{y}$) seja nula, a correspondente restrição dual (primal) pode ou não ser satisfeita com igualdade. Isto tem ligações com unicidade de soluções ótimas e degenerecência.

Folgas Complementares

Exemplo: Considere o par primal-dual:

$$
\begin{aligned}
\max \quad & 7x_1 + 2x_2 \\
\text{s.t.} \quad & -x_1 + 2x_2 \leq 4 \\
& 5x_1 + x_2 \leq 20 \\
& -2x_1 - 2x_2 \leq -7 \\
& x_1, x_2 \geq 0
\end{aligned}
$$

$$
\begin{aligned}
\min \quad & 4y_1 + 20y_2 - 7y_3 \\
\text{s.t.} \quad & -y_1 + 5y_2 - 2y_3 \geq 7 \\
& 2y_1 + y_2 - 2y_3 \geq 2 \\
& y_1, y_2, y_3 \geq 0
\end{aligned}
$$

Observe que $\boldsymbol{x} = (\frac{36}{11}, \frac{40}{11})$ é primal-viável e que $\boldsymbol{y} = (\frac{3}{11}, \frac{16}{11}, 0)$ é dual-viável. Além disso, $(7,2)^\top \boldsymbol{x} = \frac{332}{11} = \boldsymbol{y}^\top(4, 20, -7)$ e $\boldsymbol{x}$ e $\boldsymbol{y}$ são ótimas.

As variáveis de folga primal são $(s_1, s_2, s_3) = (0, 0, \frac{75}{11})$ e dual são $(t_1, t_2) = (0, 0)$. Logo, $y_i s_i = 0$ para $i \in [3]$ e $x_j t_j = 0$ para $j \in [2]$.

Aplicação: Max-Flow / Min-Cut

Seja $D = (V, E)$ um digrafo, $c : E \to \mathbb{R}_{\geq 0}$ uma função capacidade para os arcos de D e $s, t \in V$ vértices distintos.

Para um conjunto $s \in S \subseteq V \setminus \{t\}$, temos dois conjuntos de arcos ((s,t)-cortes) associados a S:

$$\partial^+(S) := \big\{(u, v) : u \in S,\, v \in V \setminus S\big\} \subseteq E,$$
$$\partial^-(S) := \big\{(v, u) : u \in S,\, v \in V \setminus S\big\} \subseteq E.$$

$\partial^+(S)$ é o (s,t)-*corte de saída* e o $\partial^-(S)$ é o (s,t)-*corte de entrada*.

A remoção de $\partial^+(S)$ elimina todos os (p, q)-caminhos (orientados) em D, com $p \in S$ e $q \in V \setminus S$. Em particular, elimina todos os (s,t)-caminhos em D.

A remoção de $\partial^-(S)$ reverte os papéis de p e q acima, e a de $\partial^+(S) \cup \partial^+(S)$ desconecta D.

Aplicação: Max-Flow / Min-Cut

A *capacidade* de $\partial^{\#}(S)$, para $\# \in \{+,-\}$, é dada por

$$c\big(\partial^{\#}(S)\big) := \sum_{e \in \partial^{\#}(S)} c(e).$$

O problema do *corte mínimo* em digrafos: dados D, c, s e t, determine $s \in S \subseteq V \setminus \{t\}$ tal que $c\big(\partial^{+}(S)\big)$ seja o menor possível.

Em outra notação:

$$\min\left\{c\big(\partial^{+}(S)\big) : s \in S \subseteq V \setminus \{t\}\right\}.$$

Seja $x : A \to \mathbb{R}_{\geq 0}$ um fluxo viável para (D, c, s, t):

- $0 \leq x(e) \leq c(e)$ para todo $e \in E$,
- $x\big(\partial^{+}(\{u\})\big) - x\big(\partial^{-}(\{u\})\big) = 0$ para todo $u \in V \setminus \{s, t\}$.

(veja slides com definições do problema do fluxo máximo na aula 1)

Aplicação: Max-Flow / Min-Cut

Para todo $s \in S \subseteq V \setminus \{t\}$, é imediato que

$$x\big(\partial^+(S)\big) - x\big(\partial^-(S)\big) \le x\big(\partial^+(S)\big) \le c\big(\partial^+(S)\big).$$

Em particular,

$$\begin{aligned}&\max\left\{x\big(\partial^+(\{s\})\big) - x\big(\partial^-(\{s\})\big) : x \text{ é um fluxo viável}\right\}\\ &\le\\ &\min\left\{c\big(\partial^+(S)\big) : s \in S \subseteq V \setminus \{t\}\right\}.\end{aligned}$$

Em palavras, o valor de um fluxo máximo em (D, c, s, t) é menor ou igual ao de um (s,t)-corte mínimo.

E isto nada mais é do que o Teorema Fraco de Dualidade.

Vamos mostrar que vale a igualdade acima (via Dualidade Forte) e caracterizar as soluções ótimas (com Folgas Complementares).

Aplicação: Max-Flow / Min-Cut

Considere o PL primal para fluxo máximo em (D, c, s, t):

$$
\begin{array}{llll}
\max\limits_{x_e} & x\big(\partial^+(\{s\})\big) - x\big(\partial^-(\{s\})\big) & & \\
\text{s.t.} & x\big(\partial^+(\{u\})\big) = x\big(\partial^-(\{u\})\big) & (u \in V \setminus \{s,t\}) & \leftarrow z_u \\
 & 0 \le x_e \le c_e & (e \in E) & \leftarrow y_e
\end{array}
$$

em que já anotamos as variáveis duais $y_e = y_{uv}$ para $e = (u, v) \in E$ e z_u para $u \in V$. Do primal, segue que $y_e \geq 0$ e z_u é livre de sinal. Seguindo o procedimento para obtenção do dual:

$$
\begin{array}{lll}
\min\limits_{y_e,\, z_u} & \sum\limits_{e \in E} c_e y_e & \\
\text{s.t.} & z_v - z_u + y_e \geq 0 & (e = (u,v) \in E) \\
 & z_s = 1,\ z_t = 0 & (\circledast) \\
 & y_e \geq 0 & (e \in E)
\end{array}
$$

No mínimo, $(\circledast)$ é igual a $z_s - z_t = 1$.

Aplicação: Max-Flow / Min-Cut

O Teorema Forte de Dualidade implica que $\max = \min$. Vamos mostrar que o mínimo do dual acima é igual à capacidade mínima de um (s,t)-corte.

Sejam $\boldsymbol{x}^*$ e $\boldsymbol{y}^*, \boldsymbol{z}^*$ soluções ótimas para o primal e dual e defina $S := \{u \in V : z^*_u \geq 1\}$. Claramente, $s \in S$ e $t \notin S$.

Se $e = (u,v) \in \partial^+(S)$, então $z^*_v - z^*_u + y^*_e \geq z^*_v - z^*_u > 0$; por Folgas Complementares, segue que $x^*_e = c_e$.

Se $e = (u,v) \in \partial^-(S)$, então $y^*_e \geq z^*_u - z^*_v > 0$; por Folgas Complementares, $x^*_e = 0$. Logo,

$$\begin{aligned} x^*\big(\partial^+(\{s\})\big) - x^*\big(\partial^-(\{s\})\big) &= \sum_{u \in S} \Big(x^*\big(\partial^+(\{u\})\big) - x^*\big(\partial^-(\{u\})\big) \Big) \\ &= x^*\big(\partial^+(S)\big) - x^*\big(\partial^-(S)\big) \\ &= c\big(\partial^+(S)\big). \end{aligned}$$

Aplicação: Max-Flow / Min-Cut

Isso mostra que o fluxo máximo é igual à capacidade do (s,t)-corte $\partial^+(S)$. Como nenhum (s,t)-corte pode ter capacidade menor que o valor de um fluxo (Dualidade Fraca), a igualdade está provada.

Teorema (Ford-Fulkerson; Elias-Feinstein-Shannon)

O valor de um (s,t)-fluxo máximo sujeito à capacidade c — em (D,c,s,t) — é igual à capacidade de (s,t)-um corte mínimo.

A prova acima ainda mostra que o dual admite uma solução ótima inteira: sejam $\overline{\boldsymbol{y}}$ e $\overline{\boldsymbol{z}}$ tais que

$$\overline{y}_e := \begin{cases} 1 & \text{se } e \in \partial^+(S), \\ 0 & \text{em caso contrário,} \end{cases} \quad \text{e} \quad \overline{z}_u := \begin{cases} 1 & \text{se } u \in S, \\ 0 & \text{em caso contrário.} \end{cases}$$

$\overline{\boldsymbol{y}}$ e $\overline{\boldsymbol{z}}$ são dual-viáveis e fornecem valor objetivo igual à $c\big(\partial^+(S)\big)$; logo, são dual-ótimas. Além disso, se $\boldsymbol{c}$ é inteiro, o primal também admite uma solução inteira (exercício).

Aplicação: Decomposição de Custos

Folgas complementares podem ser utilizadas para caracterizar otimização linear sobre intersecção de poliedros.

Para $p \geq 1$ inteiro, sejam

$$\mathcal{P}_k := \left\{ \boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}^{(k)}\boldsymbol{x} = \boldsymbol{b}^{(k)} \right\} \qquad (k \in [p])$$

poliedros e considere o PL

$$\max \left\{ \boldsymbol{c}^\top \boldsymbol{x} : \boldsymbol{x} \in \bigcap_{i=1}^{p} \mathcal{P}_i \right\}. \qquad (P)$$

Suponha que $\boldsymbol{c}$ admite uma *decomposição de custos*, isto é, existem $\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_p \in \mathbb{R}^n$ tais que $\boldsymbol{c} = \sum_{i=1}^{p} \boldsymbol{c}_i$.

Para $k \in [p]$, considere os PLs

$$\max \left\{ \boldsymbol{c}_k^\top \boldsymbol{x} : \boldsymbol{x} \in \mathcal{P}_k \right\}. \qquad (P_k)$$

Aplicação: Decomposição de Custos

Teorema (Otimização sobre intersecção de poliedros)

Uma solução $\boldsymbol{x}^* \in \mathbb{R}^n$ é ótima para (P) se e somente se existe uma decomposição de custos para $\boldsymbol{c}$ tal que $\boldsymbol{x}^*$ é ótima para todos os (P_k), $k \in [p]$.

Prova. (Suficiência) Sejam (D_k) os duais de (P_k), dados por:

$$\min \left\{ \boldsymbol{y}_k^\top \boldsymbol{b}^{(k)} : \boldsymbol{y}_k^\top \boldsymbol{A}^{(k)} \geq \boldsymbol{c}_k^\top,\ \boldsymbol{y}_k \geq \boldsymbol{0} \right\}. \qquad (D_k)$$

Sejam $\widehat{\boldsymbol{x}}$ solução ótima para (P_k), $\widehat{\boldsymbol{y}}_k$ soluções ótimas para (D_k) e $\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_p$ uma decomposição de custos para $\boldsymbol{c}$.

Por dualidade forte em (P_k)-(D_k), $\boldsymbol{c}_k^\top \widehat{\boldsymbol{x}} = \widehat{\boldsymbol{y}}_k^\top \boldsymbol{b}^{(k)}$, o que implica em $\boldsymbol{c}^\top \widehat{\boldsymbol{x}} = \sum_{k=1}^{p} \widehat{\boldsymbol{y}}_k^\top \boldsymbol{b}^{(k)}$.

Aplicação: Decomposição de Custos

Temos ainda que $\widehat{\boldsymbol{x}}$ é viável para (P) e que $(\widehat{\boldsymbol{y}}_1, \ldots, \widehat{\boldsymbol{y}}_p)$ é viável para (D), o dual de (P):

$$\min \left\{ \sum_{k=1}^{p} \widehat{\boldsymbol{y}}_k^\top \boldsymbol{b}^{(k)} : \sum_{k=1}^{p} \boldsymbol{y}_k^\top \boldsymbol{A}^{(k)} \geq \boldsymbol{c},\, \widehat{\boldsymbol{y}}_k \geq \boldsymbol{0} \right\}. \qquad (D)$$

Por dualidade fraca em (P)-(D), o resultado segue.

(Necessidade) Sejam $\widehat{\boldsymbol{x}}$ e $(\widehat{\boldsymbol{y}}_1, \ldots, \widehat{\boldsymbol{y}}_p)$ soluções ótimas para (P) e (D), respectivamente. Claramente, $\widehat{\boldsymbol{x}}$ é viável para (P_k).

Defina $\boldsymbol{c}_k = \widehat{\boldsymbol{y}}_k^\top \boldsymbol{A}^{(k)}$ de forma que $\widehat{\boldsymbol{y}}_k$ seja viável para (D_k). Note que $\boldsymbol{c}_k$ não é uma decomposição de custos para $\boldsymbol{c}$.

De qualquer forma, como $(\widehat{\boldsymbol{y}}_1, \ldots, \widehat{\boldsymbol{y}}_p)$ é viável para (D), temos que

$$\sum_{k=1}^{p} \boldsymbol{c}_k = \sum_{k=1}^{p} \widehat{\boldsymbol{y}}_k^\top \boldsymbol{A}^{(k)} \geq \boldsymbol{c}.$$

Aplicação: Decomposição de Custos

Logo, $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_k$ é uma "cobertura" de $\boldsymbol{c}$.

Por dualidade fraca em (P_k)-(D_k), segue que $\boldsymbol{c}_k^\top \widehat{\boldsymbol{x}} = \widehat{\boldsymbol{y}}_k^\top \boldsymbol{b}^{(k)}$. Somando para todo k,

$$\boldsymbol{c}^\top \widehat{\boldsymbol{x}} \leq \sum_{k=1}^{p} \boldsymbol{c}_k^\top \widehat{\boldsymbol{x}} \leq \sum_{k=1}^{p} \widehat{\boldsymbol{y}}_k^\top \boldsymbol{b}^{(k)}.$$

A dualidade forte em (P)-(D) garante que as desigualdades acima são igualdades. Logo, é o caso de que

$$\boldsymbol{c}^\top \widehat{\boldsymbol{x}} = \sum_{k=1}^{p} \widehat{\boldsymbol{y}}_k^\top \boldsymbol{b}^{(k)}$$

e, consequentemente, $\widehat{\boldsymbol{x}}$ é ótimo para (P_k) e $\widehat{\boldsymbol{y}}_k$ é ótimo para (D_k).

Aplicação: Decomposição de Custos

Suponha agora que $\sum_{k=1}^{p}(\boldsymbol{c}_k)_j > c_j$ para algum j. Logo, para o mesmo j,

$$\sum_{k=1}^{p} \widehat{\boldsymbol{y}}_k^\top \boldsymbol{A}_j^{(k)} > c_j$$

e, por folgas complementares em (P_k)-(D_k), temos que $\widehat{x}_j = 0$.

Escolhendo um ℓ e reduzindo $(\boldsymbol{c}_\ell)_j$ até que $\sum_{k=1}^{p}(\boldsymbol{c}_\ell)_j = c_j$ não altera a otimalidade de $\widehat{\boldsymbol{x}}$ em (P_k), para $k \neq \ell$.

Como $\widehat{x}_j = 0$, o teorema de folgas complementares em (P_k)-(D_k) ainda garante que $\widehat{\boldsymbol{x}}$ é ótimo em (P_ℓ), concluindo a prova. $\square$

Nota: Programação Disjuntiva

Nota: A otimização sobre a união de poliedros

$$\max \left\{ \boldsymbol{c}^\top \boldsymbol{x} : \boldsymbol{x} \in \bigcup_{i=1}^{p} \mathcal{P}_i \right\}$$

é chamada *programação disjuntiva*. Assim como o caso de intersecção (que acabamos de mostrar), a otimização sobre união também se decompõe naturalmente.

Infelizmente no caso da união, a decomposição natural não é (tão) útil, já que a união de poliedros (conjuntos convexos) raramente é um poliedro (conjunto convexo).

Através da introdução de variáveis binárias adicionais (*lifting*) e da obtenção de novas restrições para a região viável (*projecting*), é possível realizar a otimização ou um relaxamento dela.

De qualquer forma, a um custo computacional (bem) maior que o da intersecção.

Resumo: Condições de Otimalidade

Antes de prosseguirmos, convém resumirmos nossa principal conquista nesses slides: condições de otimalidade de PLs. Agora, devidamente provadas!

Sejam $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $\boldsymbol{b} \in \mathbb{R}^n$ e $\boldsymbol{c} \in \mathbb{R}^m$. Para o PL na forma padrão

$$\min \left\{ \boldsymbol{c}^\top \boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\, \boldsymbol{x} \geq \boldsymbol{0} \right\},$$

$\boldsymbol{x} \in \mathbb{R}^n$ é uma solução ótima se e somente se satisfaz as *condições de otimalidade*:

(1) $\boldsymbol{x}$ é primal-viável: $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, $\boldsymbol{x} \geq \boldsymbol{0}$ e existe $\boldsymbol{y} \in \mathbb{R}^m$ tal que

(2) $\boldsymbol{y}$ é dual-viável: $\boldsymbol{y}^\top \boldsymbol{A} = \boldsymbol{c}^\top$, e

(3) não há lacuna de dualidade (= *duality gap*): $\boldsymbol{c}^\top \boldsymbol{x} = \boldsymbol{y}^\top \boldsymbol{b}$.

Definindo as variáveis de folga $\boldsymbol{s} =: \boldsymbol{c}^\top - \boldsymbol{y}^\top \boldsymbol{A} \geq \boldsymbol{0}$, (3) pode ser substituída por

(3') folgas complementares satisfeitas: $x_i s_i = 0$, para $i \in [n]$.

Resumo: Condições de Otimalidade

As condições de otimalidade sugerem diferentes estratégias algorítmicas: iterativamente

- melhorar soluções primal-viáveis até que a lacuna de dualidade seja fechada,
- melhorar soluções dual-viáveis até que a lacuna de dualidade seja fechada,
- melhorar, alternadamente, soluções primal- e dual viáveis até que as folgas complementares sejam satisfeitas.

Cada estratégia admite variações (oriundas de detalhes dos princípios algorítmicos distintos que podem ser usados).

Exemplos "populares" das estratégias acima são os algoritmos Simplex (Primal) e Simplex Dual, e a classe de Métodos Primal-Dual.

Antes de descrevê-mo-los, no entanto, precisamos ir mais afundo na estrutura de cones e poliedros — pois há detalhes pertinentes que ainda não sabemos como tratá-los.

Outline

Dualidade Forte e Folgas Complementares

Estruturas de Cones e Poliedros (I)

Cones: Poliédricos = Finitamente Gerados

Teorema (Farkas-Minkowski-Weyl)

Um cone é poliédrico se e somente se é finitamente gerado.

Prova. (Suficiência)
Para $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_m \in \mathbb{R}^n$, provamos que $\operatorname{cone}\big(\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_m\}\big)$ é poliédrico. Podemos supor que $\operatorname{lin.span}(\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_m\}\big) = \mathbb{R}^n$.

De fato, todo meio-espaço H de $\operatorname{lin.span}(\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_m\}\big)$ pode ser estendido a um meio-espaço H' de $\mathbb{R}^n$ tal que $H = H' \cap \operatorname{lin.span}(\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_m\}\big)$.

Considere todos os meio-espaços $H = \{\boldsymbol{x} : \boldsymbol{c}^\top \boldsymbol{x} \leq 0\}$ de $\mathbb{R}^n$ tais que $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_m \in H$ e tais que o hiperplano $\{\boldsymbol{x} : \boldsymbol{c}^\top \boldsymbol{x} = 0\}$ é gerado por $n-1$ vetores linearmente independentes dentre $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_m$.

Pelo Teorema Fundamental de Inequações, $\operatorname{cone}\big(\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_m\}\big)$ é a intersecção destes meio-espaços. Como o número de meio-espaços é finito (no máximo, $\binom{m}{n-1}$), o cone é poliédrico.

Cones: Poliédricos = Finitamente Gerados

(Necessidade) Seja $\mathcal{C} = \{\boldsymbol{x} : \boldsymbol{a}_1^\top \boldsymbol{x} \leq 0, \ldots, \boldsymbol{a}_m^\top \boldsymbol{x} \leq 0\}$ um cone poliédrico. A suficiência garante que existem vetores $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_\ell$ tais que

$$\operatorname{cone}\big(\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m\}\big) = \{\boldsymbol{x} : \boldsymbol{b}_1^\top \boldsymbol{x} \leq 0, \ldots, \boldsymbol{b}_\ell^\top \boldsymbol{x} \leq 0\}. \qquad (\boxtimes)$$

Resta mostrar, então, que $\operatorname{cone}\big(\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_d\}\big) = \mathcal{C}$, pois isto implica que C é finitamente gerado.

Por $(\boxtimes)$, temos que $\boldsymbol{b}_j^\top \boldsymbol{a}_i \leq 0$, para $i \in [m]$ e $j \in [\ell]$. Logo, $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_\ell \in \mathcal{C}$ e $\operatorname{cone}\big(\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_\ell\}\big) \subseteq \mathcal{C}$.

Para a recíproca, suponha que $\exists\, \boldsymbol{y} \in \mathcal{C}$ tal que $\boldsymbol{y} \notin \operatorname{cone}\big(\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_\ell\}\big)$. Por suficiência, $\operatorname{cone}\big(\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_\ell\}\big)$ é poliédrico. Logo, existe um vetor $\boldsymbol{w}$ tal que $\boldsymbol{w}^\top \boldsymbol{b}_1, \ldots, \boldsymbol{w}^\top \boldsymbol{b}_\ell \leq 0$ e $\boldsymbol{w}^\top \boldsymbol{y} > 0$. Por $(\boxtimes)$, $\boldsymbol{w} \in \operatorname{cone}\big(\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m\}\big)$, o que implica que $\boldsymbol{w}^\top \boldsymbol{x} \leq 0$ para todo $\boldsymbol{x} \in \mathcal{C}$. Isso, no entanto, contradiz que $\boldsymbol{y} \in \mathcal{C}$ e $\boldsymbol{w}^\top \boldsymbol{y} > 0$. $\square$

Homogenização de Poliedros

Como próximo passo, vamos mostrar que todo poliedro $\mathcal{P}$ é uma soma de Minkowski entre um politopo $\mathcal{Q}$ e um cone poliédrico $\mathcal{C}$ (o Teorema de Motzkin-Weyl).

É possível mostrar este resultado diretamente, via Fourier-Motzkin, mas os vetores em $\mathcal{Q} + \mathcal{C}$ são um pouco difíceis de manipular.

Tudo fica mais fácil se homogenizarmos o espaço afim de dimensão n em um espaço linear de dimensão $n+1$, com uma variável extra, e com $\boldsymbol{x} \in \mathbb{R}^n$ sendo mapeado em $\begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix} \in \mathbb{R}^{n+1}$. Desta forma, $\mathcal{P} = \{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b},\, \boldsymbol{x} \in \mathbb{R}^n\}$ é mapeado em

$$\mathcal{C}_{\mathcal{P}} := \left\{ \begin{pmatrix} \lambda \\ \boldsymbol{x} \end{pmatrix} : \boldsymbol{x} \in \mathbb{R}^n,\, \lambda \in \mathbb{R}_{\geq 0},\, \boldsymbol{A}\boldsymbol{x} - \lambda \boldsymbol{b} \leq \boldsymbol{0} \right\},$$

e podemos utilizar o resultado de Farkas-Minkowski-Weyl em $\mathcal{C}_{\mathcal{P}}$.

Nota: $\mathcal{P}$ é a projeção de $\mathcal{C}_{\mathcal{P}} \cap \left\{ \begin{pmatrix} \lambda \\ \boldsymbol{x} \end{pmatrix} \in \mathbb{R}^{n+1} : \lambda = 1 \right\}$ em $\mathbb{R}^n$.

Poliedros = Politopos + Cones

Teorema (Motzkin-Weyl: Decomposição de Poliedros)

$\mathcal{P} \subseteq \mathbb{R}^n$ é um poliedro se e somente se $\mathcal{P} = \mathcal{Q} + \mathcal{C}$ para algum politopo $\mathcal{Q}$ e algum cone poliédrico $\mathcal{C}$.

Prova. (Necessidade)
Seja $\mathcal{P} = \{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}\} \subseteq \mathbb{R}^n$ um poliedro e considere o cone poliédrico resultante da homogenização de $\mathcal{P}$,

$$\mathcal{C}_{\mathcal{P}} := \left\{ \begin{pmatrix} \lambda \\ \boldsymbol{x} \end{pmatrix} : \boldsymbol{x} \in \mathbb{R}^n,\, \lambda \in \mathbb{R}_{\geq 0},\, \boldsymbol{A}\boldsymbol{x} - \lambda \boldsymbol{b} \leq \boldsymbol{0} \right\}.$$

Pelo Teorema de Farkas-Minkowski-Weyl, $\mathcal{C}_{\mathcal{P}}$ é finitamente gerado:

$$\operatorname{cone}\left(\left\{ \begin{pmatrix} \lambda_1 \\ \boldsymbol{x}_1 \end{pmatrix}, \ldots, \begin{pmatrix} \lambda_m \\ \boldsymbol{x}_m \end{pmatrix} \right\} \right).$$

Podemos supor que $\lambda_i \in \{0, 1\}$.

Poliedros = Politopos + Cones

Seja $\mathcal{Q} = \text{conv}\left(\{\boldsymbol{x}_i : \lambda_i = 1\}\right)$ e $\mathcal{C} = \text{cone}\left(\{\boldsymbol{x}_i : \lambda_i = 0\}\right)$. Temos que $\boldsymbol{x} \in \mathcal{P}$ se e somente se $\begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix} \in \mathcal{C}_{\mathcal{P}}$, que ocorre se e somente se

$$\begin{pmatrix} 1 \\ \boldsymbol{x} \end{pmatrix} \in \text{cone}\left(\left\{ \begin{pmatrix} \lambda_1 \\ \boldsymbol{x}_1 \end{pmatrix}, \ldots, \begin{pmatrix} \lambda_m \\ \boldsymbol{x}_m \end{pmatrix} \right\}\right).$$

Logo, $\boldsymbol{x} \in \mathcal{Q}$ ou $\boldsymbol{x} \in \mathcal{C}$ e assim, $\mathcal{P} = \mathcal{Q} + \mathcal{C}$.

(Suficiência)
Seja $\mathcal{P} = \mathcal{Q} + \mathcal{C}$ para algum politopo $\mathcal{Q} = \text{conv}\left(\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m\}\right)$ e cone poliédrico $\mathcal{C} = \text{cone}\left(\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_\ell\}\right)$. Utilizando novamente o conceito de homogenização, temos que $\boldsymbol{x}_0 \in \mathcal{P}$ se e somente se

$$\begin{pmatrix} 1 \\ \boldsymbol{x}_0 \end{pmatrix} \in \text{cone}\left(\left\{ \begin{pmatrix} 1 \\ \boldsymbol{x}_1 \end{pmatrix}, \ldots, \begin{pmatrix} 1 \\ \boldsymbol{x}_m \end{pmatrix}, \begin{pmatrix} 0 \\ \boldsymbol{y}_1 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ \boldsymbol{y}_\ell \end{pmatrix} \right\}\right).$$

Poliedros = Politopos + Cones

Pelo Teorema de Farkas-Minkowski-Weyl, o cone acima é igual a

$$\left\{ \begin{pmatrix} \lambda \\ \boldsymbol{x} \end{pmatrix} : \boldsymbol{A}\boldsymbol{x} + \lambda \boldsymbol{b} \leq \boldsymbol{0} \right\}$$

para alguma matriz $\boldsymbol{A}$ e vetor $\boldsymbol{b}$. Logo, $\boldsymbol{x}_0 \in \mathcal{P}$ se e somente se $\boldsymbol{A}\boldsymbol{x}_0 \leq -\boldsymbol{b}$. Portanto, $\mathcal{P}$ é um poliedro. $\square$

Teorema (Minkowski: Base Finita para Politopos)

$\mathcal{P} \subseteq \mathbb{R}^n$ é um politopo se e somente se $\mathcal{P}$ é um poliedro limitado.

Prova. Direta do teorema anterior com $\mathcal{C} = \emptyset$. $\square$

Nota: Devido à limitação do poliedro e o consequente não envolvimento de um cone em uma soma de Minkowski, uma prova da Base Finita para Politopos pode ser obtida facilmente via Fourier-Motzkin.