EPR202 – Métodos de Otimização Aplicados à Eng. de Produção – 2020.QS

Lista 1 – Universidade Federal do ABC

Aritanan Gruber

aritanan.gruber@ufabc.edu.br

http://professor.ufabc.edu.br/∼aritanan.gruber

Exercícios marcados com ▶ devem ser entregues via Tidia até 23/11/2020 às 00h00m.

Notação. Exercícios marcados com: ♦ são operacionais e envolvem a execução de cálculos simples ou o fornecimento de (contra) exemplos; ♥ são sobre teoria e requerem provas matemáticas; ♣ requerem a construção de algoritmos; ♠ envolvem matemática de maneira mais geral, não limitada à teoria de otimização; ★ são exercícios ligeiramente mais difíceis; finalmente, ▶ são exercícios que devem ser entregues via Tidia.

1. ▶ ♦ *This is just a drill (CBD@).* Considere o PL abaixo:

$$\begin{aligned}
\max_{x_i} \quad & -x_1 + 5x_2 + 2x_3 - 7x_4 - x_5 \\
\text{s.t.} \quad & x_2 + x_3 + x_4 \geq 13 \\
& x_1 - x_2 + 2x_4 + 2x_5 \leq 4 \\
& x_2, x_4, x_5 \geq 0 \\
& x_3 \geq -2
\end{aligned}$$

   (a) Reformule-o nas formas canônica e padrão.

   (b) Determine os PLs duais das duas formas e estabeleça equivalências.

   (c) O PL fornecido é inviável, ilimitado ou possui solução ótima? Justifique adequadamente.

2. ♦ *Absolute values (BT-1.4).* Reformule o programa abaixo como um PL.

$$\begin{aligned}
\min \quad & 2x_1 + 3|x_2 - 10| \\
\text{s.t.} \quad & |x_1 + 2| + |x_2| \leq 5
\end{aligned}$$

3. ▶ ♥ *The moment problem (BT-1.7).* Suponha que Z é uma variável aleatória que toma valores $0, 1, \ldots, K$ com probabilidades $p_0, p_1, \ldots, p_K$; isto é, $\mathbb{P}\mathrm{r}[Z = j] = p_j$. Dados os valores dos primeiros dois momentos de Z,

$$\mathbb{E}[Z] = \sum_{j=0}^{K} j\, p_j \qquad \text{e} \qquad \mathbb{E}[Z^2] = \sum_{j=0}^{K} j^2 p_j,$$

   queremos obter limitantes inferior e superior no valor do quarto momento: $\mathbb{E}[Z^4] = \sum_{j=0}^{K} j^4 p_j$. Mostre como isso pode ser feito via programação linear.

4. ▶ ♦ ♥ *Using complementary slackness (KV@).* Sejam $G = (V, E)$ um digrafo, $c : E \to \mathbb{R}_{\geq 0}$ uma função custo, $E_1, E_2 \subseteq E$ e $s, t \in V$. Considere o seguinte PL:

$$\begin{aligned}
\min \quad & \sum_{e \in E} c(e) y_e \\
\text{s.t.} \quad & y_e \geq z_w - z_v \qquad e = (v, w) \in E \\
& z_t - z_s = 1 \\
& y_e \geq 0 \qquad e \in E_1 \\
& y_e \leq 0 \qquad e \in E_2
\end{aligned}$$

   Claramente, y_e e z_u são as variáveis, com $e \in E$ e $u \in V$.

(a) Qual problema o PL acima modela? Explique!

(b) Determine o dual do PL acima.

(c) Prove que existe uma solução ótima $(\boldsymbol{y}, \boldsymbol{z})$ e um conjunto X com $s \in X \subseteq V \setminus \{t\}$ tal que

$$\begin{aligned} y_e &= 1 && \text{para } e \in \delta^+(X), \\ y_e &= -1 && \text{para } e \in \delta^-(X) \setminus E_1, \\ y_e &= 0 && \text{para todos os demais arcos } e. \end{aligned}$$

Lembre-se que $\delta^+(X) := \{(x, y) \in E : x \in X, y \in V \setminus X\}$ e $\delta^-(X) := \{(y, x) \in E : x \in X, y \in V \setminus X\}$ são os conjuntos de arcos que saem e entram em X, respectivamente.

Dica: Considere as condições de folgas complementares para os arcos entrando ou saindo de $\{v \in V : z_v \leq z_s\}$.

5. ► ♥ ♣ *Fourier-Motzkin elimination procedure (KV@).* Considere o sistema de m inequações $\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}$ em n variáveis. Multiplicando cada linha por uma constante positiva, podemos supor que a primeira coluna de $\boldsymbol{A}$ é um vetor com entradas em $\{-1, 0, 1\}$. Assim, podemos re-escrever $\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}$ como

$$\begin{aligned} (\boldsymbol{a}_i')^\top \boldsymbol{x}' &\leq b_i && (i = 1, \ldots, m_1), \\ -x_1 + (\boldsymbol{a}_j')^\top \boldsymbol{x}' &\leq b_j && (j = m_1 + 1, \ldots, m_2), \\ x_1 + (\boldsymbol{a}_k')^\top \boldsymbol{x}' &\leq b_k && (k = m_2 + 1, \ldots, m), \end{aligned}$$

em que $\boldsymbol{x}' = (x_2, \ldots, x_n)$ e $\boldsymbol{a}_1', \ldots, \boldsymbol{a}_m'$ são as linhas de $\boldsymbol{A}$ sem a primeira entrada (correspondente à primeira coluna). A variável x_1 pode ser eliminada.

(a) Prove que $\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}$ tem uma solução se e somente se o sistema

$$\begin{aligned} (\boldsymbol{a}_i')^\top \boldsymbol{x}' &\leq b_i && (i = 1, \ldots, m_1), \\ (\boldsymbol{a}_j')^\top \boldsymbol{x}' - b_j &\leq b_k - (\boldsymbol{a}_k')^\top \boldsymbol{x}' && (j = m_1 + 1, \ldots, m_2;\ k = m_2 + 1, \ldots, m) \end{aligned}$$

tem solução.

(b) Mostre que, quando iterada, a transformação acima produz um algoritmo que permite determinar uma solução para $\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}$, ou provar sua inviabilidade.

(c) Use o algoritmo acima e forneça uma prova direta da afirmação: *existe um vetor $\boldsymbol{x}$ com $\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}$ se e somente se $\boldsymbol{y}^\top \boldsymbol{b} \geq 0$ para cada vetor $\boldsymbol{y} \geq 0$ tal que $\boldsymbol{y}^\top \boldsymbol{A} = \boldsymbol{0}$.*

(d) Mostre que o item anterior implica o Teorema Forte de Dualidade: *se os poliedros*

$$\mathcal{P} := \{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}\} \qquad \text{e} \qquad \mathcal{D} := \{\boldsymbol{y} : \boldsymbol{y}^\top \boldsymbol{A} = \boldsymbol{c}, \boldsymbol{y} \geq 0\}$$

são não vazios, então $\max\{\boldsymbol{c}^\top \boldsymbol{x} : \boldsymbol{x} \in \mathcal{P}\} = \min\{\boldsymbol{y}^\top \boldsymbol{b} : \boldsymbol{y} \in \mathcal{D}\}$.

6. ► ♥ ♠ *Funções convexas (BT@).* Uma função $f : \mathbb{R}^n \to \mathbb{R}$ é *convexa* se para todos $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ e $\lambda \in [0, 1]$,

$$f\big((1 - \lambda)\boldsymbol{x} + \lambda \boldsymbol{y}\big) \leq (1 - \lambda) f(\boldsymbol{x}) + \lambda f(\boldsymbol{y}).$$

(a) Seja $c \in \mathbb{R}$ uma constante. Mostre que o conjunto $S = \{\boldsymbol{x} \in \mathbb{R}^n : f(\boldsymbol{x}) \leq c\}$ é convexo.

(b) Sejam $f_1, f_2, \ldots, f_m : \mathbb{R}^n \to \mathbb{R}$ funções convexas e considere $f(\boldsymbol{x}) = \sum_{i=1}^m f_i(\boldsymbol{x})$. Prove que f é convexa.

(c) Uma função $f : \mathbb{R}^n \to \mathbb{R}$ é *côncava* se $-f$ é convexa. Suponha que $f : \mathbb{R}^n \to \mathbb{R}$ é simultaneamente convexa e côncava. Prove que f é uma função afim.

7. ► ♥ *Teoremas de Carathéodory.* Prove que:

(a) Se $X \subseteq \mathbb{R}^n$ e $\boldsymbol{y} \in \operatorname{conv}(X)$, então existem $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{n+1} \in X$ tais que $\boldsymbol{y} \in \operatorname{conv}(\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{n+1}\})$.

(b) Se $X \subseteq \mathbb{R}^n$ e $\boldsymbol{y}, \boldsymbol{z} \in \operatorname{conv}(X)$, então existem $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n \in X$ tais que $\boldsymbol{y} \in \operatorname{conv}(\{\boldsymbol{z}, \boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\})$.

Comentários e Dicas

1. As partes (a) e (b) são mecânicas, servindo apenas para praticar as definições e os algoritmos envolvidos. A parte sobre as equivalências entre os duais em (b) espelha as equivalências entre os primais canônico e padrão; é simples mas vale à pena prestar atenção ao padrão existente — aguça a intuição.

   Já na parte (c), a inviabilidade ou ilimitação do primal P podem ser mostradas diretamente em P ou com o auxílio do dual D; no caso de solução ótima, precisamos do primal e do dual. Fizemos as duas formas (direta e via dual D) nos slides para um exemplo, pois é interessante saber operar com as duas. Recomendo que você faça o mesmo nesta questão.

   Lembre-se do Teorema Fraco de Dualidade: para um primal $P : \max\{\boldsymbol{c}^\top \boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b}, \boldsymbol{x} \ge \boldsymbol{0}\}$ e um dual $D : \min\{\boldsymbol{y}^\top \boldsymbol{b} : \boldsymbol{y}^\top \boldsymbol{A} \ge \boldsymbol{c}, \boldsymbol{y} \ge \boldsymbol{0}\}$ temos que $\boldsymbol{c}^\top \boldsymbol{x} \le \boldsymbol{y}^\top \boldsymbol{b}$, para todo $\boldsymbol{x}$ primal-viável e $\boldsymbol{y}$ dual-viável.

   Consequências diretas:

   - se P é ilimitado, então D é inviável, ($\implies$ se D é viável, P não é ilimitado,)
   - se D é ilimitado, então P é inviável,
   - se $\boldsymbol{c}^\top \boldsymbol{x} = \boldsymbol{y}^\top \boldsymbol{b}$, com $\boldsymbol{x}$ primal-viável e $\boldsymbol{y}$ dual-viável, então $\boldsymbol{x}$ e $\boldsymbol{y}$ são soluções ótimas (primal e dual).

2. Aplicação direta da técnica de eliminação de módulos via variável auxiliar:

   $$|x| \le a \iff x \le a \text{ ou } -x \le a \;\equiv\; -a \le x \le a.$$

   Veja slides sobre 'Regressão linear ℓ_1.'

3. Faça-se a pergunta: o que é dado e quais são as variáveis?

   Claramente, K, $\mathbb{E}[Z]$ e $\mathbb{E}[Z^2]$ são dados e p_j são variáveis. Se as probabilidades p_j fossem dadas, o problema seria trivial! I.é, simples avaliação de $\sum_{j=0}^{K} j^4 p_j$. Pense no problema da mochila e em variações dele.

4. Não há uma resposta única e/ou correta para o item (a), mas sim respostas que fazem ou não sentido e que são ou não precisas — o verbo 'explique' em vez de 'prove' já dava a dica. Não há necessidade de utilizar conceitos ou jargão de qualquer área de aplicação que não sejam fluxos e/ou caminhos em digrafos. O intuito desta questão é gerar questionamentos sobre a fronteira entre o que é objetivamente quantificável é o que é subjetivo (em geral, abstraído pelos modelos).

   A parte (b) é mecânica, mas interessante: observe que $y_e \ge 0$ se $e \in E_1$ e $y_e \le 0$ se $e \in E_2$; logo, E_1 e E_2 influenciam os sinais das variáveis primais $x_e^{v,w}$. A parte (c) é um uso direto de folgas complementares nos conjuntos mencionados na dica fornecida no enunciado.

5. A eliminação de Fourier-Motzkin faz o papel da eliminação Gaussiana para inequações. A diferença é: a última é polinomial no tamanho do sistema $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$; a primeira, no pior caso, é exponencial no tamanho de $\boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b}$.

   Seja $\boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b}$ o sistema em n variáveis e com m restrições após a reescrita. Rearranjando o sistema em (a), após a eliminação da variável x_1, obtemos o sistema $\boldsymbol{A}'\boldsymbol{x}' \le \boldsymbol{b}'$

   $$\begin{aligned} (\boldsymbol{a}'_i)^\top \boldsymbol{x}' &\le b_i && (i = 1, \dots, m_1), \\ (\boldsymbol{a}'_j + \boldsymbol{a}'_k)^\top \boldsymbol{x}' &\le b_k + b_j && (j = m_1 + 1, \dots, m_2;\ k = m_2 + 1, \dots, m), \end{aligned}$$

   em $n-1$ variáveis e com $m_1 + (m_2 - m_1)(m - m_2) \ge m$ restrições (quando há igualdade?).

   Você deve mostrar que $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \iff \boldsymbol{A}'\boldsymbol{x}' = \boldsymbol{b}'$. Em palavras, a eliminação de Fourier-Motzkin preserva a consistência / inconsistência do sistema. Há duas implicações a serem mostradas:

   - $\boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b} \implies \boldsymbol{A}'\boldsymbol{x}' \le \boldsymbol{b}'$: essa é imediata; basta tomar uma solução $\boldsymbol{x}$ para $\boldsymbol{A}, \boldsymbol{b}$ e mostrar que $\boldsymbol{x}'$, obtido à partir de $\boldsymbol{x} = (x_1, \boldsymbol{x}')$ pela eliminação de x_1 satisfaz $\boldsymbol{A}'\boldsymbol{x}' \le \boldsymbol{b}'$.
   - $\boldsymbol{A}'\boldsymbol{x}' \le \boldsymbol{b}' \implies \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b}$: uma forma de fazer isso é tomar uma solução $\boldsymbol{x}'$ para $\boldsymbol{A}', \boldsymbol{b}'$ e mostrar que existe um $\alpha_1 \in \mathbb{R}$ tal que $\boldsymbol{x} = (\alpha_1, \boldsymbol{x}')$ é uma solução para $\boldsymbol{A}, \boldsymbol{b}$. A pergunta crucial (e óbvia) é, qual o valor de α_1? Ele pode ser obtido da reescrita de $\boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b}$:

     $$\begin{aligned} -\alpha_1 + (\boldsymbol{a}'_j)^\top \boldsymbol{x}' &\le b_j && (j = m_1 + 1, \dots, m_2), \\ \alpha_1 + (\boldsymbol{a}'_k)^\top \boldsymbol{x}' &\le b_k && (k = m_2 + 1, \dots, m). \end{aligned}$$

A parte (b) é uma simples indução no número de variáveis n. A necessidade na parte (c) é imediata; para suficiência, modifique ligeiramente a indução que você fez em (b).

Para a parte (d), é conveniente provar primeiro uma consequência afim de (c): supondo que $\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}$ é consistente, com $\boldsymbol{c}$ um vetor e δ um escalar, toda solução $\boldsymbol{x}$ de $\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}$ satisfaz $\boldsymbol{c}^\top\boldsymbol{x} \leq \delta$ se e somente se existe um vetor $\boldsymbol{y} \geq \boldsymbol{0}$ tal que $\boldsymbol{y}^\top\boldsymbol{A} = \boldsymbol{c}$ e $\boldsymbol{y}^\top\boldsymbol{b} \leq \delta$. A prova de (d) agora decorre desta variação afim de (c) e do teorema fraco de dualidade.

6. (a): Sejam $\boldsymbol{x}, \boldsymbol{y} \in S$ e defina $\boldsymbol{z}(\lambda) := (1-\lambda)\boldsymbol{x} + \lambda\boldsymbol{y}$ para $\lambda \in [0,1]$. O que pode ser dito sobre $f(\boldsymbol{z}(\lambda))$?

   (b) ou "a soma de funções convexas é uma função convexa:" imediata, basta organizar as contas e aplicar a definição de função convexa.

   Para a parte (c), considere que uma função $g : \mathbb{R}^n \to \mathbb{R}$ é *afim* se

$$g\left(\sum_{i=1}^{m} \lambda_i \boldsymbol{x}_i\right) = \sum_{i=1}^{m} \lambda_i g(\boldsymbol{x}_i)$$

   para todo $\boldsymbol{x}_i \in \mathbb{R}^n$ e $\lambda_i \in \mathbb{R}$ tal que $\sum_{i=1}^{m} \lambda_i = 1$.

7. (a): Em outras palavras, o teorema diz que qualquer elemento em $\operatorname{conv}(X)$ pode ser representado como uma combinação convexa de não mais que $n+1$ elementos de X.

   Para $\boldsymbol{y} \in \operatorname{conv}(X)$, seja $\sum_{i=1}^{m} \lambda_i \boldsymbol{x}_i$ com $\lambda_i \in [0,1]$ e $\sum_{i=1}^{m} \lambda_i = 1$ uma representação convexa para $\boldsymbol{y}$ com o menor valor de m possível. Logo, $\lambda_i > 0$. Claramente, se $m \leq n+1$ não há nada a fazer. O caso interessante é quando $m > n+1$. É possível tratá-lo via indução (em quê?) ou contradição. Vou comentar um pouco sobre o segundo...

   Considere o subespaço vetorial S paralelo a $\operatorname{aff.span}(X)$. Temos que os vetores $\boldsymbol{x}_2 - \boldsymbol{x}_1, \ldots, \boldsymbol{x}_m - \boldsymbol{x}_1 \in S$ e como $m - 1 > n$,

$$\sum_{i=2}^{m} \mu_i(\boldsymbol{x}_i - \boldsymbol{x}_1) = 0 \qquad (\bullet)$$

   para ao menos um $\mu_i \neq 0$; isto é, os $m-1$ vetores são linearmente dependentes.

   Estude o efeito que isso tem na representação de $\boldsymbol{y}$ e mostre que $(\bullet)$ implica que $\boldsymbol{y}$ pode ser representado como combinação convexa de menos de m vetores em X, contradizendo a hipótese de minimalidade.

   (b): Parecido e pode ser reduzido ao item (a). Como?