Lista 4 - Otimização Não Linear Inteira Mista
COS886 - 2021/3

Amanda Ferreira de Azevedo - afazevedo@cos.ufrj.br

PESC/UFRJ — 10 de fevereiro de 2022

Exercício 5.1

Considere o problema de otimização

$$\begin{array}{ll} \min & x^2+1, \\ \text{s.a.} & (x-2)(x-4) \leq 0, \end{array}$$

com variável $x \in \mathbb{R}$.

1. Análise do problema primal. Informe a região viável, o valor ótimo e a solução ótima, utilizando uma abordagem geométrica.
2. Lagrangiano e função dual. Plote o objetivo x^2+1 por x. No mesmo gráfico, informe a região viável, o valor ótimo e a solução ótima e plote o Lagrangiano $\mathcal{L}(x,\lambda)$ por x para alguns valores positivos de λ. Verifique a propriedade do limite inferior ($p^\star \geq \inf_x \mathcal{L}(x,\lambda)$ para $\lambda \geq 0$). Derive e apresente o gráfico da função dual lagrangiana g.
3. Problema dual lagrangiano. Formule o problema dual e verifique que é um problema côncavo de maximização. Encontre o valor ótimo e a solução ótima do dual. A dualidade forte é satisfeita?
4. Análise de sensibilidade. Seja $p^\star$ o valor ótimo do problema

$$\begin{array}{ll} \min & x^2+1 \\ \text{s.a.} & (x-2)(x-4) \leq u \end{array}$$

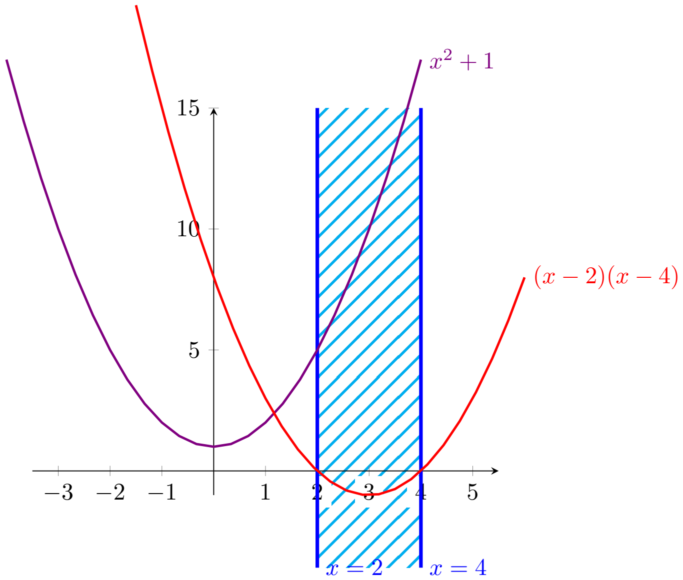
como uma função de parâmetro u. Plote $p^\star(u)$. Verifique que $\frac{dp^\star(0)}{du} = -\lambda^\star$.

Análise do problema primal

Da restrição do problema, temos

$$(x-2)(x-4) \leq 0 \Leftrightarrow 2 \leq x \leq 4$$

Assim, a região viável será delimitada pelo intervalo fechado $[2,4]$, como mostra a Figura abaixo.



Note que toda a reta $x = 2$ será a solução ótima uma vez que $x^2 + 1$ é monótona e aumenta para $x > 0$. Dessa forma, a solução ótima será $p^* = 2^2 + 1 = 5$.

$\boxed{\text{Lagrangiano e função dual}}$

Seja o Lagrangiano $L : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ associado à restrição considerada, como

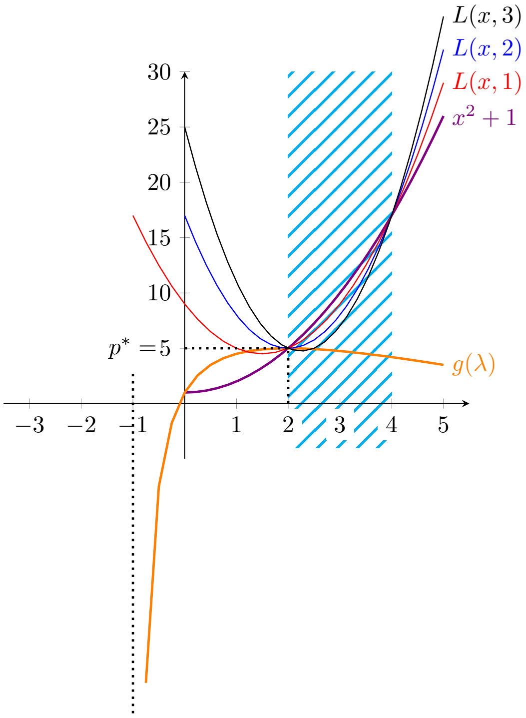
$$\begin{aligned} L(x, \lambda) &= \underbrace{x^2 + 1}_{\text{fobj}} + \lambda[(x-2)(x-4)] \\ &= x^2 + 1 + \lambda x^2 - 6\lambda x + 8\lambda \\ &= x^2(1+\lambda) - 6\lambda x + 8\lambda + 1 \end{aligned}$$

Onde λ é o multiplicador de Lagrange. O problema Dual Lagrangiano $g : \mathbb{R} \to \mathbb{R}$ será o valor mínimo do Lagragiano com relação a x. Note que, pela dualidade fraca, $g(\lambda) \leq L(\lambda, 2)$, uma vez que $L(\lambda, 2) = (\lambda+1)2^2 - 12\lambda + 8\lambda + 1 = 5 = p^*$. Para encontrar o mínimo de L vamos derivar com relação a x e igualar a zero. Dessa forma,

$$\begin{aligned} L'(\lambda, x) = 2x(1+\lambda)x - 6\lambda &= 0 \\ 2x(1+\lambda) &= 6\lambda \\ 2x + 2\lambda x &= 6\lambda \\ x(1+\lambda) &= 3\lambda \\ x &= \frac{3\lambda}{1+\lambda} \end{aligned}$$

Com isso, podemos definir a função Dual Lagrangiana,

$$\begin{aligned} g(\lambda) &= \min \quad L(x, \lambda) \\ &= \left(\frac{3\lambda}{1+\lambda}\right)^2 (1+\lambda) - 6\lambda\frac{3\lambda}{1+\lambda} + 8\lambda + 1 \\ &= \frac{9\lambda^2 - 18\lambda^2 + \lambda + 1 + 8\lambda^2 + 8\lambda}{1+\lambda} \\ &= \frac{-\lambda^2 + 9\lambda + 1}{1+\lambda} \end{aligned}$$



Observe que $g(\lambda)$ não está definida para valores $\lambda \leq -1$. Podemos notar também que a função dual é côncava em que $p^* = 5$ acontece exatamente quando $\lambda = 2$ e para outros valores de λ menor que p^*.

$\boxed{\text{Problema dual lagrangiano}}$

O problema Dual será exatamente a maximização de $g(\lambda)$, ou seja

$$\begin{array}{ll} \text{max} & \dfrac{-\lambda^2 + 9\lambda + 1}{1+\lambda} \\ \text{sujeito a} & \lambda \geq 0 \end{array}$$

A solução ótima deste problema dá exatamente $d^* = p^*$ quando $\lambda = 2$, como vimos anteriormente. Também vimos, pelo gráfico da função, que é de fato uma função côncava. Além disso, a dualidade forte é satisfeita, uma vez que $d^* = p^*$ e o problema é **estritamente** viável dado que $p^* > -\infty$.

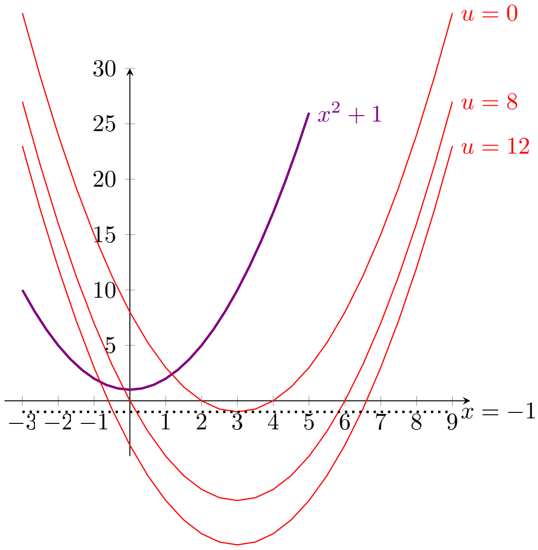
$\boxed{\text{Análise de sensibilidade}}$

Partindo do modelo

$$\begin{array}{ll} \text{min} & x^2 + 1 \\ \text{s.a.} & (x-2)(x-4) \leq u \end{array}$$

Desenvolvendo a restrição,

$$x^2 - 6x + 8 \leq u$$



Notamos que o menor valor que $x^2 - 6x + 8$ pode assumir é quando $y = -1$ e $x = 3$. Dessa forma, o problema não está definido para $u < -1$. Agora, vamos descobrir um intervalo para x para qualquer u,

$$x^2 - 6x + 8 = u \tag{1}$$

$$x^2 - 6x + 8 - u = 0 \tag{2}$$

$$x = \frac{6 \pm \sqrt{4(1+u)}}{2} \tag{3}$$

$$x = 3 \pm \sqrt{1+u} \tag{4}$$

Dessa forma, a região viável estará entre $[3 - \sqrt{1+u}, 3 + \sqrt{1+u}]$. Note, pela figura, que a partir de $u > 8$, o valor do máximo será majorado pela raiz de $x^2 + 1$. Dessa forma, para $-1 \leq u \leq 8$ a solução ótima será $x^*(u) = 3 - \sqrt{1+u}$ e para $u > 8$, $x^*(u) = 0$. Resumindo, substituindo na função objetivo

$$p^\star(u) = \begin{cases} \infty & u < -1 \\ 11 + u - 6\sqrt{1+u} & -1 \leq u \leq 8 \\ 1 & u \geq 8 \end{cases}$$

Exercício 5.13

Relaxação lagrangiana de um PL booleano. Um PL booleano é um problema de otimização dado na forma

$$\begin{array}{rl} \min & c^\top x, \\ \text{s.a.} & Ax \preceq b, \\ & x_i \in \{0,1\}, \quad i = 1, \dots, n, \end{array}$$

e, em geral, é bem dificil de resolver. No exercício 4.15 é apresentado uma relaxação LP desse problema

$$\begin{array}{rl} \min & c^\top x \\ \text{s.a.} & Ax \preceq b, \\ & 0 \leq x_i \leq 1, \quad i = 1, \dots, n, \end{array}$$

que é mais fácil de resolver e retorna um limite inferior do valor ótimo do PL booleano. Nesse problema vamos apresentar outro limite inferior do PL booleano e trabalhar na relação entre os dois limites inferiores.

1. Relaxaçáo lagrangiana. O PL booleano pode ser reformulado como o problema

$$\begin{array}{rl} \min & c^\top x, \\ \text{s.a.} & Ax \preceq b, \\ & x_i(1 - x_i) = 0, \quad i = 1, \dots, n, \end{array}$$

   que possui restrições quadráticas de igualdade. Encontre o dual lagrangiano deste problema. O valor ótimo do problema dual (que é convexo) retorna um limite inferior no valor ótimo do PL booleano. Esse método de encontrar um limite inferior do valor ótimo é chamado de relaxaçáo lagrangiana.

2. Mostre que, neste caso, o limite inferior recebido pela relaxação lagrangiana e pela relaxação linear são iguais.

Relaxaçáo lagrangiana

Dualizando as restrições com os multiplicadores $\lambda \geq 0$ e $\mu \geq 0$, o Lagrangiano é definido por

$$L(x, \lambda, \mu) = \underbrace{\sum_{i=1}^{n} c_i x_i}_{c^T x} + \underbrace{\sum_{i=1}^{n} \lambda_i (a_i x_i - b)}_{\mu^T(Ax-b)} - \mu \underbrace{\sum_{i=1}^{n} x_i + x_i^2}_{x^T \mathbf{diag}(\mu) x}$$

Dessa forma, podemos reescrever como

$$L(x, \lambda, \mu) = c^T x + \mu^T(Ax - b) - \mu^T x + x^T \mathbf{diag}(\mu) x$$

Ou seja,

$$x^T \mathbf{diag}(\mu) x + (c + A^T \lambda - \mu)^T x - b^T \lambda$$

Agora para encontrarmos o Dual, temos que minimizar com respeito a x.
Pela definição, temos que

$$\alpha = x^T A x = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij} x_i x_j$$

Derivando α com respeito ao k-ésimo elemento de x, temos

$$\frac{\partial \alpha}{\partial x_k} = \sum_{j=1}^{n} a_{kj} x_j + \sum_{i=1}^{n} a_{ik} x_i$$

Fazendo para todo $k = 1, 2, \dots, n$, temos

$$\frac{\partial \alpha}{\partial x} = x^T A^T + x^T A = x^T(A^T + A)$$

Como $A = \mathbf{diag}(\mu)$, teremos que

$$\frac{\partial \alpha}{\partial x} = 2x^T \mathbf{diag}(\mu) = 2x^T \mu$$

Continuando, teremos

$$L'(x, \lambda, \mu) = 2x^T \mu + c + A^T \lambda - \mu$$

Igualando a zero,

$$\begin{aligned} 2x^T \mu + c + A^T \lambda - \mu = 0 \\ 2x^T \mu &= \mu - c - A^T \lambda \\ x &= \frac{\mu - c - A^T \lambda}{2\mu} \end{aligned}$$

Substituindo em L,

$$\begin{aligned} \sum_{i=1}^{n} \mu_i x_i^2 - \sum_{i=1}^{n} (c_i + a_i^T \lambda - \mu_i) x_i - b^T \lambda &= \sum_{i=1}^{n} x_i(\mu_i x_i - c_i + a_i^T \lambda - \mu_i) - b^T \lambda \\ &= \sum_{i=1}^{n} \frac{\mu_i - c_i - a_i^T \lambda}{2\mu_i} (\mu_i x_i - c_i + a_i^T \lambda - \mu_i) - b^T \lambda \\ &= -b^T \mu - \frac{1}{4} \sum_{i=1}^{n} \frac{(c_i + a_i^T \lambda - \mu_i)^2}{\mu_i} \end{aligned}$$

Dessa forma, a função do dual será

$$g(\lambda, \mu) = -b^T \mu - \frac{1}{4} \sum_{i=1}^{n} \frac{(c_i + a_i^T \lambda - \mu_i)^2}{\mu_i}$$

E portanto, o problema dual será

$$\begin{array}{ll} \text{max} & -b^T \mu - \frac{1}{4} \sum_{i=1}^{n} \frac{(c_i + a_i^T \lambda - \mu_i)^2}{\mu_i} \\ \text{sujeito a} & \mu \succeq 0 \quad \lambda \geq 0 \end{array}$$

Mostre que, neste caso, o limite inferior recebido pela relaxação lagrangiana e pela relaxação linear são iguais.

Demonstração: O problema "relaxado"é descrito como

$$\begin{array}{rl} \min & c^\top x \\ \text{s.a.} & Ax \preceq b, \\ & 0 \leq x_i \leq 1, \quad i = 1, \ldots, n, \end{array}$$

Não consegui fazer. :(

■

Exercício 5.26

Considere o QCQP

$$\begin{array}{ll} \min & x_1^2 + x_2^2 \\ \text{s.a.} & (x_1 - 1)^2 + (x_2 - 1)^2 \leq 1 \\ & (x_1 - 1)^2 + (x_2 + 1)^2 \leq 1 \end{array}$$

com variável $x \in \mathbb{R}^2$.

1. Faça o desenho da região viável e as curvas de nível da função objetivo. Encontre o ponto ótimo $x^\star$ e o valor ótimo $p^\star$.

2. Informe as condições KKT. Diga se existem multiplicadores de Lagrange $\lambda_1^\star$ e $\lambda_2^\star$ que provem que $x^\star$ é ótimo.

3. Encontre o problema dual lagrangiano. A dualidade forte é satisfeita?

Faça o desenho da região viável e as curvas de nível da função objetivo. Encontre o ponto ótimo $x^\star$ e o valor ótimo $p^\star$.

Desenhando a região viável e as curvas de nível representadas pela função objetivo, temos

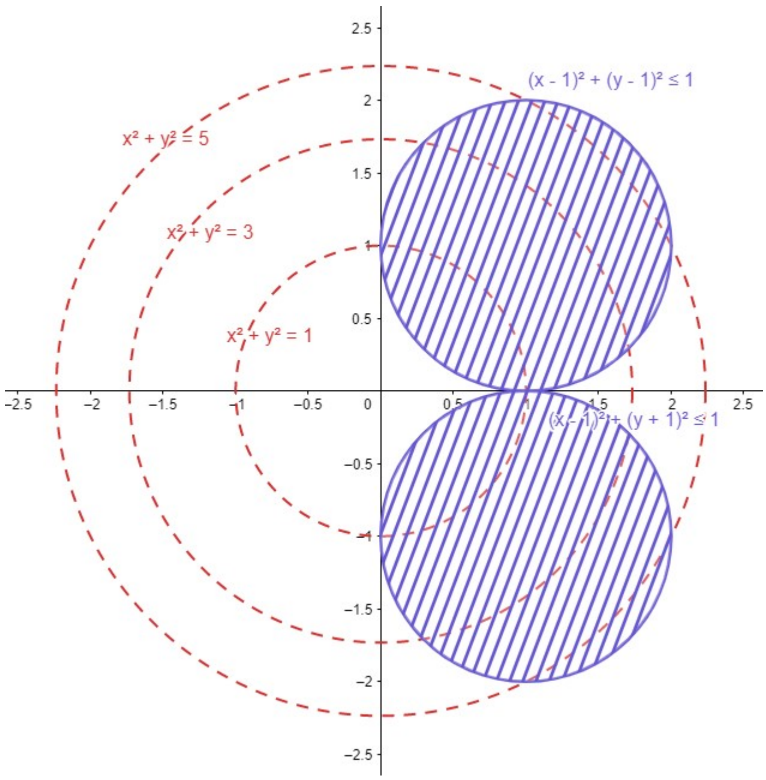


Figura 1: Região viável

A região viável será exatamente a interseção dos dois círculos violetas, que é exatamente no ponto $x = (1, 0)$. Como só há um ponto viável, esse é o ponto que nos dá a solução ótima. Ou seja, $p^* = 1$.

Informe as condições KKT.

Primeiro, precisamos ver que x^* satisfaz as restrições do problema primal

$$(x_1 - 1)^2 + (x_2 - 1)^2 \leq 1, \quad (x_1 - 1)^2 + (x_2 + 1)^2 \leq 1$$

Ambos os lados dão exatamente 1. Agora, precisamos mostrar que as variáveis duais associadas a estas restrições são não-negativas, isto é, ser dual viável.

$$\lambda_1 \geq 0, \lambda_2 \geq 0$$

Em seguida, devemos mostrar que x^* satisfaz as condições de complementariedade

$$\begin{aligned}
\lambda_1 \left[(x_1-1)^2+(x_2-1)^2-1\right] &= 0 \\
\lambda_2 \left[(x_1-1)^2+(x_2+1)^2-1\right] &= 0 \\
\lambda_1 \left[(x_1-1)^2+(x_2-1)^2-1\right] &= \lambda_2 \left[(x_1-1)^2+(x_2+1)^2-1\right] \\
-2\lambda_1+2\lambda_2 &= 0
\end{aligned}$$

Com essa equação, podemos ter diversas soluções pra λ_1, λ_2 que satisfazem a essa igualdade.

Por fim, devemos ver que x^* minimiza o Lagrangiano quando fixamos λ na solução ótima do dual. Em outras palavras, o gradiante do Lagrangiano com respeito a x se anula.

$$\begin{aligned}
2x_1+2\lambda_1(x_1-1)+2\lambda_1(x_2-1) &= 0 \\
2x_2++2\lambda_2(x_1-1)+2\lambda_2(x_2+1) &= 0
\end{aligned}$$

Da primeira equação, quando substituímos o ponto $x^* = (1,0)$ temos que $2 = 0$. Logo, as condições KKT não são satisfeitas e não existem λ_1 e λ_2 que provem que x^* é ótimo.

Encontre o problema dual lagrangiano. A dualidade forte é satisfeita?

A função dual Lagrangiana será dada por

$$g(\lambda_1,\lambda_2) = \min_{x_1,x_2} L(x_1,x_2,\lambda_1,\lambda_2)$$

$$\begin{aligned}
L(x_1,x_2,\lambda_1,\lambda_2) &= x_1^2+x_2^2+\lambda_1\left((x_1-1)^2+(x_2-1)^2-1\right)+\lambda_2\left((x_1-1)^2+(x_2+1)^2-1\right) \\
&= (1+\lambda_1+\lambda_2)\,x_1^2+(1+\lambda_1+\lambda_2)\,x_2^2-2(\lambda_1+\lambda_2)\,x_1-2(\lambda_1-\lambda_2)\,x_2+\lambda_1+\lambda_2
\end{aligned}$$

Vamos derivar e igualar a zero para encontrar os pontos críticos

$$L'(x_1,x_2,\lambda_1,\lambda_2) = 2(1+\lambda_1+\lambda_2)x_1+2(1+\lambda_1+\lambda_2)x_2-2(\lambda_1+\lambda_2)-2(\lambda_1-\lambda_2)$$

Agora igualando a zero

$$\begin{aligned}
2(1+\lambda_1+\lambda_2)x_1+2(1+\lambda_1+\lambda_2)x_2-2(\lambda_1+\lambda_2)-2(\lambda_1-\lambda_2) &= 0 \\
(1+\lambda_1+\lambda_2)x_1+(1+\lambda_1+\lambda_2)x_2 &= (\lambda_1+\lambda_2)+(\lambda_1-\lambda_2) \\
(1+\lambda_1+\lambda_2)x_1 &= (\lambda_1+\lambda_2) \\
(1+\lambda_1+\lambda_2)x_2 &= (\lambda_1-\lambda_2) \\
x_1 &= \frac{(\lambda_1+\lambda_2)}{(1+\lambda_1+\lambda_2)} \\
x_2 &= \frac{(\lambda_1-\lambda_2)}{(1+\lambda_1+\lambda_2)}
\end{aligned}$$

Substituindo em L, teremos a função dual Lagrangiana

$$
\begin{aligned}
g(\lambda_1, \lambda_2) &= (1+\lambda_1+\lambda_2)\left(\frac{(\lambda_1+\lambda_2)}{(1+\lambda_1+\lambda_2)}\right)^2 + (1+\lambda_1+\lambda_2)\left(\frac{(\lambda_1-\lambda_2)}{(1+\lambda_1+\lambda_2)}\right)^2 \\
&- 2\,(\lambda_1+\lambda_2)\,\frac{(\lambda_1+\lambda_2)}{(1+\lambda_1+\lambda_2)} - 2\,(\lambda_1-\lambda_2)\,\frac{(\lambda_1-\lambda_2)}{(1+\lambda_1+\lambda_2)} + \lambda_1+\lambda_2 \\
&= -\frac{(\lambda_1+\lambda_2)^2 + (\lambda_1-\lambda_2)^2}{1+\lambda_1+\lambda_2} + \lambda_1+\lambda_2
\end{aligned}
$$

Pra ser ponto de mínimo, como a segunda derivada dá $4(1+\lambda_1+\lambda_2)$, então $1+\lambda_1+\lambda_2 \geq 0$. Dessa forma, o problema dual Lagrangiano será dado por

$$
\begin{array}{ll}
\text{max} & -\frac{(\lambda_1+\lambda_2)^2+(\lambda_1-\lambda_2)^2}{1+\lambda_1+\lambda_2} + \lambda_1+\lambda_2 \\
\text{sujeito a} & \lambda_1, \lambda_2 \geq 0
\end{array}
$$

Exercício 5.27

Mínimos quadrados com restrições de igualdade. Considere o problema de mínimos quadrados com restrições de igualdade

$$\begin{array}{ll} \min & \|Ax-b\|_2^2, \\ \text{s.a.} & Gx = h, \end{array}$$

onde $A \in \mathbb{R}^{m\times n}$ com posto $A = n$, e $G \in \mathbb{R}^{p\times n}$ com posto $G = p$.

1. Encontre a função dual lagrangiana $g(\nu)$.
2. Informe as condições KKT.
3. A partir das condições KKT, apresente a expressão analítica para $\nu^\star$.

Encontre a função dual lagrangiana $g(\nu)$.

Seja $\nu \in R^p$ as variáveis duais associadas as restrições de igualdade. Temos o seguinte Lagrangiano

$$\begin{aligned} L(x,\nu) &= \|Ax-b\|^2 + v^T(Gx-h) \\ &= (Ax-b)^T(Ax-b) + v^TGx - v^Th \\ &= (x^TA^T - b^T)(Ax-b) + v^TGx - v^Th \\ &= x^TA^TAx - x^TA^Tb - b^TAx + b^Tb + v^TGx - v^Th \\ &= x^TA^TAx + (G^Tv - 2A^Tb)^Tx + b^Tb - v^Th \end{aligned}$$

Derivando L temos,

$$L'(x,\nu) = 2A^TAx + G^T\nu - 2A^Tb$$

Uma vez que $rank(A) = n$, podemos inverter A^TA. Igualando a zero, temos

$$x = (A^TA)^{-1}\left(A^Tb - \frac{1}{2}G^Tv\right)$$

Dessa forma, a função dual lagrangiana será

$$g(\nu) = -(1/4)\left(G^T\nu - 2A^Tb\right)^T\left(A^TA\right)^{-1}\left(G^T\nu - 2A^Tb\right) + b^Tb - v^Th$$

Encontre a função dual lagrangiana $g(\nu)$.

Se encontrarmos um ponto primal $x^* \in \mathbb{R}^n$ e um ponto dual $v^* \in R^p$ que satisfazem as condições de KKT então, por ser um problema de otimização convexa, temos que x^* e ν^* é o ponto ótimo, isto é, a dualidade forte acontece.

Primeiro, devemos verificar as **condições de estacionariedade**. Essa condição é exatamente a derivada do Lagrangiano que encontramos anteriormente no ponto dual nu^*.

$$L'(x,\nu*) = 2A^TAx + G^T\nu^* - 2A^Tb$$

Agora, podemos identificar a solução primal x^* que será, pela outro item

$$x^* = (A^TA)^{-1}\left(A^Tb - \frac{1}{2}G^Tv^*\right)$$

Assumiremos que será **primal viável**, isto é

$$Gx^* = h$$

E **dual viável**.

$\boxed{\text{A partir das condições KKT, apresente a expressão analítica para } \nu^{\star}.}$

Para encontrar v^* vamos substituir x^* nas restrições de igualdade

$$G(A^TA)^{-1}\left(A^Tb - \frac{1}{2}G^Tv^*\right) = h$$
$$G(A^TA)^{-1}A^Tb - G(A^TA)^{-1}\frac{1}{2}G^Tv^* = h$$

Isolando v^* e usando da hipótese que $rank(G) = p$ para concluir que $G(A^TA)^{-1}G^T$ é invertível, temos

$$\nu^* = -2\left(G\left(A^TA\right)^{-1}G^T\right)^{-1}\left(h - G\left(A^TA\right)^{-1}A^Tb\right)$$

Dessa forma, encontramos um ponto x^* e um ponto nu^* que satisfaz as condições de KKT, dessa forma, podemos assumir que $x^* = \nu^*$ é a solução ótima e a dualidade forte acontece. Para encontramos exatamente o valor de x^* basta substituir ν^*.

Exercício 5.43

Dual de um SOCP. Mostre que o dual do SOCP

$$\begin{array}{ll} \min & f^\top x, \\ \text{s.a.} & \|A_i x + b_i\|_2 \le c_i^\top x + d_i, \quad i = 1, \dots, m, \end{array}$$

com variáveis $x \in \mathbb{R}^n$, pode ser dado da seguinte forma

$$\begin{array}{ll} \max & \sum_{i=1}^m \left(b_i^\top u_i - d_i v_i\right), \\ \text{s.a.} & \sum_{i=1}^m \left(A_i^\top u_i - c_i v_i\right) + f = 0, \\ & \|u_i\|_2 \le v_i, \quad i = 1, \dots, m, \end{array}$$

com as variáveis $u_i \in \mathbb{R}^{n_i}, v_i \in \mathbb{R}, i = 1, \dots, m$. Os dados do problema são $f \in \mathbb{R}^n, A_i \in \mathbb{R}^{n_i \times n}$, $b_i \in \mathbb{R}^{n_i}, c_i \in \mathbb{R}$ e $d_i \in \mathbb{R}, i = 1, \dots, m$. Apresente o dual das seguintes formas:

1. Introduza novas variáveis $y_i \in \mathbb{R}^{n_i}$ e $t_i \in \mathbb{R}$ e igualdades

$$\begin{array}{ll} y_i = A_i x + b_i, & i = 1, \dots, m \\ t_i = c_i^\top x + d_i, & i = 1, \dots, m \end{array}$$

   e apresente o dual Lagrangiano.

2. Comece com a formulação cônica do SOCP e use o cone dual. Utilize o fato de que o cone de segunda ordem é o próprio dual.

Introduza novas variáveis e igualdades e apresente o dual Lagrangiano.

Introduzindo as novas variáveis, ficamos com

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \|y_i\|_2 \le t_i, \quad i = 1, \dots, m \\ & y_i = A_i x + b_i, \quad i = 1, \dots, m \\ & t_i = c_i^T x + d_i, \quad i = 1, \dots, m \end{array}$$

O Lagrangiano será

$$\begin{aligned} L(x, y, t, \lambda, \nu, \mu) &= c^T x + \sum_{i=1}^m \lambda_i \left(\|y_i\|_2 - t_i\right) + \sum_{i=1}^m \nu_i^T \left(y_i - A_i x - b_i\right) + \sum_{i=1}^m \mu_i \left(t_i - c_i^T x - d_i\right) \\ &= \left(c - \sum_{i=1}^m A_i^T \nu_i - \sum_{i=1}^m \mu_i c_i\right)^T x + \sum_{i=1}^m \left(\lambda_i \|y_i\|_2 + \nu_i^T y_i\right) + \sum_{i=1}^m \left(-\lambda_i + \mu_i\right) t_i \\ &\quad - \sum_{i=1}^n \left(b_i^T \nu_i + d_i \mu_i\right) \end{aligned}$$

Derivando com relação a x e igualando a zero, temos

$$\left(c - \sum_{i=1}^m A_i^T \nu_i - \sum_{i=1}^m \mu_i c_i\right)^T = 0$$
$$\sum_{i=1}^m (A_i^T \nu_i + \mu_i c_i) = c$$

Minimizando com respeito a y_i, assumimos primeiramente que $\|\nu_i\|_2 \le \lambda_i$. Da desigualdade de Cauchy-Schwarz, temos

$$-\nu_i^T y_i \le \|\nu_i\|_2 \|y_i\|_2 \le \lambda_i \|y_i\|_2$$

O que implica que

$$\lambda_i \|y_i\|_2 + \nu_i^T y_i \geq 0$$

Segue que,

$$\inf_{y_i} \lambda_i \|y_i\|_2 + \nu_i^T y_i = 0$$

Agora, podemos assumir que $\|\nu_i\|_2 > \lambda_i$. Note que λ_i é restrito a ser não-negativo no problema dual. Se tomarmos $y_i = -s\nu_i$ para algum $s > 0$, então

$$\begin{aligned} \lambda_i \|y_i\|_2 + \nu_i^T y_i &= \lambda_i s \|\nu_i\|_2 - s \|\nu_i\|_2^2 \\ &= s \|\nu_i\|_2 \underbrace{(\lambda_i - \|\nu_i\|_2)}_{<0} \end{aligned}$$

Que pode ser tão negativo quanto se queira tomando s grande. Logo, neste caso temos

$$\inf_{y_i} \lambda_i \|y_i\|_2 + \nu_i^T y_i = -\infty$$

Dessa forma,

$$\inf_{y_i} \left(\lambda_i \|y_i\|_2 + \nu_i^T y_i\right) = \begin{cases} 0 & \|\nu_i\|_2 \leq \lambda_i \\ -\infty & \text{caso contrário} \end{cases}$$

O mínimo com relação a t_i será

$$-\lambda_i + \mu_i = 0 \Rightarrow \mu_i = \lambda_i$$

Com isso, a função Dual Lagrangiana será

$$g(\lambda, \nu, \mu) = \begin{cases} -\sum_{i=1}^{n} \left(b_i^T \nu_i + d_i \mu_i\right) & \sum_{i=1}^{m} \left(A_i^T \nu_i + \mu_i c_i\right) = c, \\ & \|\nu_i\|_2 \leq \lambda_i, \quad \mu = \lambda \\ -\infty & \text{caso contrário} \end{cases}$$

$\boxed{\text{Comece com a formulação cônica do SOCP e use o cone dual}}$

Não entendi. :(

Exercício 5.39

Relaxações de Programação Semidefinida (PSD) do problema de partição bidirecional. Considere o problema de partição bidirecional, apresentado na página 219 do livro do Boyd,

$$\begin{array}{ll} \min & x^\top W x \\ \text{s.a.} & x_i^2 = 1, \quad i = 1, \dots, n \end{array}$$

com variável $x \in \mathbb{R}^n$. O dual lagrangiano desse problema (não-convexo) é dado pelo PSD

$$\begin{array}{ll} \max & -\mathbf{1}^\top \nu, \\ \text{s.a.} & W + \operatorname{diag}(\nu) \succeq 0, \end{array}$$

com variável $\nu \in \mathbb{R}^n$. O valor ótimo desse PSD apresenta um limite inferior do valor ótimo do problema de partição. Nesse exercício vamos derivar outro PSD que apresenta um limite inferior no valor ótimo do problema de partição bidirecional e explorar a conexão entre esses dois PSDs.

1. Problema de partição bidirecional na forma matricial. Mostre que a partição bidirecional pode ser escrita como

$$\begin{array}{ll} \min & \operatorname{tr}(WX) \\ \text{s.a.} & X \succeq 0 \\ & \text{posto } X = 1 \\ & X_{ii} = 1, \quad i = 1, \dots, n, \end{array}$$

   com $X \in \mathbb{S}^n$. Dica: Mostre que X é viável se e somente se tem a forma $X = xx^\top$, onde $x \in \mathbb{R}^n$ satisfaz $x_i \in \{-1, 1\}$.

2. Relaxação PSD do problema de partição bidirecional. Usando a formulação do item acima, podemos formular a relaxação

$$\begin{array}{ll} \min & \operatorname{tr}(WX), \\ \text{s.a.} & X \succeq 0, \\ & X_{ii} = 1, \quad i = 1, \dots, n, \end{array}$$

   com $X \in \mathbb{S}^n$. Esse problema é um PSD, e portanto pode ser resolvido eficientemente. Explique o motivo do seu valor ótimo retornar um limite inferior para o valor ótimo do problema de partição bidirecional. O que podemos falar se o ponto ótimo $X^\star$ tiver posto igual a um?

3. Agora temos dois PSDs que retornam um limite inferior para o valor ótimo do problema de partição bidirecional: A relaxação PSD dada na Equação 2, e o dual lagrangiano no problema de partição bidirecional dada na Equação 1. Qual é a relação entre os dois PSDs? O que podemos falar sobre os limites inferiores encontrados por eles? Dica: Relacione-os por dualidade.

Se $X = xx^T$ então naturalmente $(xx^T)_{ii} = x_i^2$. Dessa forma, a restrição $X_{ii} = 1$ equivale a exatamente $x_i^2 = 1$. Note também que

$$\mathbf{tr}(WX) = \mathbf{tr}(Wxx^T) = \mathbf{tr}(xWx^T) = xWx^T$$

Ideias:

- Como a solução ótima $x_i \in \{-1, 1\}$, então as entradas de X serão ou 1 ou -1, pode explicar o posto=1?

Não consegui fazer. :(

Exercício 5.28

Utilizando teoria de dualidade, prove que a solução ótima do LP

$$\min 47x_1 + 93x_2 + 17x_3 - 93x_4,$$
$$\text{s.a.} \begin{bmatrix} -1 & -6 & 1 & 3 \\ -1 & -2 & 7 & 1 \\ 0 & 3 & -10 & -1 \\ -6 & -11 & -2 & 12 \\ 1 & 6 & -1 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \preceq \begin{bmatrix} -3 \\ 5 \\ -8 \\ -7 \\ 4 \end{bmatrix},$$

é única, dada por $x^\star = (1, 1, 1, 1)$.

Substituindo o vetor x^* nas equações, conseguimos observar que x^* de fato é uma solução primal viável. Agora, associando variáveis duais $z = (z_1, z_2, z_3, z_4, z_5)$ à cada uma das restrições acima, temos o seguinte problema Dual:

$$\max 3z_1 - 5z_2 + 8z_3 + 7z_4 - 4z_5$$

$$\begin{aligned} z_1 + z_2 + 6z_4 - z_5 &\leq 47 \\ 6z_1 + 2z_2 - 3z_3 + 11z_4 - 6z_5 &\leq 93 \\ -z_1 - 7z_2 + 10z_3 + 2z_4 + z_5 &\leq 17 \\ -3z_1 - z_2 + z_3 - 12z_4 + 3z_5 &\leq -93 \end{aligned}$$

com $z_1, z_2, z_3, z_4, z_5 \geq 0$

Usando as restrições de complementariedade, tal que

$$z^*(Ax^* - b) = 0$$

Vemos que apenas a última equação não atende na igualdade quando aplicada x^*. Dessa forma, $z_5 = 0$. Com isso, usando novamente as condições de complementariedade, podemos montar o seguinte sistema:

$$\begin{aligned} z_1 + z_2 + 6z_4 - z_5 &= 47 \\ 6z_1 + 2z_2 - 3z_3 + 11z_4 - 6z_5 &= 93 \\ -z_1 - 7z_2 + 10z_3 + 2z_4 + z_5 &= 17 \\ -3z_1 - z_2 + z_3 - 12z_4 + 3z_5 &= -93 \end{aligned}$$

Fazendo $z_5 = 0$, temos

$$\begin{aligned} z_1 + z_2 + 6z_4 &= 47 \\ 6z_1 + 2z_2 - 3z_3 + 11z_4 &= 93 \\ -z_1 - 7z_2 + 10z_3 + 2z_4 &= 17 \\ -z_1 - z_2 + z_3 - 12z_4 &= -93 \end{aligned}$$

Resolvendo esse sistema linear com 4 variáveis e 4 restrições, chegamos a solução $z_1 = 3, z_2 = 2, z_3 = 2$ e $z4 = 7$. Dessa forma, a solução dual será $z = (3, 2, 2, 7, 0)$. Substituindo na função objetivo, temos

$$3z_1 - 5z_2 + 8z_3 + 7z_4 - 4z_5 = 64$$

Que é exatamente o valor da função objetivo do primal quando aplicado o ponto x^*. Portanto, a dualidade forte está sendo respeitada e o ponto $z^* = (3, 2, 2, 7, 0)$ é um certificado de otimalidade para x^*.

Exercício 5.21

Um problema convexo que a dualidade forte falha. Considere o problema de otimização

$$\begin{aligned} \min \quad & e^{-x} \\ \text{s.a.} \quad & \frac{x^2}{y} \leq 0 \end{aligned}$$

com variáveis x e y e o domínio $\mathcal{D} = \{(x, y) \mid y > 0\}$.

1. Verifique que é um problema de otimização convexa. Encontre o valor ótimo.
2. Informe o problema dual lagrangiano, e encontre a solução ótima $\lambda^\star$ e o valor ótimo $d^\star$ do problema dual. Qual é o gap de dualidade ótimo?
3. A condição de Slater é satisfeita para este problema? Explique sua resposta.
4. Qual é o valor ótimo $p^\star(u)$ do problema perturbado

$$\begin{aligned} \min \quad & e^{-x} \\ \text{s.a.} \quad & \frac{x^2}{y} \leq u \end{aligned}$$

como função de u ? Verifique que a inequação global de sensibilidade

$$p^\star(u) \geq p^\star(0) - \lambda^\star u$$

não é satisfeita.

Verifique que é um problema de otimização convexa. Encontre o valor ótimo.

A função e^{-x} é uma função convexa, uma vez que $(e^{-x})'' = (-e^{-x})' = e^{-x} > 0$ para todo $x \in \mathbb{R}$.
Como o domínio D são de pontos em que a coordenada y é estritamente maior que zero, uma solução viável para esse problema será apenas quando $x = 0$, uma vez que $\frac{x^2}{y} \geq 0$. Dessa forma, independente de y, a solução ótima será $p^* = e^0 = 1$.

Informe o problema dual lagrangiano, e encontre a solução ótima $\lambda^\star$ e o valor ótimo $d^\star$ do problema dual.

Calculando o Lagragiano, temos

$$L(\lambda, (x, y)) = e^{-x} + \lambda \left(\frac{x^2}{y}\right)$$

Como é uma soma de termos positivos, assumindo que $\lambda \geq 0$, o menor valor será quando essa soma atingir 0. Dessa forma, $d^* = 0$. O GAP de dualidade é exatamente $p^* - d^* = 1$.

A condição de Slater é satisfeita para este problema? Explique sua resposta.

Se $(x, y) \in D$ é viável, então $x^2 \leq 0$, e dessa forma, $x = 0$. Em outras palavras, a região viável é exatamente $\{(0, y) | y > 0\}$, que não contém um ponto estritamente viável. Dessa forma, as condições de Slater não são satisfeitas.

Qual é o valor ótimo $p^\star(u)$ do problema perturbado?

Se $u = 0$, voltamos ao problema anterior e $p^*(u) = 1$. Se $u > 0$, então $p^*(u) = 0$, uma vez que quando $x \to \infty$, a função objetivo vai a zero. Se $u \leq 0$ o problema é inviável. Verificando a inequação global,

$$p^*(u) \geq 1 - \lambda u$$

Vimos que $p^*(u)$ não assume valores negativos. Dessa forma, para a equação ser satisfeita, $\lambda u = 1$, o que é um absurdo.

Exercício 10

Seja o QP abaixo

$$\begin{aligned} p^\star := \min \quad & x^\top P x, \\ \text{s.a.} \quad & Ax \preceq b, \end{aligned}$$

onde $P \in \mathbb{S}^n_{++}$. 1. Mostre que o dual lagrangiano do problema (3) é dado por

$$\begin{aligned} d^\star := \max & -\frac{1}{4}\lambda^\top A P^{-1} A^\top \lambda - b^\top \lambda \\ \text{s.a.} \quad & \lambda \succeq 0 \end{aligned}$$

2. Mostre que se o problema (3) é inviável, então $p^\star = d^\star = +\infty$.

$\boxed{\text{Mostre que o dual lagrangiano do problema}}$

O Lagrangiano pode ser escrito como

$$\begin{aligned} L(x, \lambda) &= x^T P x + \lambda(A^T x - b) \\ &= x^T P x + x\lambda^T A - b^T \lambda \end{aligned}$$

Derivando com respeito a x e igualando a zero, temos

$$2x^T P + \lambda^T A = 0$$

Como P é simétrica e positiva definida, temos

$$\begin{aligned} 2x^T P &= -\lambda^T A \\ 2x^T &= -\lambda^T A P^{-1} \\ x^T &= -\frac{1}{2}\lambda^T A P^{-1} \end{aligned}$$

Substituindo em L, temos

$$\begin{aligned} g(\lambda) &= \frac{1}{2}\lambda^T A P^{-1} P(-\frac{1}{2}\lambda^T A P^{-1})^T + (-\frac{1}{2}\lambda^T A P^{-1})^T \lambda^T A - b^T \lambda \\ &= -\frac{1}{4}\lambda^\top A P^{-1} A^\top \lambda - b^\top \lambda \end{aligned}$$

Dessa forma, o problema dual Lagrangiano será

$$\begin{aligned} \max & -\frac{1}{4}\lambda^\top A P^{-1} A^\top \lambda - b^\top \lambda \\ \text{s.a.} \quad & \lambda \succeq 0 \end{aligned}$$

$\boxed{\text{Mostre que se o problema (3) é inviável, então } p^\star = d^\star = +\infty.}$

Para o problema (3) ser inviável, teremos uma solução x tal que $b - Ax \preceq 0$.
O Lagrangiano pode ser escrito como

$$\begin{aligned} L(x, \lambda) &= x^T P x + \lambda(b - A^T x) \\ &= x^T P x + b^T \lambda - x\lambda^T A \end{aligned}$$

O problema dual será

$$
\begin{aligned}
d^* := \max \ & \frac{1}{4}\lambda^\top A P^{-1} A^\top \lambda + b^\top \lambda \\
\text{s.a.} \quad & \lambda \succeq 0
\end{aligned}
$$

Claramente como só há uma restrição de não-negatividade, o máximo desse problema será ∞. Pela dualidade forte, $p^* = d^*$.

Exercício 5.19

Soma dos maiores elementos de um vetor. Defina $f : \mathbb{R}^n \to \mathbb{R}$ como

$$f(x) = \sum_{i=1}^{r} x_{[i]}$$

sendo r um inteiro entre 1 e n e $x_{[1]} \geq x_{[2]} \geq \ldots \geq x_{[r]}$ são componentes de x ordenados de forma não-crescente. Em outras palavras, $f(x)$ é a soma dos r maiores elementos de x. Nesse problema, podemos utilizar a restrição

$$f(x) \leq \alpha$$

No capítulo 3, página 80 do livro do Boyd, essa é uma restrição convexa e é equivalente a um conjunto de $n!/(r!(n-r)!)$ inequações lineares

$$x_{i_1} + \cdots + x_{i_r} \leq \alpha, \quad 1 \leq i_1 < i_2 < \cdots < i_r \leq n.$$

O objetivo é apresentar uma representação mais compacta desse problema.

1. Dado um vetor $x \in \mathbb{R}^n$, mostre que $f(x)$ é igual ao valor ótimo do PL

$$\begin{array}{ll} \max & x^\top y \\ \text{s.a.} & 0 \preceq y \preceq \mathbf{1}, \\ & \mathbf{1}^\top y = r, \end{array}$$

   com $y \in \mathbb{R}^n$ como variável.

2. Derive o dual como PL do item acima. Mostre que pode ser escrito como

$$\begin{array}{ll} \min & rt + \mathbf{1}^\top u \\ \text{s.a.} & t\mathbf{1}^\top + u \succeq x \\ & u \succeq 0 \end{array}$$

   sendo $t \in \mathbb{R}$ e $u \in \mathbb{R}^n$ variáveis do problema. Por dualidade, esse LP tem o mesmo valor ótimo que o LP do item acima, em outras palavras, $f(x)$. Portanto, temos o seguinte resultado: x satisfaz $f(x) \leq \alpha$ se e somente se existe $t \in \mathbb{R}, u \in \mathbb{R}^n$ tal que

$$rt + \mathbf{1}^\top u \leq \alpha, \quad t\mathbf{1}^\top + u \succeq x, \quad u \succeq 0$$

   Essas condições formam um conjunto de $2n + 1$ inequações em $2n + 1$ variáveis x, u, t.

Dado um vetor $x \in \mathbb{R}^n$, mostre que $f(x)$ é igual ao valor ótimo do PL

Vamos assumir que os elementos do vetor x estão ordenados de forma decrescente

$$x_1 \geq x_2 \geq \cdots \geq x_n$$

Uma solução viável para o problema acontece quando tomamos r elementos do vetor y sendo igualados a 1 e o restante igual a 0.

Para obtermos o **maior** valor na função objetivo devemos escolher, em especial, os r elementos de y que correspondem aos r maiores valores de x. Em outras palavras,

$$y_1 = y_2 = \cdots = y_r = 1$$

e

$$y_{r+1} = \cdots = y_n = 0$$

E com isso, a função objetivo nos retornará a seguinte solução ótima

$$x_1 + x_2 + \cdots + x_r$$

Que é de fato a soma dos r maiores elementos de x.

$\boxed{\text{Derive o dual como PL do item acima}}$

Associando multiplicadores λ, u, t as restrições $1^T y = r$, $y \geq 0$ e $y \leq 1$, temos o seguinte Lagrangiano

$$\begin{aligned} L(y, \lambda, u, t) &= x^T y + \lambda^T y - u^T(y - 1) - t(1^T y - r) \\ &= 1^T u + rt + (x + \lambda - u - t1)^T y \end{aligned}$$

Estamos com uma função do tipo $c + ty$, que é uma reta. Desse modo, só atingirá o máximo quando $x + \lambda - u - t1 = 0$, o que nos dará uma função constante $1^T u + rt$. Assim,

$$g(\lambda, u, t) = \begin{cases} 1^T u + rt & x + \lambda - u - t1 = 0 \\ \infty & \text{caso contrário} \end{cases}$$

Dessa forma, o problema dual pode ser considerado como

$$\begin{array}{ll} \min & 1^T u + rt \\ \text{sujeito a} & -\lambda + u + t1 = x \\ & \lambda \succeq 0, \quad u \succeq 0 \end{array}$$

Note que λ atua como uma variável de folga no problema, dessa forma, podemos reescrever como

$$\begin{array}{ll} \min & 1^T u + rt \\ \text{sujeito a} & u + t1 \succeq x \\ & \quad u \succeq 0 \end{array}$$