Universidade Federal de São Carlos
Departamento de Engenharia de Produção

Otimização Linear Contínua e Discreta (Tópicos Avançados em PCSP)

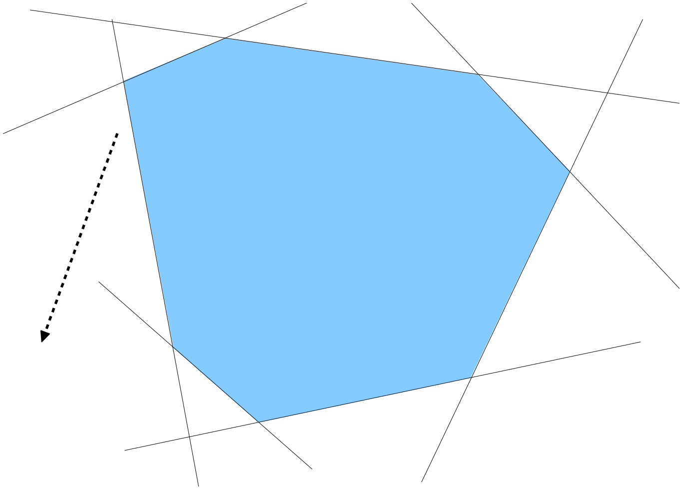
PPGEP, UFSCar - Semestre 01/2022
Prof. Dr. Pedro Munari (munari@dep.ufscar.br)

Tópico 6.2: Métodos de pontos interiores: KKT perturbado, barreira logarítmica e trajetória central

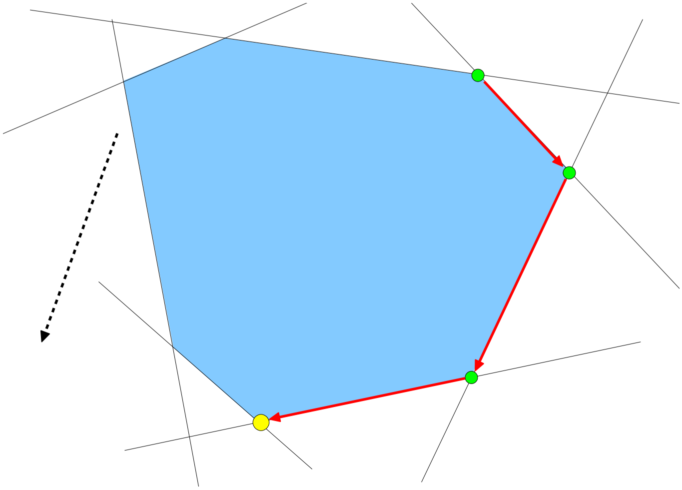
Objetivos deste tópico

- Estudar os conceitos de barreira logarítmica, trajetória central e vizinhança em métodos de pontos interiores;
- Vamos ver como esses conceitos se relacionam com o KKT perturbado e como dão origem ao algoritmo primal-dual.

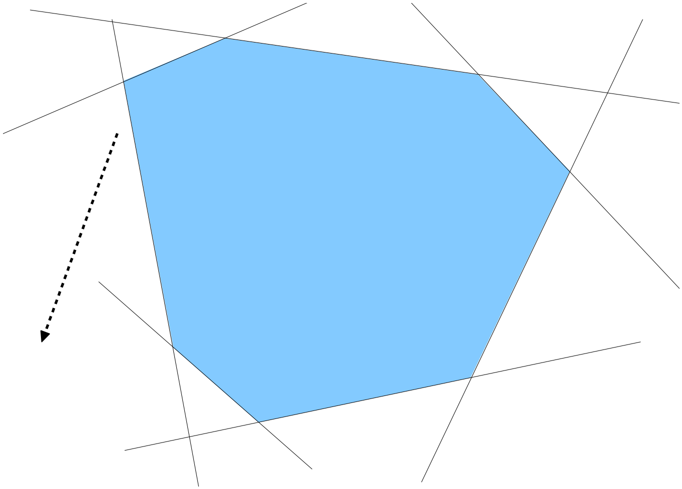
Métodos tipo simplex



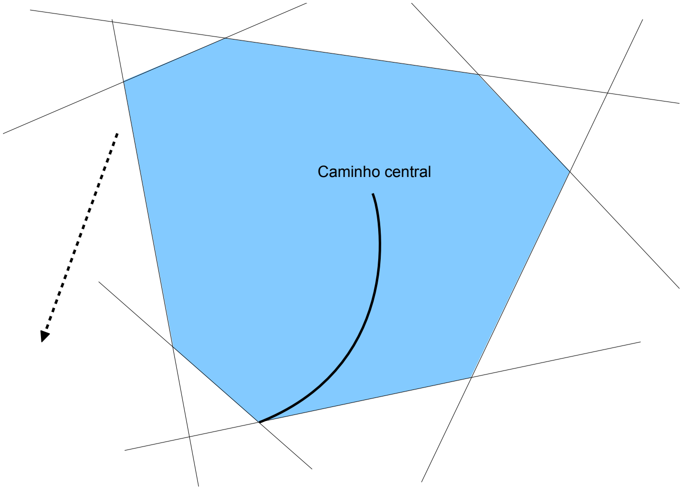
Métodos tipo simplex



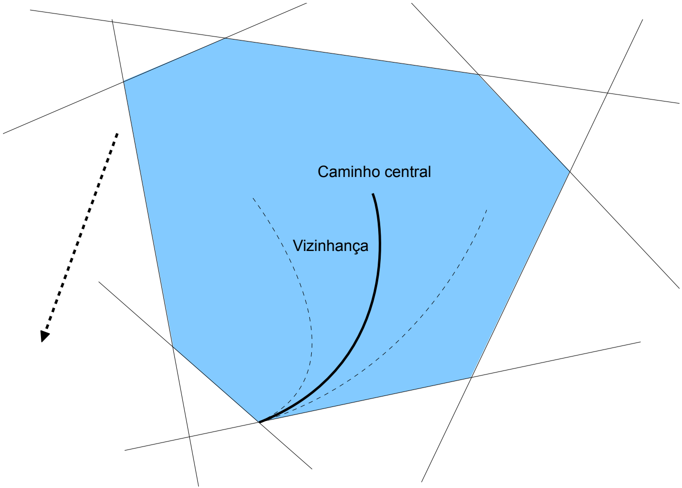
Métodos de pontos interiores



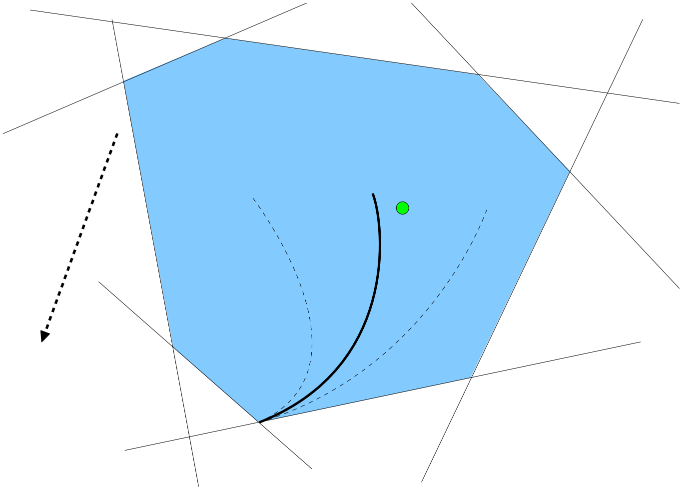
Métodos de pontos interiores



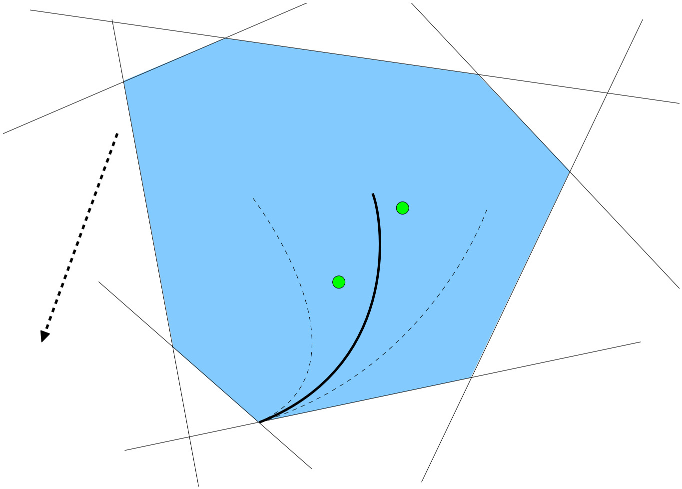
Métodos de pontos interiores



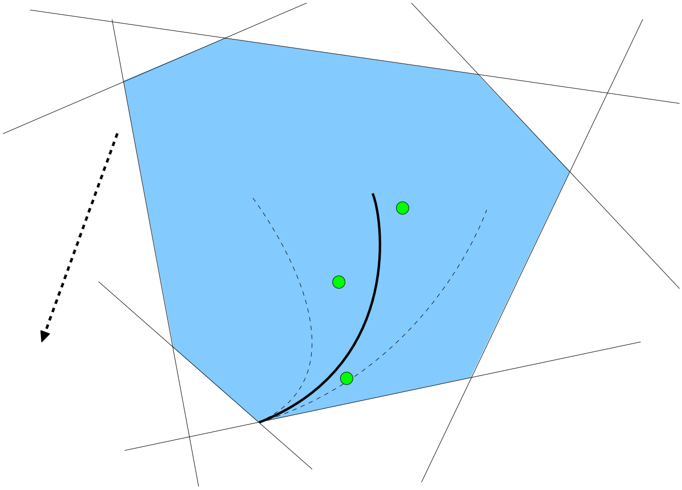
Métodos de pontos interiores



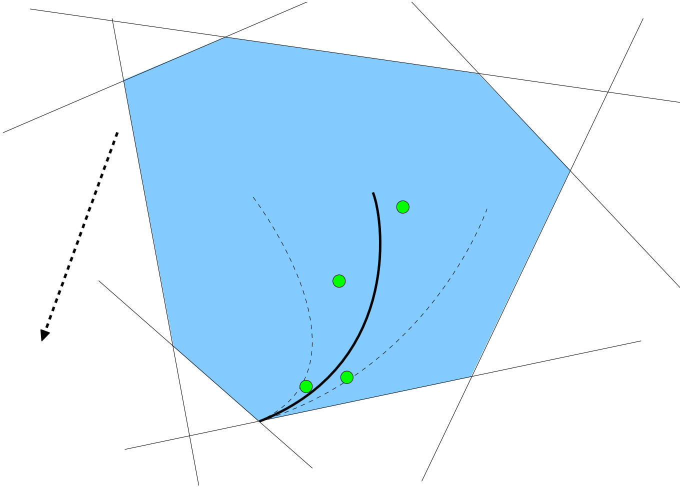
Métodos de pontos interiores



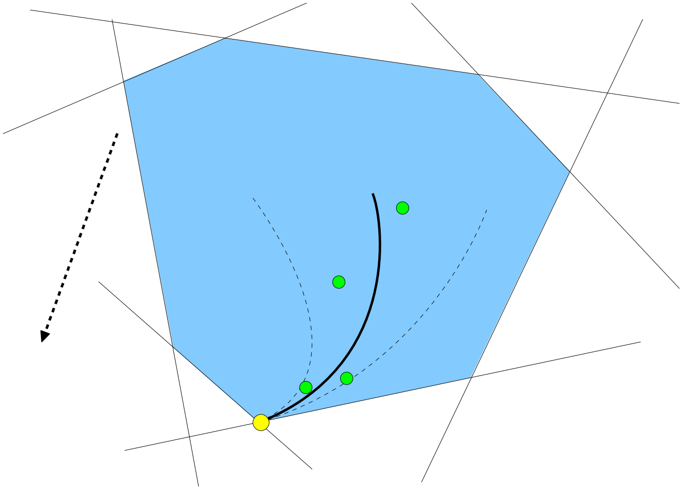
Métodos de pontos interiores



Métodos de pontos interiores



Métodos de pontos interiores



Métodos de pontos interiores

Considere o seguinte problema na forma padrão:

$$\begin{aligned} \text{min} \quad & c^T x \\ \text{s.a} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

Métodos de pontos interiores

Considere o seguinte problema na forma padrão:

$$\begin{aligned} \text{min} \quad & c^T x \\ \text{s.a} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

E se em vez de achar o ótimo, quiséssemos achar o centro da região factível?

Métodos de pontos interiores

Considere o seguinte problema na forma padrão:

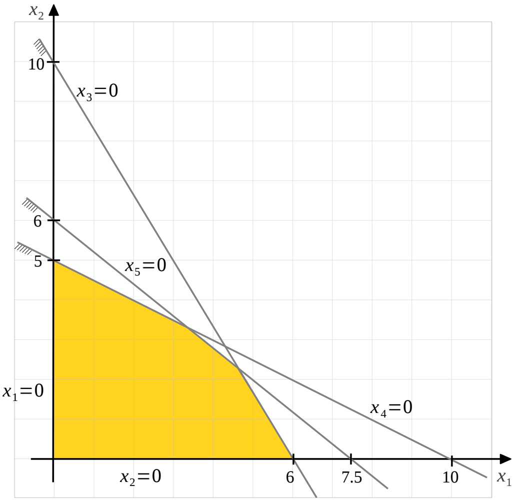
$$\begin{aligned} \text{min} \quad & c^T x \\ \text{s.a} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

E se em vez de achar o ótimo, quiséssemos achar o centro da região factível?

$$\begin{aligned} \text{max} \quad & \sum_{j=1}^{n} \ln x_j \\ \text{s.a} \quad & Ax = b \end{aligned}$$

Métodos de pontos interiores

$$\begin{aligned} 0{,}5x_1 + 0{,}3x_2 + x_3 &= 3 \\ 0{,}1x_1 + 0{,}2x_2 + x_4 &= 1 \\ 0{,}4x_1 + 0{,}5x_2 + x_5 &= 3 \\ x_1, x_2, x_3, x_4, x_5 &\geq 0 \end{aligned}$$



Métodos de pontos interiores

Considere o seguinte problema na forma padrão:

$$\begin{aligned} \text{min} \quad & c^T x \\ \text{s.a} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

E se em vez de achar o ótimo, quiséssemos achar o centro da região factível?

$$\begin{aligned} \text{min} \quad & -\sum_{j=1}^{n} \ln x_j \\ \text{s.a} \quad & Ax = b \end{aligned}$$

Métodos de pontos interiores

Podemos balancear centralidade e otimalidade:

Métodos de pontos interiores

Podemos balancear centralidade e otimalidade:

$$\begin{aligned} \text{min} \quad & c^T x - \mu \sum_{j=1}^{n} \ln x_j \\ \text{s.a} \quad & Ax = b \end{aligned}$$

Métodos de pontos interiores

Podemos balancear centralidade e otimalidade:

$$\begin{array}{ll} \text{min} & c^T x - \mu \sum_{j=1}^{n} \ln x_j \\ \text{s.a} & Ax = b \end{array}$$

- $\ln x_j$ recebe o nome de *barreira logarítmica*;

Métodos de pontos interiores

Podemos balancear centralidade e otimalidade:

$$\begin{array}{ll} \text{min} & c^T x - \mu \sum_{j=1}^{n} \ln x_j \\ \text{s.a} & Ax = b \end{array}$$

- $\ln x_j$ recebe o nome de *barreira logarítmica*;
- μ é o *parâmetro de barreira*

Métodos de pontos interiores

Podemos balancear centralidade e otimalidade:

$$\begin{array}{ll} \text{min} & c^T x - \mu \sum_{j=1}^{n} \ln x_j \\ \text{s.a} & Ax = b \end{array}$$

- $\ln x_j$ recebe o nome de *barreira logarítmica*;
- μ é o *parâmetro de barreira* e determina qual o peso da centralidade em relação à otimalidade;

Métodos de pontos interiores

Podemos balancear centralidade e otimalidade:

$$\begin{array}{ll} \text{min} & c^T x - \mu \sum_{j=1}^{n} \ln x_j \\ \text{s.a} & Ax = b \end{array}$$

- $\ln x_j$ recebe o nome de *barreira logarítmica*;
- μ é o *parâmetro de barreira* e determina qual o peso da centralidade em relação à otimalidade;
- Se $\mu = 0$, buscamos por otimalidade apenas;

Métodos de pontos interiores

Podemos balancear centralidade e otimalidade:

$$\begin{array}{ll} \text{min} & c^T x - \mu \sum_{j=1}^{n} \ln x_j \\ \text{s.a} & Ax = b \end{array}$$

- $\ln x_j$ recebe o nome de *barreira logarítmica*;
- μ é o *parâmetro de barreira* e determina qual o peso da centralidade em relação à otimalidade;
- Se $\mu = 0$, buscamos por otimalidade apenas;
- Se $\mu \to \infty$, buscamos por centralidade apenas;

Métodos de pontos interiores

Podemos balancear centralidade e otimalidade:

$$\begin{array}{ll} \text{min} & c^T x - \mu \sum_{j=1}^{n} \ln x_j \\ \text{s.a} & Ax = b \end{array}$$

- $\ln x_j$ recebe o nome de *barreira logarítmica*;
- μ é o *parâmetro de barreira* e determina qual o peso da centralidade em relação à otimalidade;
- Se $\mu = 0$, buscamos por otimalidade apenas;
- Se $\mu \to \infty$, buscamos por centralidade apenas;
- E para valores intermediários de μ?

Métodos de pontos interiores

Problema primal com barreira logarítmica:

Métodos de pontos interiores

Problema primal com barreira logarítmica:

$$\begin{aligned} \text{min} \quad & c^T x - \mu \sum_{j=1}^{n} \ln x_j \\ \text{s.a} \quad & Ax = b \end{aligned}$$

Métodos de pontos interiores

Problema primal com barreira logarítmica:

$$\begin{aligned} \text{min} \quad & c^T x - \mu \sum_{j=1}^{n} \ln x_j \\ \text{s.a} \quad & Ax = b \end{aligned}$$

- Como ficam as condições de KKT para esse problema?

Condições KKT

Considere o problema de Otimização (geral)

$$\begin{aligned} \text{min} \quad & f(x) \\ \text{s.a} \quad & h(x) = 0 \\ & g(x) \leq 0 \end{aligned}$$

com $f : \mathbb{R}^n \to \mathbb{R}$, $h : \mathbb{R}^n \to \mathbb{R}^l$ e $g : \mathbb{R}^n \to \mathbb{R}^m$.

Definimos a função Lagrangiana do problema:

$$\mathcal{L}(x, p, s) = f(x) + p^T h(x) + s^T g(x)$$

com $s^T \geq 0$.

Condições KKT

Teorema: Condições necessárias de primeira ordem (condições KKT).

- Suponha que x seja um ótimo local do problema de otimização geral, que as funções f, g e h sejam continuamente diferenciáveis e que certas condições de qualificação sejam válidas. Então existe um vetor de multiplicadores de Lagrange (p, s) tal que as seguintes condições são satisfeitas:

$$
\begin{aligned}
\nabla_x \mathcal{L}(x, p, s) &= 0 \\
h_k(x) &= 0, \; k = 1, \ldots, l \\
g_j(x) &\leq 0, \; j = 1, \ldots, m \\
s_j g_j(x) &= 0, \; j = 1, \ldots, m \\
s_j &\geq 0, \; j = 1, \ldots, m
\end{aligned}
$$

Condições KKT

- Para o problema primal com barreira logarítmica:

Condições KKT

- Para o problema primal com barreira logarítmica:

  $f(x) =$

Condições KKT

- Para o problema primal com barreira logarítmica:

$$f(x) = c^T x - \mu \sum_{j=1}^{n} \ln x_j;$$

Condições KKT

- Para o problema primal com barreira logarítmica:

$$f(x) = c^T x - \mu \sum_{j=1}^{n} \ln x_j;$$

$$h(x) =$$

Condições KKT

- Para o problema primal com barreira logarítmica:

$$f(x) = c^T x - \mu \sum_{j=1}^{n} \ln x_j;$$

$$h(x) = b - Ax;$$

Condições KKT

- Para o problema primal com barreira logarítmica:

$$f(x) = c^T x - \mu \sum_{j=1}^{n} \ln x_j;$$

$$h(x) = b - Ax;$$

$$\mathcal{L}(x, p, s) =$$

Condições KKT

- Para o problema primal com barreira logarítmica:

$$f(x) = c^T x - \mu \sum_{j=1}^{n} \ln x_j;$$

$$h(x) = b - Ax;$$

$$\mathcal{L}(x, p, s) = c^T x - \mu \sum_{j=1}^{n} \ln x_j + p^T (b - Ax),\ x_j > 0$$

Condições KKT

- Para o problema primal com barreira logarítmica:

$$f(x) = c^T x - \mu \sum_{j=1}^{n} \ln x_j;$$

$$h(x) = b - Ax;$$

$$\mathcal{L}(x,p,s) = c^T x - \mu \sum_{j=1}^{n} \ln x_j + p^T (b - Ax),\ x_j > 0$$

$$\nabla_{x_j} \mathcal{L}(x,p,s) =$$

Condições KKT

- Para o problema primal com barreira logarítmica:

$$f(x) = c^T x - \mu \sum_{j=1}^{n} \ln x_j;$$

$$h(x) = b - Ax;$$

$$\mathcal{L}(x, p, s) = c^T x - \mu \sum_{j=1}^{n} \ln x_j + p^T(b - Ax),\ x_j > 0$$

$$\nabla_{x_j} \mathcal{L}(x, p, s) = c_j - a_j^T p - \mu \frac{1}{x_j},\ x_j > 0$$

Condições KKT

- Para o problema primal com barreira logarítmica:

$$f(x) = c^T x - \mu \sum_{j=1}^{n} \ln x_j;$$

$$h(x) = b - Ax;$$

$$\mathcal{L}(x,p,s) = c^T x - \mu \sum_{j=1}^{n} \ln x_j + p^T(b - Ax),\ x_j > 0$$

$$\nabla_{x_j} \mathcal{L}(x,p,s) = c_j - a_j^T p - \mu \frac{1}{x_j},\ x_j > 0$$

- Fazendo $s_j = \mu / x_j$, obtemos:

Condições KKT

- Para o problema primal com barreira logarítmica:

$$f(x) = c^T x - \mu \sum_{j=1}^{n} \ln x_j;$$

$$h(x) = b - Ax;$$

$$\mathcal{L}(x,p,s) = c^T x - \mu \sum_{j=1}^{n} \ln x_j + p^T(b - Ax),\ x_j > 0$$

$$\nabla_{x_j} \mathcal{L}(x,p,s) = c_j - a_j^T p - \mu \frac{1}{x_j},\ x_j > 0$$

- Fazendo $s_j = \mu / x_j$, obtemos:

$$\nabla_x \mathcal{L}(x,p,s) = c - A^T p - s$$

Condições KKT

Teorema: Condições necessárias de primeira ordem (condições KKT).

- Suponha que x seja um ótimo local do problema de otimização geral, que as funções f, g e h sejam continuamente diferenciáveis e que certas condições de qualificação sejam válidas. Então existe um vetor de multiplicadores de Lagrange (p, s) tal que as seguintes condições são satisfeitas:

$$
\begin{aligned}
\nabla_x \mathcal{L}(x, p, s) &= 0 \\
h_k(x) &= 0,\ k = 1, \ldots, l \\
g_j(x) &\leq 0,\ j = 1, \ldots, m \\
s_j g_j(x) &= 0,\ j = 1, \ldots, m \\
s_j &\geq 0,\ j = 1, \ldots, m
\end{aligned}
$$

Condições KKT

Logo, obtemos as condições KKT perturbadas:

$$\begin{aligned} Ax &= b \\ A^T p + s &= c \\ x_j s_j &= \mu, \ j = 1, \ldots, n \\ x, s &> 0 \end{aligned}$$

Condições KKT

Logo, obtemos as condições KKT perturbadas:

$$\begin{aligned} Ax &= b \\ A^T p + s &= c \\ XSe &= \mu e \\ x, s &> 0 \end{aligned}$$

Condições KKT

Logo, obtemos as condições KKT perturbadas:

$$\begin{aligned} Ax &= b \\ A^T p + s &= c \\ XSe &= \mu e \\ x, s &> 0 \end{aligned}$$

com $X = diag(x_1, x_2, \ldots, x_n)$,

Condições KKT

Logo, obtemos as condições KKT perturbadas:

$$\begin{array}{rcl} Ax & = & b \\ A^T p + s & = & c \\ XSe & = & \mu e \\ x, s & > & 0 \end{array}$$

com $X = diag(x_1, x_2, \ldots, x_n)$, $S = diag(s_1, s_2, \ldots, s_n)$,

Condições KKT

Logo, obtemos as condições KKT perturbadas:

$$\begin{aligned} Ax &= b \\ A^T p + s &= c \\ XSe &= \mu e \\ x, s &> 0 \end{aligned}$$

com $X = diag(x_1, x_2, \ldots, x_n)$, $S = diag(s_1, s_2, \ldots, s_n)$, $e = (1, 1, \ldots, 1)^T$.

Métodos de pontos interiores

Problema dual com barreira logarítmica:

Métodos de pontos interiores

Problema dual com barreira logarítmica:

$$\begin{array}{ll} \text{max} & b^T p + \mu \sum_{j=1}^{n} \ln s_j \\ \text{s.a} & A^T p + s = c \end{array}$$

Métodos de pontos interiores

Problema dual com barreira logarítmica:

$$\begin{array}{ll} \text{max} & b^T p + \mu \sum_{j=1}^{n} \ln s_j \\ \text{s.a} & A^T p + s = c \end{array}$$

- Como ficam as condições de KKT para esse problema?

Condições KKT

- Para o problema de barreira logarítmica dual:

Condições KKT

- Para o problema de barreira logarítmica dual:

  $f(p, s) =$

Condições KKT

- Para o problema de barreira logarítmica dual:

  $$f(p, s) = b^T p + \mu \sum_{j=1}^{n} \ln s_j;$$

Condições KKT

- Para o problema de barreira logarítmica dual:

$$f(p,s) = b^T p + \mu \sum_{j=1}^{n} \ln s_j;$$

$$h(p,s) =$$

Condições KKT

- Para o problema de barreira logarítmica dual:

  $f(p,s) = b^T p + \mu \sum_{j=1}^{n} \ln s_j;$

  $h(p,s) = c - A^T p - s;$

Condições KKT

- Para o problema de barreira logarítmica dual:

$$f(p, s) = b^T p + \mu \sum_{j=1}^{n} \ln s_j;$$

$$h(p, s) = c - A^T p - s;$$

$$\mathcal{L}(p, s, x) =$$

Condições KKT

- Para o problema de barreira logarítmica dual:

$$f(p,s) = b^T p + \mu \sum_{j=1}^{n} \ln s_j;$$

$$h(p,s) = c - A^T p - s;$$

$$\mathcal{L}(p,s,x) = b^T p - \mu \sum_{j=1}^{n} \ln s_j + x^T(c - A^T p - s),\ s_j > 0$$

Condições KKT

- Para o problema de barreira logarítmica dual:

$$f(p,s) = b^T p + \mu \sum_{j=1}^{n} \ln s_j;$$

$$h(p,s) = c - A^T p - s;$$

$$\mathcal{L}(p,s,x) = b^T p - \mu \sum_{j=1}^{n} \ln s_j + x^T(c - A^T p - s),\ s_j > 0$$

$$\nabla_p \mathcal{L}(p,s,x) =$$

Condições KKT

- Para o problema de barreira logarítmica dual:

$$f(p,s) = b^T p + \mu \sum_{j=1}^{n} \ln s_j;$$

$$h(p,s) = c - A^T p - s;$$

$$\mathcal{L}(p,s,x) = b^T p - \mu \sum_{j=1}^{n} \ln s_j + x^T(c - A^T p - s), \; s_j > 0$$

$$\nabla_p \mathcal{L}(p,s,x) = b - Ax$$

Condições KKT

- Para o problema de barreira logarítmica dual:

$$f(p,s) = b^T p + \mu \sum_{j=1}^{n} \ln s_j;$$

$$h(p,s) = c - A^T p - s;$$

$$\mathcal{L}(p,s,x) = b^T p - \mu \sum_{j=1}^{n} \ln s_j + x^T(c - A^T p - s), \; s_j > 0$$

$$\nabla_p \mathcal{L}(p,s,x) = b - Ax$$

$$\nabla_{s_j} \mathcal{L}(p,s,x) =$$

Condições KKT

- Para o problema de barreira logarítmica dual:

$$f(p,s) = b^T p + \mu \sum_{j=1}^{n} \ln s_j;$$

$$h(p,s) = c - A^T p - s;$$

$$\mathcal{L}(p,s,x) = b^T p - \mu \sum_{j=1}^{n} \ln s_j + x^T(c - A^T p - s),\ s_j > 0$$

$$\nabla_p \mathcal{L}(p,s,x) = b - Ax$$

$$\nabla_{s_j} \mathcal{L}(p,s,x) = \mu \frac{1}{s_j} - x_j,\ s_j > 0$$

Métodos de pontos interiores

Logo, obtemos as mesmas condições KKT perturbadas:

$$\begin{aligned} Ax &= b \\ A^T p + s &= c \\ s_j x_j &= \mu, \; j = 1, \ldots, n \\ x, s &> 0 \end{aligned}$$

Métodos de pontos interiores

Logo, obtemos as mesmas condições KKT perturbadas:

$$\begin{aligned} Ax &= b \\ A^T p + s &= c \\ s_j x_j &= \mu,\ j = 1, \ldots, n \\ x, s &> 0 \end{aligned}$$

- Para cada valor de $\mu > 0$, temos uma solução única $(x(\mu), p(\mu), s(\mu))$ das condições KKT perturbadas;

Métodos de pontos interiores

Logo, obtemos as mesmas condições KKT perturbadas:

$$
\begin{aligned}
Ax &= b \\
A^T p + s &= c \\
s_j x_j &= \mu, \; j = 1, \dots, n \\
x, s &> 0
\end{aligned}
$$

- Para cada valor de $\mu > 0$, temos uma solução única $(x(\mu), p(\mu), s(\mu))$ das condições KKT perturbadas;
- $x(\mu)$ é chamado de μ-centro do problema primal;

Métodos de pontos interiores

Logo, obtemos as mesmas condições KKT perturbadas:

$$\begin{aligned} Ax &= b \\ A^T p + s &= c \\ s_j x_j &= \mu, \;\; j = 1, \ldots, n \\ x, s &> 0 \end{aligned}$$

- Para cada valor de $\mu > 0$, temos uma solução única $(x(\mu), p(\mu), s(\mu))$ das condições KKT perturbadas;
- $x(\mu)$ é chamado de μ-centro do problema primal;
- $(p(\mu), s(\mu))$ é chamado de μ-centro do problema dual;

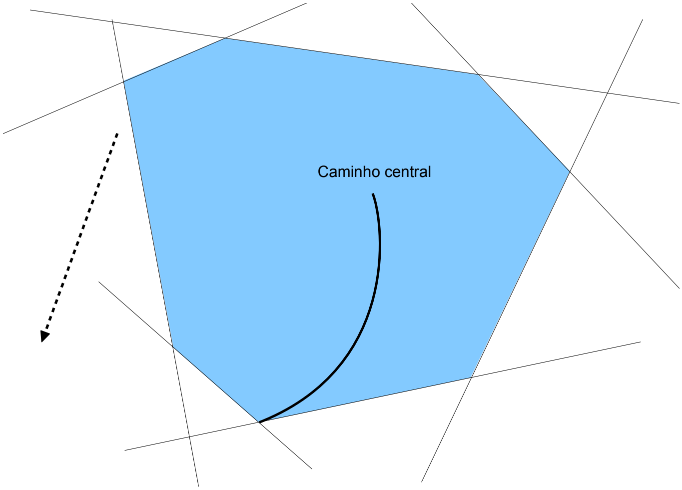
Métodos de pontos interiores

Logo, obtemos as mesmas condições KKT perturbadas:

$$\begin{aligned} Ax &= b \\ A^T p + s &= c \\ s_j x_j &= \mu, \ j = 1, \ldots, n \\ x, s &> 0 \end{aligned}$$

- Para cada valor de $\mu > 0$, temos uma solução única $(x(\mu), p(\mu), s(\mu))$ das condições KKT perturbadas;
- $x(\mu)$ é chamado de μ-centro do problema primal;
- $(p(\mu), s(\mu))$ é chamado de μ-centro do problema dual;
- O conjunto de todos os μ-centros definem a caminho/trajetória central (*central path*) em suas respectivas regiões factíveis.

Métodos de pontos interiores



Método primal-dual de pontos interiores

- O caminho central serve como um guia “seguro” para a otimalidade;

Método primal-dual de pontos interiores

- O caminho central serve como um guia "seguro" para a otimalidade;
- Ao seguir essa trajetória, as soluções primais e duais de cada iteração caminham em direção à solução ótima,

Método primal-dual de pontos interiores

- O caminho central serve como um guia “seguro” para a otimalidade;
- Ao seguir essa trajetória, as soluções primais e duais de cada iteração caminham em direção à solução ótima, *sem se aproximar prematuramente* da fronteira da região factível;

Método primal-dual de pontos interiores

- O caminho central serve como um guia “seguro” para a otimalidade;
- Ao seguir essa trajetória, as soluções primais e duais de cada iteração caminham em direção à solução ótima, *sem se aproximar prematuramente* da fronteira da região factível;
- Dessa forma, esses métodos são conhecidos como métodos seguidores de caminho (*path-following methods*);

Método primal-dual de pontos interiores

- O caminho central serve como um guia "seguro" para a otimalidade;
- Ao seguir essa trajetória, as soluções primais e duais de cada iteração caminham em direção à solução ótima, *sem se aproximar prematuramente* da fronteira da região factível;
- Dessa forma, esses métodos são conhecidos como métodos seguidores de caminho (*path-following methods*);
- Seguir estritamente o caminho central ($s_j x_j = \mu, \ j = 1, \ldots, n$) seria caro, em geral, pois teríamos que resolver o KKT perturbado a cada iteração;

Método primal-dual de pontos interiores

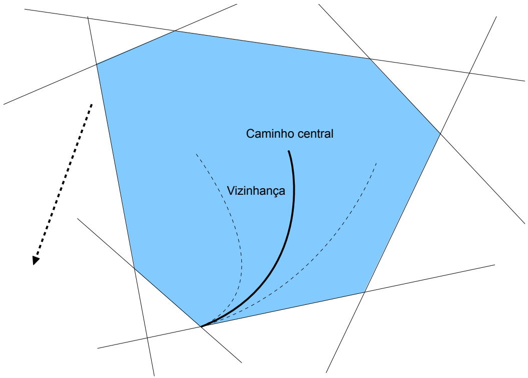
- O caminho central serve como um guia "seguro" para a otimalidade;
- Ao seguir essa trajetória, as soluções primais e duais de cada iteração caminham em direção à solução ótima, *sem se aproximar prematuramente* da fronteira da região factível;
- Dessa forma, esses métodos são conhecidos como métodos seguidores de caminho (*path-following methods*);
- Seguir estritamente o caminho central ($s_j x_j = \mu,\ j = 1, \ldots, n$) seria caro, em geral, pois teríamos que resolver o KKT perturbado a cada iteração;
- Ao invés, exigimos que os pares $s_j x_j$ sejam em média iguais a μ, isto é, $\mu = s^T x/n$;

Método primal-dual de pontos interiores

- O caminho central serve como um guia "seguro" para a otimalidade;
- Ao seguir essa trajetória, as soluções primais e duais de cada iteração caminham em direção à solução ótima, *sem se aproximar prematuramente* da fronteira da região factível;
- Dessa forma, esses métodos são conhecidos como métodos seguidores de caminho (*path-following methods*);
- Seguir estritamente o caminho central ($s_j x_j = \mu, \ j = 1, \ldots, n$) seria caro, em geral, pois teríamos que resolver o KKT perturbado a cada iteração;
- Ao invés, exigimos que os pares $s_j x_j$ sejam em média iguais a μ, isto é, $\mu = s^T x / n$;
- Por isso, μ também é conhecido como *medida de dualidade*, dado que $s^T x = c^T x - b^T p$.

Método primal-dual de pontos interiores

- Além disso, os métodos mantêm as soluções em uma vizinhança do caminho central (veremos à frente diferentes tipos de vizinhança).



- Obrigado pela atenção!
- Dúvidas?