

Pontos interiores Para Programação Linear

$$PL: \min_x c^t x \quad \text{s.a.} \quad Ax = b, \quad x \geq 0.$$

$$SP(\mu_k) = \min_x c^t x - \mu_k \sum_{i=1}^n \ln(x_i)$$

$$\text{s.a.} \quad Ax = b.$$

KKT de $SP(\mu_k)$:

$$\begin{cases} c - \mu_k \begin{bmatrix} 1/x_1 \\ \vdots \\ 1/x_n \end{bmatrix} + A^t y = 0 \\ Ax - b = 0 \end{cases}$$

$$F(x, y) = \begin{bmatrix} c - \mu_k X^{-1} e + A^t y \\ Ax - b \end{bmatrix} = 0$$

onde

$$X = \text{diag}(x) = \begin{bmatrix} x_1 & & 0 \\ & \ddots & \\ 0 & & x_n \end{bmatrix} \quad x > 0 \quad \Rightarrow \quad X^{-1} = \begin{bmatrix} 1/x_1 & & 0 \\ & \ddots & \\ 0 & & 1/x_n \end{bmatrix}$$

$$e = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}.$$

$$\text{Note que } X^{-1}e = \begin{bmatrix} 1/x_1 & 0 \\ \vdots & \vdots \\ 0 & 1/x_m \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} 1/x_1 \\ \vdots \\ 1/x_m \end{bmatrix}$$

Direção Newtoniana:

$$d^k \text{ solução } F'(x^k, y^k) d = -F(x^k, y^k).$$

Obtemos que

$$F'(x, y) = \begin{bmatrix} \mu_k X^{-2} & A^t \\ A & 0 \end{bmatrix}$$

Veja que, à medida que μ_k se aproxima de 0, algum x_i potencialmente vai à zero. Neste caso, a parcela $\mu_k X^{-2}$ ficará instável, dado que

$$\mu_k X^{-2} = \begin{bmatrix} \frac{\mu_k}{x_1^2} & & 0 \\ & \ddots & \\ 0 & & \frac{\mu_k}{x_m^2} \end{bmatrix}.$$

O segredo para ganhar estabilidade numérica é controlar os termos $\frac{\mu_k}{x_i^2}$.

Definimos $z_i = \frac{\mu_k}{x_i}$, $\forall i$. Note que

$$\mu_k X^{-1} e = \begin{bmatrix} \frac{\mu_k}{x_1} \\ \vdots \\ \frac{\mu_k}{x_n} \end{bmatrix} = z. \text{ Assim, o}$$

sistema $F(x, y) = 0$ fica

$$\tilde{F}(x, y, z) = \begin{bmatrix} c - z + A^t y \\ Ax - b \\ x_1 z_1 - \mu_k \\ \vdots \\ x_n z_n - \mu_k \end{bmatrix} = 0$$

$$(z_i = \frac{\mu_k}{x_i} \iff_{x_i > 0} x_i z_i - \mu_k = 0)$$

A direcção Newtoniana para $\tilde{F}(x, y, z) = 0$ terá uma componente adicional relativa à z .

$$\tilde{F}'(x, y, z) d = -\tilde{F}(x, y, z).$$

$$\tilde{F}'(x, y, z) = \begin{bmatrix} 0 & A^t & -I_m \\ A & 0 & 0 \\ \begin{bmatrix} z_1 & 0 & \dots & 0 \\ 0 & z_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & z_m \end{bmatrix} & 0 & \begin{bmatrix} x_1 & & 0 \\ & \ddots & \\ 0 & & x_m \end{bmatrix} \end{bmatrix}$$

Denotando

$$X = \text{diag}(x) = \begin{bmatrix} x_1 & & 0 \\ & \ddots & \\ 0 & & x_m \end{bmatrix}, \quad Z = \text{diag}(z) = \begin{bmatrix} z_1 & & 0 \\ & \ddots & \\ 0 & & z_m \end{bmatrix},$$

obtemos

$$\tilde{F}'(x, y, z) = \begin{bmatrix} 0 & A^t & -I \\ A & 0 & 0 \\ Z & 0 & X \end{bmatrix}$$

Sistema Newtoniano:

$$\begin{bmatrix} 0 & A^t & -I \\ A & 0 & 0 \\ Z & 0 & X \end{bmatrix} \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix} = - \begin{bmatrix} c - z + A^t y \\ Ax - b \\ XZe - \mu e \end{bmatrix}$$

não permiti frações!

Este sistema é base para os métodos de pontos interiores "primal-dual".

↳ primal-dual afim escala

↳ seguidor de caminhos

↳ Mehrotra

↳ primal-dual com múltiplas correções

(CPLEX, GUROBI, etc)

Tópicos PO
(2021/1)

"Método" pontos interiores para PL:

(*) (x_0, y_0, z_0) ponto inicial; $k \leftarrow 0$
 $\mu_0 > 0$ penalizador inicial;
 $x_0 > 0, z_0 > 0$ ($z_i = \mu/x_i > 0$)

$k \leftarrow k+1$

$x^{k+1} = x^k + \alpha_k^p d_x^k$
 $y^{k+1} = y^k + \alpha_k^d d_y^k$
 $z^{k+1} = z^k + \alpha_k^d d_z^k$

decreça μ : (*)
 $0 < \mu^{k+1} < \mu^k$

PARAR (*) → SIM
 NÃO

Calcule $\alpha_k^p \in (0,1]$ e $\alpha_k^d \in (0,1]$
 tais que $x^k + \alpha_k^p d_x^k > 0$ e
 $z^k + \alpha_k^d d_z^k > 0$ (*)

(*) Calcule a direção de Newton $d = (d_x, d_y, d_z)$, resolvendo
 $\tilde{F}'(x^k, y^k, z^k) d = -\tilde{F}(x^k, y^k, z^k)$

* Como se calcula na prática d^k ?
↳ tratando o sistema Newtoniano, e usando Cholesky com reordenamento.

* Como se calcula α_k^p e α_k^d ?
↳ buscamos as componentes de $(d_x^k)_i$ e $(d_z^k)_i$ que potencialmente tomam x_i^{k+1} e z_i^{k+1} negativos, e tomamos α_k^p e α_k^d como "passos máximos" que garantam $x_i^{k+1} > 0$ e $z_i^{k+1} > 0$

* Como inicializar (x^0, y^0, z^0) ?
Em tese, pode ser qualquer ponto em que $x^0 > 0$ e $z^0 > 0$ (p.ex., $x_i^0 = 1, \forall i$, $z_i^0 = 1, \forall i$, e $y = 0$). Porém na prática existem heurísticas para uma boa inicialização.

(*) Como parar?

↳ SUCESSO: \approx KKT. (\Leftrightarrow solução ótima)^{PL}

↳ INSUCESSO: • nº máx. iterações atingido;

• cálculo d falhou;

• falta de progresso;

• $\|(x, z)\|$ muito grande

↳ PL provavelmente é

ilimitado ou inviável.

(*) Como diminuir μ ?

MUITO IMPORTANTE para estabilidade do método. Lembre-se que no sistema

$\tilde{F}(x, y, z) = 0$, há as expressões

$$x_1 z_1 - \mu = 0, \dots, x_n z_n - \mu = 0.$$

Logo, os produtos $x_i z_i$ ficam entrados no valor μ . Se μ for muito rápido

para zero, corremos o risco de obtermos produtos $x_i z_i$ desbalanceados, o que gera mal condicionamento do sistema

$\tilde{F}(x, y, z) d = -\tilde{F}(x, y, z)$. Se controlamos a velocidade de decaimento de μ , os produtos $x_i z_i$ tendem a ficar próximos entre si, favorecendo a estabilidade numérica.

Uma ideia que funciona bem é tomar μ em função das médias dos produtos

$x_i z_i$, $\forall i$, isto é,

$$\frac{\sum_{i=1}^m x_i z_i}{m}.$$

Detalhes nas aulas de "Tópicos em PO" 2021/1, veja página da disciplina.