Universidade Estadual de Campinas
FACULDADE DE ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO
DT-FEEC-UNICAMP

Método de Pontos Interiores em Programação Linear

1 Introdução

1.1 Pontos Interiores X Simplex

- Ambos eficientes
- Simplex: muitas iterações "simples" pelas arestas
- Pontos Interiores: poucas iterações "caras" pelos pontos interiores.

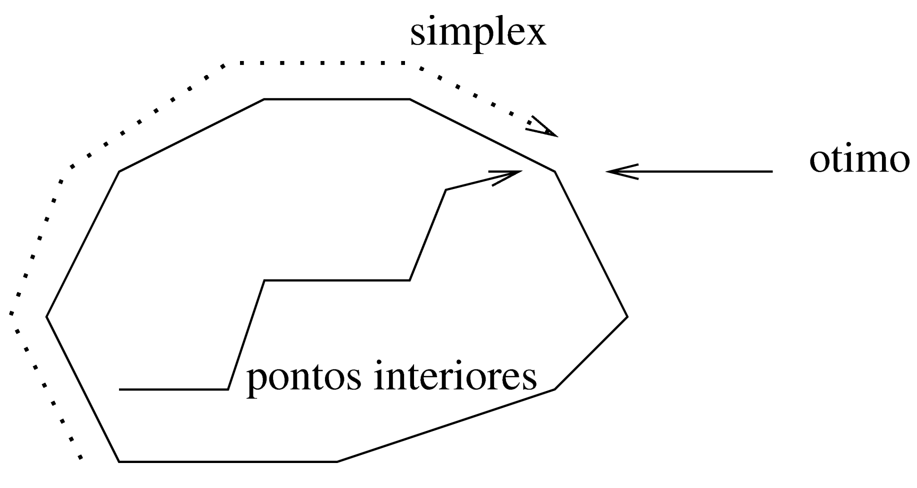

Figure 1: Simplex X Pontos Interiores

1.2 Notação

- $A, B, C...$ onde $A = \{a_{ij}\}$ - Matrizes
- $a, b, c...$ onde $a = \{a_i\}$ - Vetores colunas

1.3 Programação Linear - Formas Padrões

$$Primal : \begin{cases} min & c^t x \\ sa & Ax = b \\ & x \geq 0 \end{cases}$$

$$Dual : \begin{cases} max & b^t y \\ sa & A^t y \leq c \\ & y \ \ livre \end{cases} \quad ou \ \ Dual : \begin{cases} max & b^t y \\ sa & A^t y + z = c \\ & z \geq 0, y \ \ livre \end{cases}$$

1.4 Definições

- **Def. 1:** Ponto **interior**: x^o tal que $x^o > 0$ é ponto interior do primal.
- **Def. 2:** Ponto **factível**: x^o tal que $Ax^o = b, x^o \geq 0$ é um ponto factível do primal.
- **Def. 3:** Ponto **interior factível** (satisfaz 1 e 2): x^o tal que $Ax^o = b, x^o > 0$ é um ponto interior factível do primal.
- **Def. 4: Gap** (ou GAP): diferença entre o valor do primal e do dual. Ex.: $GAP = c^t x - b^t y$.

1.5 Condições de Otimalidade

i) Primal factibilidade: $b - Ax = 0, x \geq 0$

ii) Dual factibilidade: $c - A^t y - z = 0, z \geq 0$

iii) Folga Complementar: $x_i z_i = 0$

1.6 Problemas de Mínimos Quadrados

Problema: $Min_{x \in R} \Phi(x) = \frac{1}{2} \| b - Ax \|^2$

Logo:

$$\Phi(x) = \frac{1}{2}(b - Ax)^t (b - Ax) = \frac{1}{2}(b^t b - b^t Ax - x^t A^t b + x^t A^t Ax)$$

$$\nabla \Phi(x) = 0 \quad \Rightarrow \quad x = (A^t A)^{-1} A^t b$$

$\nabla^2 \Phi(x) = A^t A > 0 \quad \Rightarrow \quad$ é mínimo global.

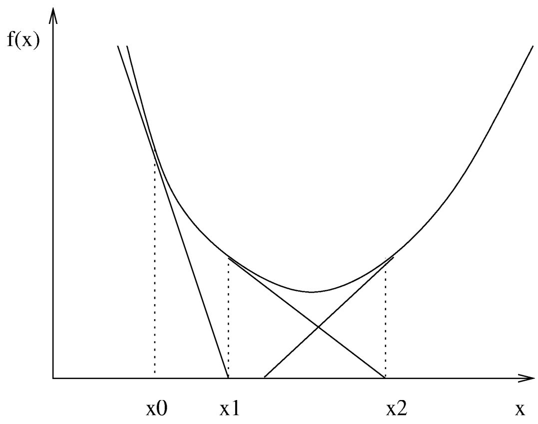

Figure 2: Método de Newton

1.7 Métodos de Gradiente e de Newton

- Método Gradiente (monovariável):

$$\Phi(x) \cong \Phi(x^0) + \nabla\Phi^t(x^0)(x - x^0)$$

$$\Phi(x) = 0 \;\Rightarrow\; \nabla\Phi^t(x^0)x = \nabla\Phi^t(x^0)x^0 - \Phi(x^0)$$

$$x = x^0 - \frac{\Phi(x^0)}{\nabla\Phi(x^0)}$$

- Método de Newton:

$$\Phi(x) \cong \Phi(x^k) + \nabla\Phi(x^k)(x - x^k) + \frac{1}{2}(x - x^k)^t H(x^k)(x - x^k)$$

onde $H(x^k) = \nabla(\nabla\Phi(x^k))$ (Hessiana de $\Phi(x)$)

$$\nabla\Phi(x) = 0 \;\Rightarrow\; \nabla\Phi(x^k) + H(x^k)(x - x^k) = 0$$

$$x^{k+1} = x^k - \alpha[H(x^k)]^{-1}\nabla\Phi(x^k)$$

- **Lema 1:** Uma direção d para melhorar a função objetivo é factível para restrições de igualdade do tipo $Ax = b$ se $Ad = 0$ (pois $A(x + d) = b \;\Rightarrow\; Ad = 0$).
- **Premissa 1:** Algoritmos de pontos interiores começam em pontos interiores factíveis e movem de ponto em ponto interior factível, em direção à solução ótima.

2 Método Primal-Afim-Escala

2.1 Base Teórica

$$X(nxn) = diag(x) = \begin{bmatrix} x_1 & 0 & ... & 0 \\ 0 & x_2 & ... & 0 \\ . & . & ... & . \\ . & . & ... & . \\ 0 & 0 & ... & x_n \end{bmatrix}$$

$$Min_{y,z}\ \frac{1}{2} \| Xz \|^2 \quad sa \quad z = c - A^t y$$

$$Min_{y,z} \frac{1}{2} \| Xz \|^2 = Min_z \frac{1}{2} \| X(c - A^t y) \|^2$$

$$Z = \| X(c - A^t y) \|^2$$

$$= c^t X^t X c - c^t X^t X A^t y - y^t A X^t X c + y^t A X^t X A^t y$$

$X^t = X$, $\nabla_y Z = 0$: $-AX^2 c + AX^2 A^t y = 0$

$$z = c - A^t (AX^2 A^t)^{-1} AX^2 c$$

Teorema (Dikin): Dados x tal que $Ax = b, x > 0, y = (AX^2A^t)^{-1}AX^2c$, então a direção $d = -X^2 z$ é uma direção de **descida factível**.

Assim, $d = -X^2 z$ minimiza $\| Xz \|^2$:

$$x^{k+1} = x^k + \alpha^k d^k \text{ com:}$$

$$\alpha^k = \tau min_{\delta_i^k<0}(-\frac{x_i^k}{\delta_i^k}), 0 < \tau < 1, d^k = \{\delta_i^k\}$$

2.2 Algoritmo

1. Dados $\tau \in (0,1)$ e $x^0 \mid Ax^0 = b,\ x^0 > 0, k = 0$:
2. Faça até convergir:
   (a) $y^k = (A(X^k)^2A^t)^{-1}A(X^k)^2c$
   (b) $z^k = c - A^ty^k$
   (c) $d^k = -(X^k)^2z^k$
   (d) $\alpha^k = \tau min_{\delta_i^k<0}\{-\frac{x_i^k}{\delta_i^k}\}$
   (e) $x^{k+1} = x^k + \alpha^kd^k$
   (f) $k \leftarrow k + 1$
3. Fim Faça

- a) Critérios de convergências:
  - Gap relativo: $\frac{\|X^k z^k\|}{1+\|b^t y^k + c^t x^k\|} \leq \epsilon$
  - Variação do valor da função objetivo: $\frac{\|c^t x^{k+1} - c^t x^k\|}{1+\|c^t x^k\|} \leq \epsilon$
- b) Ponto inicial interior factível x^0, resolver:

$$(PM): \begin{cases} min & c^t x + M\sigma \\ sa & Ax + p\sigma & = & b \\ & (x, \sigma) & \geq & 0 \end{cases}$$

onde: $p = b - Ax^0$. Iniciar com $(x^0, 1)$.

- c) Cálculo de y^k:

$$(A(X^k)^2 A^t) y^k = A(X^k)^2 c$$

- d) **afim:** espaço afim, **escala:** $\tilde{x} = X^{-1}x$:

$$(P) \begin{cases} min & c^t x \\ sa & Ax \leq b \\ & x \geq 0 \end{cases} \rightarrow \; x = X\tilde{x} \; \Rightarrow \; (\tilde{P}) \begin{cases} min & \tilde{c}^t \tilde{x} \\ sa & \tilde{A}\tilde{x} \leq b \\ & \tilde{x} \geq 0 \end{cases}$$

$\tilde{c} = Xc$, $\tilde{A} = AX$.

2.3 Projeção Afim-Escala

Problema:

$$\begin{cases} max & c^t x \\ sa & Ax = b \\ & x \geq 0 \end{cases}$$

Direção d tal que:

$$\begin{cases} max & c^t(x^0 + d) \\ sa & A(x^0 + d) = b \\ & \| \ d \ \|^2 = 1 \end{cases}$$

Lagrangeano:

$$L(d, y, \lambda) = c^t(x^0 + d) + \lambda(1 - d^t d) - y^t(A(x^0 + d) - b)$$

Dual:

$$Min_{d,y,\lambda} L(d, y, \lambda).$$

Otimalidade:

i) $\frac{\partial L}{\partial d} = c - 2\lambda d - A^t y = 0$

ii) $\frac{\partial L}{\partial \lambda} = 1 - d^t d = 0$

iii) $\frac{\partial L}{\partial y} = A(x^0 + d) - b = 0$

iii): $Ax^0 - b + Ad = 0 \;\Rightarrow\; Ad = 0$

$\lambda = \frac{1}{2}$:

i): $c - d - A^t y = 0 \;\Rightarrow\; d = c - A^t y$

$\Rightarrow \quad y = (AA^t)^{-1}Ac$

$d = c - A^t(AA^t)^{-1}Ac = (I - A^t(AA^t)^{-1}A)c = Pc$

$P = (I - A^t(AA^t)^{-1}A) =$ **matriz de projeção ortogonal ao espaço nulo de** A.

Figura 3: $d = d_a + d_b$, $d_a = A^t x$

$Ad_b = A(d - d_a) = 0$ (espaço nulo de A).

Assim:

$Ad = Ad_a = AA^t x$

$x = (AA^t)^{-1}Ad \;\Rightarrow\; d_a = A^t(AA^t)^{-1}Ad$

$d_b = d - A^t(AA^t)^{-1}Ad = (I - A^t(AA^t)^{-1}A)d = P(A)d$

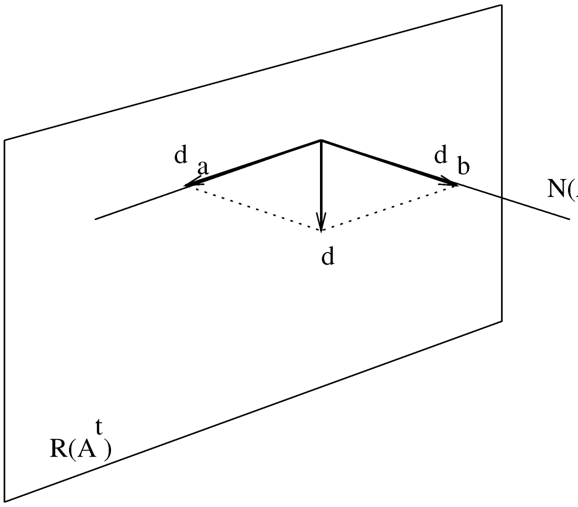

Figure 3: Matriz de projeção ortogonal de A

Problema $(\tilde{P})$:

$$\tilde{d} = -\tilde{P}\tilde{c} = -(I - \tilde{A}^t(\tilde{A}\tilde{A}^t)^{-1}\tilde{A})\tilde{c}$$

$\tilde{A} = AX,\ \tilde{c} = Xc$, iteração k:

$$\tilde{d}^k = -(I - X^kA^t(AX^kX^kA^t)^{-1}AX^k)X^kc$$

$$= -X^kc + X^kA^t(A(X^k)^2A^t)^{-1}A(X^k)^2c$$

Logo:

$$\begin{cases} d^k &= X^k\tilde{d}^k = -(I - (X^k)^2A^t(A(X^k)^2A^t)^{-1}A).(X^k)^2c \\ x^{k+1} &= x^k + \alpha^k d^k \end{cases}$$

Primal-afim-escala:

$$\begin{cases} d^k &= -(X^k)^2 z^k \\ z^k &= c - A^t y^k \\ y^k &= (A(X^k)^2 A^t)^{-1} A(X^k)^2 c \end{cases}$$

$$\Rightarrow \quad \begin{cases} d^k &= -(I - (X^k)^2 A^t (A(X^k)^2 A^t)^{-1} A).(X^k)^2 c \\ x^{k+1} &= x^k + \alpha^k d^k \end{cases}$$

2.4 Exemplo (Frannie's Firewood Problem)

Frannie vende 3 "cordas" de lenha todo final do ano. Pode vender meia "corda" a $U\$90$ ou uma "corda" a $U\$150$. Como maximizar o lucro?

$$Modelo: \begin{cases} max & 90x_1 + 150x_2 \\ sa & \frac{1}{2}x_1 + x_2 \leq 3 \\ & x_1, x_2 \geq 0 \end{cases}$$

$x_3 =$ variável de folga, $(1, \frac{1}{2}, 2) =$ ponto inicial interior factível

Resultado: tabela 1, figura 4.

x_1	x_2	x_3	$c^t x$
$c_1 = 90$	$c_2 = 150$	$c_3 = 0$	$b = 3$
1.00	0.50	2.00	165.00
1.73	0.80	1.34	275.51
2.82	1.13	0.46	423.40
3.92	1.00	0.04	502.84
5.66	0.14	0.03	530.45
5.88	0.05	0.01	537.25
5.98	0.01	0.003	539.22
5.99	0.005	0.000	539.79
6.00	0.001	0.000	539.95
6.00	0.000	0.000	539.99

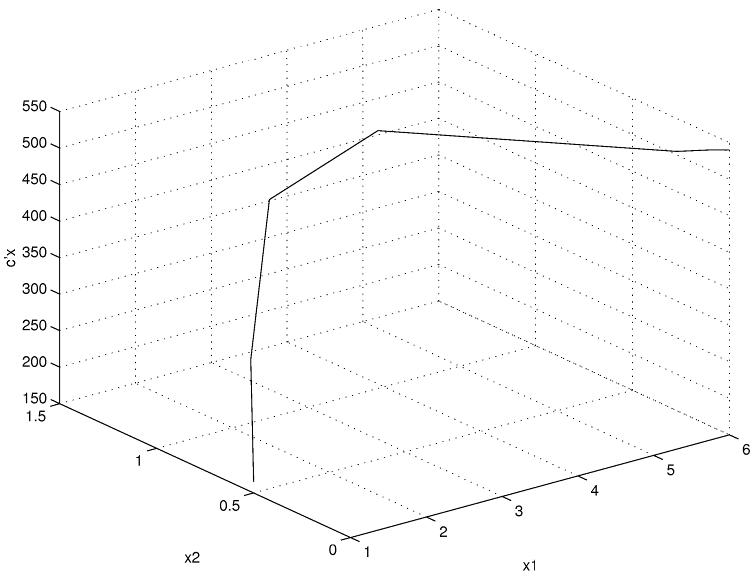

Figure 4: Evolução dos pontos interiores do problema de Frannie

3 Método Dual-Afim-Escala

3.1 Introdução

Problema:

$$\begin{cases} min & \frac{1}{2} \| Zx \|^2 \\ sa & Ax = b \end{cases} \qquad Z = diag(z_i)$$

Lagrangeano:

$$L(x, w) = \frac{1}{2} \parallel Zx \parallel^2 + w^t(b - Ax) = \frac{1}{2}(x^t ZZx) + w^t(b - Ax)$$

Condições de otimalidade:

$$\begin{cases} \frac{\partial L}{\partial x} = 0 & \rightarrow & Z^2 x - A^t w = 0 \\ \frac{\partial L}{\partial w} = 0 & \rightarrow & b - Ax = 0 \end{cases}$$

Ou seja:

$$x = Z^{-2} A^t w; \quad (AZ^{-2}A^t)w = b \Rightarrow x = Z^{-2}A^t(AZ^{-2}A^t)^{-1}b$$

Hessiana:

$$H = \begin{bmatrix} Z^2 & -A \\ -A^t & 0 \end{bmatrix} = -AA^t < 0$$

$$z = c - A^t y \Rightarrow dz = -A^t dy$$

Direção $dz = -Z^2x$:

$dz = -Z^2[Z^{-2}A^t(AZ^{-2}A^t)^{-1}b]$

$= -A^t(AZ^{-2}A^t)^{-1}b = -A^tdy$

Assim:

$dy = (AZ^{-2}A^t)^{-1}b$

Teorema: Dados (y, z) tais que $z = c - A^ty$, $z \geq 0$, $x = Z^{-2}A^t(AZ^{-2}A^t)^{-1}b$, a direção dada por:

$(dy, dz) = (AZ^{-2}A^t)^{-1}b, -Z^2x)$

é dual factível e é de subida.

3.2 Algoritmo

1. Dados (y^0, z^0) tal que $A^ty^0 + z^0 = c,\ z^0 > 0$ e $\tau \in (0, 1),\ k = 0$,
2. Faça até convergência:
   (a) $dy^k = (A(Z^k)^{-2}A^t)^{-1}b$
   (b) $dz^k = -A^tdy^k$
   (c) $x^k = -(Z^k)^{-2}dz^k$

(d) $\alpha^k = \tau min_{\delta z_i^k < 0}\{-\frac{z_i^k}{\delta z_i^k}\}$

(e) $y^{k+1} = y^k + \alpha^k dy^k$

(f) $z^{k+1} = z^k + \alpha^k dz^k$ (ou $z^{k+1} = c - A^t y^{k+1}$)

(g) $k \leftarrow k+1$

3. Fim Faça

- a) Critério de convergência originalmente utilizado:

  $\frac{\|b^t y^k - b^t y^{k+1}\|}{max(1, \|b^t y^k\|)} \leq \epsilon.$

- b) Ponto inicial dual factível (y^0, z^0), resolver o problema:

$$\begin{cases} max & b^t y - M\sigma \\ sa & A^t y + z - e\sigma = c \\ & z \geq 0 \end{cases}$$

  aplicando o método dual-afim-escala até $\sigma < 0$, com valor inicial y^0 qualquer (livre), $\sigma^0 = -2min_j(c_j - A_j^t y^0)$.

- c) $z^{k+1} = c - A^t y^{k+1} = c - A^t(y^k + \alpha^k dy^k) =$

  $c - A^t y^k - \alpha^k A^t dy^k = z^k + \alpha^k dz^k$

- d) O cálculo de x^k é dispensável, a não ser que se utilize no critério de convergência.

- e) Como y é livre, não é feito teste de barreira.
- f) Grande custo computacional no cálculo de $A(Z^k)^{-2}A^t$.

4 Método Primal-Dual-Afim-Escala

4.1 Introdução

Condições de otimalidade:

$$F(x,y,z) = \begin{bmatrix} F_p \\ F_d \\ F_a \end{bmatrix} = \begin{bmatrix} Ax - b \\ A^t y + z - c \\ XZe \end{bmatrix} = 0$$

Aproximação linear fornece:

$(x,y,z) \cong (x^0,y^0,z^0) - J^{-1}(x^0,y^0,z^0)F(x^0,y^0,z^0)$

pois:

$F(x,y,z) \cong F(x^0,y^0,z^0) + J(x^0,y^0,z^0)[(x,y,z)-$

$(x^0,y^0,z^0))] = 0$

$$-F(x^0,y^0,z^0) = \begin{bmatrix} b - Ax^0 \\ c - A^t y^0 - z^0 \\ -X^0 Z^0 e \end{bmatrix} = \begin{bmatrix} r_p \\ r_d \\ r_a \end{bmatrix} = r$$

e:

$$J(x^0,y^0,z^0) = \begin{bmatrix} \nabla F_p^t \\ \nabla F_d^t \\ \nabla F_a^t \end{bmatrix} = \begin{bmatrix} A & 0 & 0 \\ 0 & A^t & I \\ Z^0 & 0 & X^0 \end{bmatrix}$$

Assim:

$$(x,y,z) = (x^0,y^0,z^0) + \begin{bmatrix} A & 0 & 0 \\ 0 & A^t & I \\ Z^0 & 0 & X^0 \end{bmatrix}^{-1} \begin{bmatrix} r_p \\ r_d \\ r_a \end{bmatrix} = (x^0,y^0,z^0) + d$$

onde:

$$d = \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix} = \begin{bmatrix} A & 0 & 0 \\ 0 & A^t & I \\ Z^0 & 0 & X^0 \end{bmatrix}^{-1} \begin{bmatrix} r_p \\ r_d \\ r_a \end{bmatrix}$$

Pode-se resolver o sistema:

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^t & I \\ Z^0 & 0 & X^0 \end{bmatrix} \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix} = \begin{bmatrix} r_p \\ r_d \\ r_a \end{bmatrix}$$

Considerando:

$$\begin{cases} Adx & = & r_p \\ A^t dy + dz & = & r_d \\ Z^0 dx + X^0 dz & = & r_a \end{cases}$$

$dz = r_d - A^t dy \quad \Rightarrow \quad :$

$$Z^0 dx + X^0 dz = Z^0 dx + X^0 (r_d - A^t dy) = r_a$$

$$Z^0 dx - X^0 A^t dy = r_a - X^0 r_d$$

Ou seja:

$$-(X^0)^{-1} Z^0 dx + A^t dy = -(X^0)^{-1} r_a + r_d$$

que fornece:

$$\begin{bmatrix} A & 0 \\ -D & A^t \end{bmatrix} \begin{bmatrix} dx \\ dy \end{bmatrix} = \begin{bmatrix} r_p \\ r_d - (X^0)^{-1} r_a \end{bmatrix}$$

onde $D = X^{-1}Z$. Note que:

$$\begin{cases} Adx & = & r_p \\ -Ddx + A^t dy & = & r_d - X^{-1} r_a \\ dx & = & D^{-1}[A^t dy - r_d + X^{-1} r_a] \end{cases}$$

e:

$$D^{-1} A^t dy = dx + D^{-1}(r_d - X^{-1} r_a)$$

$$(AD^{-1}A^t) dy = Adx + AD^{-1} r_d - AD^{-1} X^{-1} r_a$$

$$\Rightarrow \quad (AD^{-1}A^t) dy = r_p + AD^{-1} r_d - AZ^{-1} r_a$$

4.2 Algoritmo

1. Dados (x^0, y^0, z^0) tal que $(x^0, z^0) > 0$ e $\tau \in (0, 1)$, $k = 0$,
2. Faça até convergência:
   (a) $r_p^k = b - Ax^k$
   (b) $r_d^k = c - A^t y^k - z^k$
   (c) $r_a^k = -X^k Z^k e$
   (d) $D^k = [(X^k)^{-1} Z^k]$
   (e) $dy^k = [A(D^k)^{-1} A^t]^{-1} [r_p^k + A(D^k)^{-1} r_d^k - A(Z^k)^{-1} r_a^k]$
   (f) $dx^k = (D^k)^{-1} [A^t dy^k - r_d^k + (X^k)^{-1} r_a^k]$
   (g) $dz^k = (X^k)^{-1} [r_a^k - Z^k dx^k]$
   (h) $\rho_p^k = min_{\delta x_i^k < 0} \{-\frac{x_i^k}{\delta x_i^k}\}$
   (i) $\rho_d^k = min_{\delta z_i^k < 0} \{-\frac{z_i^k}{\delta z_i^k}\}$
   (j) $\alpha_p^k = min(\tau \rho_p^k, 1)$
   (k) $\alpha_d^k = min(\tau \rho_d^k, 1)$
   (l) $x^{k+1} = x^k + \alpha_p^k dx^k$
   (m) $y^{k+1} = y^k + \alpha_d^k dy^k$
   (n) $z^{k+1} = z^k + \alpha_d^k dz^k$
   (o) $k \leftarrow k + 1$
3. Fim Faça

- A convergência pode ser testada sobre o valor de $\| F \|^2$.
- O ponto inicial (x^0, y^0, z^0) não precisa ser factível. Recomenda-se:
  - Para o primal:
    Fazendo $x = A^t\tilde{y}$, $Ax = b \ \rightarrow \ AA^t\tilde{y} = b$

    $\Rightarrow \ \tilde{y} = (AA^t)^{-1}b$
    $A^t\tilde{y} = A^t(AA^t)^{-1}b$
    $x = A^t(AA^t)^{-1}b$
    $\epsilon_1 = max(-min_i x_i, \epsilon_2, \frac{\|b\|_1}{\epsilon_2\|A\|_1})$, onde $\epsilon_2 > 0$
    $x_i^0 = max(x_i, \epsilon_1)$
    $(\| b \|_1 = \Sigma_{i=1}^m \mid b_i \mid, \| A \|_1 = max_j \Sigma_{i=1}^m \mid a_{ij} \mid)$
  - Para dual:
    $A^t y + z = c, y$ livre e $z > 0 \ \Rightarrow \ y^0 = 0$, e:

$$z^0 = \begin{cases} z_i + \epsilon_3 & se \ \ z_i \geq 0 \\ -z_i & se \ \ z_i \leq -\epsilon_3 \\ \epsilon_3 & se \ \ -\epsilon_3 \leq z_i \leq 0 \end{cases}$$

    onde $\epsilon_3 = 1 + \| c \|_1$. Estes pontos procuram ser bem posicionados, longe da fronteira ($x_i z_i$ não muito pequenos).
- O tamanho do passo para y é o mesmo de para z (α_d^k) para garantir que ocorra $c - A^t y^k - z^k = 0$ na convergência:

$$c - A^t dy^{k+1} - z^{k+1} = c - A^t(y^k + \alpha_d^k dy^k) - (z^k + \alpha_d^k dz^k)$$

$$= (c - A^t y^k - z^k) - \alpha_d^k (A^t dy^k - dz^k)$$

- Como não precisa de ponto inicial factível, este método é melhor que o primal ou dual (não precisa de fase I).

5 Método Primal-Dual Clássico

5.1 Introdução

Primal-dual afim-escala permite que x e z aproximem ràpidamente das fronteiras $\Rightarrow$ ineficiente.

Primal-Dual Clássico acrescenta uma perturbação na condição de complementariedade:

$x_i z_i = \mu_i$

Novas condições de otimalidade:

$$\begin{cases} b - Ax & = 0 \\ c^t - A^t y - z & = 0 \\ \mu e - XZe & = 0 \end{cases}$$

μ tal que $lim_{k \to \infty} \mu^k = 0$

Estimação de μ_k:

$\gamma^k = Tr[(X^k)^t Z^k]$ onde: $Tr[X] =$ traço de X

Na maioria das implementações, adota-se:

$\mu^k = \sigma^k(\frac{\gamma^k}{n}), \ \ \sigma^k \in (0,1)$

Nota: $\sigma^k = 0 \Rightarrow$ primal-dual afim-escala, $\sigma^k = 1 \Rightarrow$

direção de centragem pois:

$$\mu e - XZe = 0$$

$$\frac{\gamma^k}{n} e - XZe = 0 \quad \Rightarrow \quad X^t Ze = \frac{X^t Z}{n} e$$

A cada iteração, tem-se o sistema $J(x^k, y^k, z^k)d^k = r^k$, ou seja:

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^t & I \\ Z^k & 0 & X^k \end{bmatrix} \begin{bmatrix} dx^k \\ dy^k \\ dz^k \end{bmatrix} = \begin{bmatrix} r_p^k \\ r_d^k \\ r_c^k \end{bmatrix} = \begin{bmatrix} b - Ax^k \\ c - A^t - z^k \\ \mu^k e - (X^k)^t Z^k e \end{bmatrix}$$

Tem-se assim, duas diferenças em relação ao método primal-dual afim-escala:

- a) troca de r_a^k por r_c^k
- b) cálculo de μ^k

5.2 Algoritmo

1. Dados (x^0, y^0, z^0) tal que $(x^0, z^0) > 0$, $\tau \in (0,1), \sigma \in (0,1)$ e $k = 0$,
2. Faça até convergência:
   (a) $\gamma^k = Tr[X^k Z^k]$
   (b) $\mu^k = \sigma(\frac{\gamma^k}{n})$
   (c) $r_p^k = b - Ax^k$
   (d) $r_d^k = c - A^t y^k - z^k$
   (e) $r_c^k = \mu^k e - X^k Z^k e$
   (f) $D^k = [(X^k)^{-1} Z^k]$
   (g) $dy^k = [A(D^k)^{-1}A^t]^{-1}[r_p^k + A(D^k)^{-1}r_d^k - A(Z^k)^{-1}r_c^k]$
   (h) $dx^k = (D^k)^{-1}[A^t dy^k - r_d^k + (X^k)^{-1} r_c^k]$
   (i) $dz^k = (X^k)^{-1}[r_c^k - Z^k dx^k]$
   (j) $\rho_p^k = min_{\delta x_i^k < 0}\{-\frac{x_i^k}{\delta x_i^k}\}$
   (k) $\rho_d^k = min_{\delta z_i^k < 0}\{-\frac{z_i^k}{\delta z_i^k}\}$
   (l) $\alpha_p^k = min(\tau \rho_p^k, 1)$
   (m) $\alpha_d^k = min(\tau \rho_d^k, 1)$
   (n) $x^{k+1} = x^k + \alpha_p^k dx^k$
   (o) $y^{k+1} = y^k + \alpha_d^k dy^k$
   (p) $z^{k+1} = z^k + \alpha_d^k dz^k$
   (q) $k \leftarrow k + 1$
3. Fim Faça

- O critério de convergência e a inicialização deste método podem ser os mesmos do método primal-dual afim-escala. É recomendável inicializar com μ^0 alto.

- Dependendo dos valores de σ e τ, obtemos algoritmos de diversas naturezas (complexidade polinomial, convergência superlinear, etc.).

- Os valores típicos de τ estão entre $(0.995, 0.99995)$.

- Quando $\gamma^k < 1$, recomenda-se utilizar $\mu^k = \frac{(\gamma^k)^2}{n}$ para procurar acelerar a convergência.

6 Método Preditor-Corretor

6.1 Introdução

Baseado em 3 componentes:

- direção afim-escala $\tilde{d}$ (direção de Newton, preditor).
- direção de centragem, definido pelo σ do primal-dual clássico.
- direção de correção $\hat{d}$, que tenta compensar a aproximação linear de Newton.

Idéia: calcular a direção afim-escala e estudar o progresso ao longo desta direção, atuando na perturbação μ (centragem) e na correção não-linear.

No ponto (x, y, z):

$$(1) \begin{cases} Ad\tilde{x} & = & r_p \\ A^t d\tilde{y} + d\tilde{z} & = & r_d \\ Zd\tilde{x} + Xd\tilde{z} & = & r_a = -XZe \end{cases}$$

Obtém-se, então, o ponto $(\tilde{x}, \tilde{y}, \tilde{z})$, onde:

$$\begin{bmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \end{bmatrix} = \begin{bmatrix} x + d\tilde{x} \\ y + d\tilde{y} \\ z + d\tilde{z} \end{bmatrix}$$

A seguir, determinar a direção $(d\hat{x}, d\hat{y}, d\hat{z})$:

$$
(2)\begin{cases} Ad\hat{x} & = & 0 \\ A^t d\hat{y} + d\hat{z} & = & 0 \\ Zd\hat{x} + Xd\hat{z} & = & \mu e - (D\tilde{x}D\tilde{z})e = r_c \end{cases}
$$

onde $D\tilde{x} = diag(d\tilde{x})$ e $D\tilde{z} = diag(d\tilde{z})$. Finalmente, a direção final (dx, dy, dz) é determinada somando (1) e (2):

$$
\begin{cases} A(d\tilde{x} + d\hat{x}) & = & r_p \\ A^t(d\tilde{y} + d\hat{y}) + (d\tilde{z} + d\hat{z}) & = & r_d \\ Z(d\tilde{x} + d\hat{x}) + X(d\tilde{z} + d\hat{z}) & = & r_a + r_c = r_s \end{cases}
$$

onde:

$$
\begin{cases} r_a & = & -XZe \\ r_c & = & \mu e - (D\tilde{x}D\tilde{z})e \end{cases}
$$

6.2 Algoritmo

1. Dados (x^0, y^0, z^0) tal que $(x^0, z^0) > 0$, $\tau \in (0, 1)$ e $k = 0$,
2. Faça até convergência:
   (a) $r_p^k = b - Ax^k$
   (b) $r_d^k = c - A^t y^k - z^k$
   (c) $r_a^k = -X^k Z^k e$
   (d) $D^k = [(X^k)^{-1} Z^k]$
   (e) $d\tilde{y}^k = [A(D^k)^{-1}A^t]^{-1}[r_p^k + A(D^k)^{-1}r_d^k - A(Z^k)^{-1}r_a^k]$
   (f) $d\tilde{x}^k = (D^k)^{-1}[A^t d\tilde{y}^k - r_d^k + (X^k)^{-1}r_a^k]$
   (g) $d\tilde{z}^k = (X^k)^{-1}[r_a^k - Z^k d\tilde{x}^k]$
   (h) $\tilde{\rho}_p^k = min_{\delta\tilde{x}_i^k<0}\{-\frac{x_i^k}{\delta\tilde{x}_i^k}\}$
   (i) $\tilde{\rho}_d^k = min_{\delta\tilde{z}_i^k<0}\{-\frac{z_i^k}{\delta\tilde{z}_i^k}\}$
   (j) $\tilde{\alpha}_p^k = min(\tau\tilde{\rho}_p^k, 1)$
   (k) $\tilde{\alpha}_d^k = min(\tau\tilde{\rho}_d^k, 1)$
   (l) $\tilde{\gamma}^k = (x^k + \tilde{\alpha}_p^k d\tilde{x}^k)^t(z^k + \tilde{\alpha}_d^k d\tilde{z}^k)$, $\gamma^k = Tr[X^k Z^k]$
   (m) $\sigma^k = \begin{cases} (\frac{\tilde{\gamma}^k}{\gamma^k})^3 & se \quad \gamma^k > 1 \\ (\frac{\gamma^k}{\sqrt{n}}) & se \quad \gamma^k \leq 1 \end{cases}$
   (n) $\mu^k = \sigma^k(\frac{\gamma^k}{n})$
   (o) $r_s^k = r_a^k + \mu^k e - (D\tilde{x}^k)(D\tilde{z}^k)e$
   (p) $dy^k = [A(D^k)^{-1}A^t]^{-1}[r_p^k + A(D^k)^{-1}r_d^k - A(Z^k)^{-1}r_s^k]$
   (q) $dx^k = (D^k)^{-1}[A^t dy^k - r_d^k + (X^k)^{-1}r_s^k]$

(r) $dz^k = (X^k)^{-1}[r_s^k - Z^k dx^k]$

(s) $\rho_p^k = min_{\delta x_i^k < 0}\{-\frac{x_i^k}{\delta x_i^k}\}$

(t) $\rho_d^k = min_{\delta z_i^k < 0}\{-\frac{z_i^k}{\delta z_i^k}\}$

(u) $\alpha_p^k = min(\tau \rho_p^k, 1)$

(v) $\alpha_d^k = min(\tau \rho_d^k, 1)$

(w) $x^{k+1} = x^k + \alpha_p^k dx^k$

(x) $y^{k+1} = y^k + \alpha_d^k dy^k$

(y) $z^{k+1} = z^k + \alpha_d^k dz^k$

(z) $k \leftarrow k + 1$

3. Fim Faça

Nota:

- O critério de convergência e a inicialização deste método podem ser os mesmos do método primal-dual afim-escala.
- Dois sistemas lineares precisam ser resolvidos, utilizando a mesma relação: $A(D^k)^{-1}A^t = L^k(L^k)^t$.
- Espera-se que o esfôrço para resolver dois sistemas lineares seja recompensado pela redução no número de iterações.
- Este é o método com melhores resultados teóricos e práticos (tem convergência quadrática).

7 Método de Barreira Logarítimica

7.1 Introdução

Seja o problema:

$$(P1): \begin{cases} max & c^t x \\ sa & Ax = b \\ & x \geq 0 \end{cases}$$

Substituindo a restrição $x \geq 0$ na forma:

$$(\tilde{P}1): \begin{cases} max & c^t x + \mu^t f(x) \\ sa & Ax = b \\ onde: & f(x) = ln(x) = \begin{bmatrix} ln(x_1) \\ ln(x_2) \\ . \\ . \\ . \\ ln(x_n) \end{bmatrix} \end{cases}$$

μ pode ser assumido um escalar ($\mu \in \mathbf{R}$).

Exemplo:

$$\begin{cases} min & z = 5x_1 + 3x_2 & \\ sa & x_1 + x_2 \geq 1 & (eq.1) \\ & 0 \leq x_1 \leq 2 & (eq.2) \\ & 0 \leq x_2 \leq 2 & (eq.3) \end{cases}$$

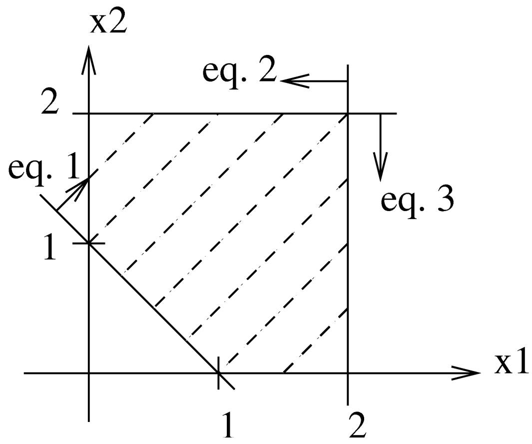

Figure 5: Região factível do problema

Região factível na figura 5. Com as variáveis de folga:

$$\begin{cases} min & z = 5x_1 + 3x_2 \\ sa & x_1 + x_2 - x_3 = 1 \\ & x_1 + x_4 = 2 \\ & x_2 + x_5 = 2 \\ & x_i \geq 0, i = 1, 2, 3, 4, 5 \end{cases}$$

Com a adição da função barreira:

$$\begin{cases} min\ \tilde{z} & = 5x_1 + 3x_2 - \mu(ln(x_1) + ln(x_2) + ln(x_3) + \\ & ln(x_4) + ln(x_5)) \\ sa & x_1 + x_2 - x_3 = 1 \\ & x_1 + x_4 = 2 \\ & x_2 + x_5 = 2 \end{cases}$$

Vamos analisar duas situações:

- a) $x_1 = x_2 = \frac{5}{4}$ (aproximadamente no meio da região factível)
  $z = 5x_1 + 3x_2 = 10$

$\tilde{z} \cong 7.237$

- b) $x_1 = 1.999,\ x_2 = 0.001$ (quase na fronteira)
  $z = 9.998$
  $\tilde{z} \cong 134.3$

$(P1)$ pode ser colocado como sendo:

$$(P2) : \begin{cases} max & f(x) = c^t x + \mu ln x \\ sa & Ax = b \end{cases}$$

e o problema de busca de melhor direção factível (em torno do ponto x^k) pode ser colocado como:

$$(P3) : \begin{cases} max & \Delta f(x^k, dx^k) \cong \nabla f^t(x^k) dx^k + \frac{1}{2}(dx^k)^t J(x^k) dx^k \\ sa & A dx^k = 0 \end{cases}$$

Note que:

$$\begin{cases} \nabla f(x^k) & = c + \mu \frac{1}{x^k} \\ J(x^k) & = \{a_{ij}\} \\ a_{ij} & = \begin{cases} -\frac{\mu}{(x_i^k)^2} & se\ i = j \\ 0 & se\ c.c. \end{cases} \end{cases}$$

Aplicando o método primal afim-escala ao $(P3)$, obtemos (do lagrangeano associado):

$dx^k = \frac{1}{\mu} X^k P^k (c^k + \mu e)$

onde $X^k = diag(x^k)$, $c^k = X^k c$ e P^k é a matriz de projeção $P^k = I - (A^k)^t[A^k(A^k)^t]^{-1}A^k$ com $A^k = AX^k$.

Próximo ponto interior:

$x^{k+1} = x^k + \alpha^k dx^k$, com $dx^k = X^k P^k(c^k + \mu e)$

Melhor tamanho do passo no método de Newton $= \frac{1}{\mu}$, $\Rightarrow$:

$\alpha^k = min\{\frac{1}{\mu}, \tau\lambda_{max}\}$, onde: $\lambda_{max} = min_{\delta_{x_i^k}<0}\{-\frac{x_i^k}{\delta_{x_i^k}}\}$

μ influencia fortemente no fator de convergência.

Valor típico: $\tau \in [.9, .999]$

7.2 Algoritmo

1. Dados $\tau \in (0,1)$, $x^0 > 0$, μ^0 grande, $k = 0$:
2. Faça até convergir:
   (a) $X^k = diag\{x^k\}$
   (b) $c^k = X^k c$
   (c) $A^k = AX^k$
   (d) $P^k = I - (A^k)^t[A^k(A^k)^t]^{-1}A^k$
   (e) $dx^k = X^k P^k(c^k + \mu^k e)$
   (f) $\lambda^k_{max} = min_{\delta_{x_i^k<0}}\{-\frac{x_i^k}{\delta_{x_i^k}}\}$

(g) $\alpha^k = min\{\frac{1}{\mu^k}, \tau\lambda^k_{max}\}$

(h) $x^{k+1} = x^k + \alpha^k dx^k$

(i) $\mu^{k+1} = f(\mu^k, \Delta c^t x^k)$

(j) $k \leftarrow k+1$

3. Fim Faça

Nota:

- a) Critério de convergências: pode ser feito sobre a variação do valor de $c^t x^k$ ($\Delta c^t x^k \leq \epsilon$) ou sobre o valor de dx^k ($\Delta dx^k \leq \epsilon$).
- b) O ponto inicial interior x^0 pode ser encontrado utilizando o método de M grande, como no caso do primal afim-escala.
- c) O valor de $\Delta\mu^k$ pode ser tal que:

$$\Delta\mu^k = \begin{cases} 0 & se \;\; \| \Delta c^t x^k \| \;\; grande \\ \beta \| \Delta c^t x^k \| & se \;\; \| \Delta c^t x^k \| \;\; pequeno \end{cases}$$

ou considerando:

$$\mu^{k+1} = \beta\mu^k, \quad 0 < \beta < 1$$

O valor de $\beta \in \mathbf{R}$, $\beta < \mathbf{1}$ define a velocidade de convergência.

8 Problemas com Variáveis Canalizadas

8.1 Introdução

Seja o problema:

$$(P1)\begin{cases} Min & c^t x \\ s.a & Ax = b \\ & l \leq x \leq u \end{cases}$$

Fazendo $x = \tilde{x} + l$:

$$(P2)\begin{cases} Min & c^t \tilde{x} + c^t l \\ s.a & A\tilde{x} = \tilde{b} \\ & 0 \leq \tilde{x} \leq \tilde{u} \end{cases}$$

onde $\tilde{b} = b - Al$ e $\tilde{u} = u - l$. Eliminando $c^t l$ e fazendo $(\tilde{x}, \tilde{b}, \tilde{u}) \rightarrow (x, b, u)$:

$$(P3)\begin{cases} Min & Z = c^t x \\ s.a & Ax = b \\ & 0 \leq x \leq u \end{cases}$$

que pode ser escrito como:

$$(P4)\begin{cases} Min & Z = c^t x + 0^t v \\ s.a & Ax + 0v = b \\ & Ix + Iv = u \\ & (x, v) \geq 0 \end{cases}$$

onde v é variável de folga. O problema dual fica:

$$(D1)\begin{cases} Max & \phi = b^t y + u^t \tilde{w} \\ s.a & A^t y + I\tilde{w} \leq c \\ & 0y + I\tilde{w} \leq 0 \end{cases}$$

Fazendo $w = -\tilde{w}$ e adicionando variável de folga z:

$$(D2)\begin{cases} Max & \phi = b^t y - u^t w \\ s.a & A^t y - w + z = c \\ & (z, w) \geq 0 \end{cases}$$

Condições de otimalidade:

$$Primal : \begin{cases} Ax = b \\ x + v = u \\ (x, v) \geq 0 \end{cases}$$

$$Dual : \begin{cases} A^t y - w + z = c \\ (z, w) \geq 0 \end{cases}$$

$$Folga\ Complementar : \begin{cases} XZe = 0 \\ VWe = 0 \end{cases}$$

Re-escrevendo:

$$F(x,y,z,w,v) = \begin{bmatrix} b - Ax \\ c - A^t y + w - z \\ u - x - v \\ \mu e - XZe \\ \mu e - VWe \end{bmatrix} = \begin{bmatrix} r_p \\ r_d \\ r_u \\ r_c \\ r_b \end{bmatrix} = r$$

Série de Taylor em torno de $(x^0, y^0, z^0, w^0, v^0)$:

$F(x,y,z,w,v) \cong F(x^0, y^0, z^0, w^0, v^0) +$
$J(x^0, y^0, z^0, w^0, v^0)[(x,y,z,w,v) - (x^0, y^0, z^0, w^0, v^0)]$

$-Jd = r$

ou seja:

$$\begin{bmatrix} A & 0 & 0 & 0 & 0 \\ 0 & A^t & I & -I & 0 \\ I & 0 & 0 & 0 & I \\ Z & 0 & X & 0 & 0 \\ A & 0 & 0 & V & W \end{bmatrix} \begin{bmatrix} d_x \\ d_y \\ d_z \\ d_w \\ d_v \end{bmatrix} = \begin{bmatrix} r_p \\ r_d \\ r_u \\ r_c \\ r_b \end{bmatrix}$$

O sistema fica:

$$
\begin{cases}
(1) \;\; Ad_x = r_p \\
(2) \;\; A^t d_y - d_w + d_z = r_d \\
(3) \;\; d_x + d_v = r_u \\
(4) \;\; Zd_x + Xd_z = r_c \\
(5) \;\; Vd_w + Wd_v = r_b
\end{cases}
$$

De (2) e (5):

$$(6) \;\; A^t d_y + d_z + V^{-1}Wd_v = r_d + V^{-1}r_b$$

De (3), (4) e (6):

$$A^t d_y - (X^{-1}Z + V^{-1}W)d_w = \\ r_d + V^{-1}r_b - X^{-1}r_c - V^{-1}Wr_u$$

Definindo $D = X^{-1}Z + V^{-1}W$:

$$d_x = D^{-1}\left[A^t d_y - r_d - V^{-1}r_b + X^{-1}r_c + V^{-1}Wr_u\right]$$

De (1):

$$AD^{-1}A^t d_y = r_p + AD^{-1}[r_d + V^{-1}r_b - X^{-1}r_c - V^{-1}Wr_u]$$

8.2 Algoritmo

1. Dados $(x^0, v^0, y^0, z^0, w^0)$ tal que $(x^0, v^0, z^0, w^0) > 0$, $\tau \in (0,1)$, $\sigma \in (0,1)$ e $k = 0$,

2. Faça até convergência:

   (a) $\gamma^k = (x^k)^t z^k + (v^k)^t w^k$

   (b) $\mu^k = \frac{\sigma\gamma^k}{n_p}$

   (c) $r_p^k = b - Ax^k$

   (d) $r_d^k = c - A^t y^k - z^k + w^k$

   (e) $r_u = u - x^k - v^k$

   (f) $r_c^k = \mu^k e - X^k Z^k e$

   (g) $r_b^k = \mu^k e - V^k W^k e$

   (h) $D^k = [(X^k)^{-1} Z^k]$

   (i) $dy^k = [A(D^k)^{-1}A^t]^{-1}[r_p^k + A(D^k)^{-1}(r_d^k - (X^k)^{-1}r_c^k +$

   (j) $\qquad (V^k)^{-1}r_b^k - (V^k)^{-1}W^k r_u^k)]$

   (k) $dx^k = (D^k)^{-1}[A^t dy^k - r_d^k + (X^k)^{-1}r_c^k - (V^k)^{-1}r_b^k +$

   (l) $\qquad (V^k)^{-1}W^k r_u^k]$

   (m) $dz^k = (X^k)^{-1}[r_c^k - Z^k dx^k]$

   (n) $d_v^k = r_u^k - d_x^k$

   (o) $d_w^k = (V^k)^{-1}(r_b^k - W^k d_v^k)$

   (p) $\rho_p^k = min\{min_{\delta x_i^k < 0}\{-\frac{x_i^k}{\delta x_i^k}\}, min_{\delta v_i^k < 0}\{-\frac{v_i^k}{\delta v_i^k}\}\}$

   (q) $\rho_d^k = min\{min_{\delta z_i^k < 0}\{-\frac{z_i^k}{\delta z_i^k}\}, min_{\delta w_i^k < 0}\{-\frac{w_i^k}{\delta w_i^k}\}\}$

(r) $\alpha_p^k = min(\tau\rho_p^k, 1)$
(s) $\alpha_d^k = min(\tau\rho_d^k, 1)$
(t) $x^{k+1} = x^k + \alpha_p^k dx^k$
(u) $v^{k+1} = v^k + \alpha_p^k dv^k$
(v) $y^{k+1} = y^k + \alpha_d^k dy^k$
(w) $z^{k+1} = z^k + \alpha_d^k dz^k$
(x) $w^{k+1} = w^k + \alpha_p^k dw^k$
(y) $k \leftarrow k + 1$

3. Fim Faça

Nota:

- O critério de convergência pode ser:
  - a) $\parallel r^k \parallel < \epsilon$
  - b) $\parallel X^k Z^k \parallel < \epsilon$ e $\parallel V^k W^k \parallel < \epsilon$
- Pode-se inicializar com $(x^0, v^0, y^0, z^0, w^0)$ tal que:
  - $\tilde{x} = A^t(AA^t)^{-1}b$
  - $\tilde{v} = u - \tilde{x}$
  - $\epsilon_2 > 0$
  - $\epsilon_1 = max\{-min_i\ \tilde{x}_i, -min_i\ \tilde{v}_i, \epsilon_2, \frac{\|b+u\|_1}{\epsilon_2\|A\|_1}\}$
  - $x_i^0 = max\{max_i\{\tilde{x}_i\}, \epsilon_1\}$
  - $v_i^0 = max\{max_i\{\tilde{v}_i\}, \epsilon_1\}$

- $y_i^0 = 0$
- Considerando $c = [\gamma_1 \ \gamma_2 ... \gamma_n]$:
  * a) se $\gamma_i = 0$, fazer $z_i^0 = w_i^0 = \epsilon_2 > 0$
  * b) se $\gamma_i > 0$, fazer $z_i^0 = 2\gamma_i$ e $w_i^0 = \gamma_i$
  * c) se $\gamma_i < 0$, fazer $w_i^0 = -2\gamma_i$ e $z_i^0 = -\gamma_i$

- Se todas as variáveis forem canalizadas, $n_p = 2n$ (n variáveis x e n variáveis z).
- As adaptações para métodos primal-afim-escala e dual-afim-escala são semelhantes ao visto para primal-dual-clássico.

9 Comentários sobre Sistemas Lineares

Nos métodos de pontos interiores, são precisos determinar matrizes do tipo: $B = AD^{-1}A^t$, onde A é uma matriz dada e $D > 0$ é uma matriz diagonal que varia a cada iteração e é tal que:

- Primal Afim-Escala: $D = X^{-2} \quad (y^k = [A(X^k)^2A^t]^{-1}A(X^k)^2c)$
- Dual Afim-Escala: $D = Z^2 \quad (dy^k = [A(Z^k)^{-2}A^t]b)$
- Primal-Dual Afim-Escala:
  $D = X^{-1}Z$
  $(dy^k = [A(D^k)^{-1}A^t]^{-1}[r_p^k + A(D^k)^{-1}r_d^k - A(Z^k)^{-1}r_c^k])$

É uma etapa computacionalmente "cara". A seguir, vamos analisar alguns aspectos que poderão melhorar o comportamento computacional.

Como $D > 0$ é diagonal, $D^t > 0$ também é diagonal e podemos escrever:

$$B = AD^{-1}A^t = \tilde{A}\tilde{A}^t > 0 \text{ com } \tilde{A} = AD^{-\frac{1}{2}}$$

Assim, o elemento β_{ij} de B é dado por (ver figura 6):

$$\beta_{ij} = (\tilde{a}_i)^t(\tilde{a}_j) = \Sigma_{k=1}^n \tilde{a}_{ik}\tilde{a}_{jk} = \Sigma_{k=1}^n a_{ik}a_{jk}\delta_{kk}^{-1}$$

onde $\{\delta_{ij}\} = D$. Para cada β_{ij}, temos $[(2$ produtos $+$ uma soma$)n] = 3n$ operações. Como $a_{ik}a_{jk}$ é constante para todas as

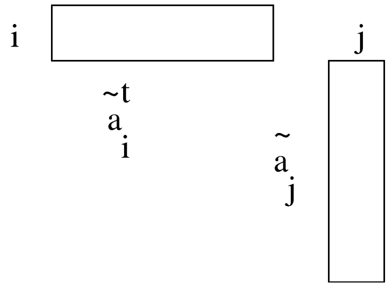

Figure 6: $(\tilde{a}_i^t)(\tilde{a}_j)$

iterações, se armazenarmos este produto, o cálculo de β_{ij} vai exigir $[(1 \text{ produto} + 1 \text{ soma})n] = 2n$ operações.

Existem $\frac{m(m+1)}{2}$ elementos diferentes em B (simétrica, portanto, $m + (m-1) + ... + 1 = \frac{m+1}{2}m$). Assim, o número de operações são:

Sem armazenamento:	$3n\frac{m(m+1)}{2}$	operações
Com armazenamento:	$nm(m+1)$	operações

A matriz B é simétrica e definida positiva. Logo, a resolução do sistema:

$$Bd = b$$

fica facilitada, se decompormos na forma:

$$LUd = b$$

onde: $\begin{cases} L = & matriz\ triangular\ inferior \\ U = & matriz\ triangular\ superior \end{cases}$

A resolução é processada, fazendo:

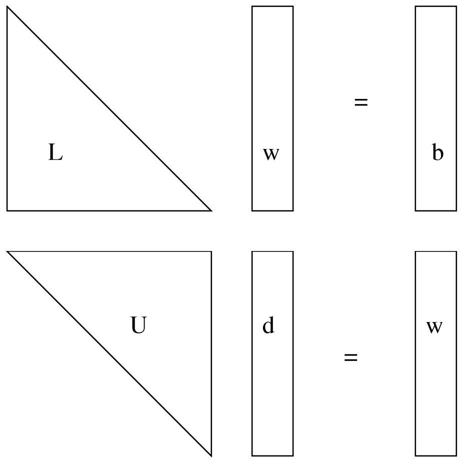

Figure 7: Decomposição LU - Solução por substituição

$$L(Ud) = b \quad \begin{cases} Lw = b \\ Ud = w \end{cases}$$

A decomposição LU não é única mas a decomposição $B = L\Delta U$ é única (onde Δ é diagonal e L e U são matrizes unitários. É possível mostrar que, para matrizes simétricas, $U = L^t$, ou seja, $B = L\Delta L^t$. Como $B > 0 \Rightarrow \Delta > 0$, podemos escrever que $\Delta = \Delta^{\frac{1}{2}}\Delta^{\frac{1}{2}}$, o que permite escrever:

$$B = L\Delta^{\frac{1}{2}}\Delta^{\frac{1}{2}}L^t = \tilde{L}\tilde{L}^t$$

que é conhecido como decomposição de Cholesky. Isso permite escrevermos:

$$Bd = b \Rightarrow \tilde{L}\tilde{L}^t d = b \Rightarrow \begin{cases} \tilde{L}w = b \\ \tilde{L}^t d = w \end{cases}$$

FEEC, 13 de outubro de 2004.

Akebo Yamakami
DT-FEEC-UNICAMP