Métodos de pontos interiores para PL

Motivação histórica: complexidade.

- método simplex (anos 1940-50): não polinomial
- " dos elipsóides (Khachiyan, 1979)
  - ↳ primeiro método polinomial ($\Rightarrow$ PL é polin.)
  - ↳ numericamente muito ruim!
  - ↳ não é tipo pontos interiores...
- algoritmo de Karmakar (1984)
  - ↳ 1° pontos interiores polinomial que "funciona".

- método primal afim escala (Dikin, 1967) $\lfloor 2$
  - 1º pontos interiores
  - não polinomial :(
  - atua no problema primal
  - não é bom numericamente...
- método dual afim escala
  - parecido com o primal, mas atua no problema dual
  - funciona bem, compete com o simplex!

- método primal dual afim escala
  - ↳ baseia-se na aplicação do método de Newton ao sistema das condições de otimalidade (viabilidade primal + dual + complementaridade; ou sistema KKT)
  - ↳ Isto é, atualiza variáveis primais e duais (multiplicadores de Lagrange) __ao mesmo tempo__ (daí "primal dual")

- ↳ é polinomial
- ↳ <u>não</u> funciona tão bem ...

- método primal dual seguidor de caminhos
  - ↳ Newton ...
  - ↳ corrige defeitos do primal dual afim escala (é + estável numericamente)
  - ↳ é polinomial, funciona bem.

- método preditor–corretor (Mehrotra, 1989)
  - é "primal dual"
  - combina a direção afim (direção "preditora") com uma direção "corretora".
  - é polinomial e funciona muito bem!
  - é a base das implementações modernas (CPLEX, GUROBI etc).
  - adequado à PL's grandes.

PRELIMINARES

(1) Condições de otimalidade para PL.

PL na forma padrão:

$$P: \min_{x} \; c^t x$$
$$\text{s.a. } Ax = b$$
$$x \geq 0$$

A $m \times n$

posto $A = m \leq n$

Lembre-se que <u>todo</u> PL pode ser escrito na forma padrão.

Problema dual:

$$\max_{y} \; b^t y \quad \text{s.a.} \; A^t y \leq c \quad \xrightarrow{+\text{folga}} \quad \boxed{\begin{array}{l} \max_{y,z} \; b^t y \\ \text{s.a.} \; A^t y + z = c \\ \qquad z \geq 0 \end{array}}$$

D:

As condições de otimalidade vêm da teoria de dualidade em PL: <u>x é ótimo para P</u> se, <u>e só se</u>, existir (y, z) tal que

$Ax = b$, $x \geq 0$ (viabilidade primal)

$A^t y + z = c$, $z \geq 0$ (viabilidade dual)

$x_i z_i = 0$, $\forall i$ (complementaridade).

Essas condições são necessárias e suficientes para que x seja ótimo. (disciplina P.O.)

↳ Outra visão: são as condições KKT, suficientes dado que P é convexo (Otimização 1)

Tomando

$$X = \mathrm{diag}(x) = \begin{bmatrix} x_1 & & 0 \\ & \ddots & \\ 0 & & x_n \end{bmatrix}, \quad Z = \mathrm{diag}(z)$$

e $e = [1 \cdots 1]^t$, podemos escrever

$$\boxed{\begin{array}{l} Ax = b \,, \quad x \geq 0 \\ A^t y + z = c \,, \quad z \geq 0 \\ XZe = 0 \end{array}}$$

(verifique!)

② Método de Newton (Otimização 1).

Sistema (não linear) $F(w) = 0$.

Ponto corrente: w^k.

Próximo ponto: $w^{k+1} = w^k + d^k$, onde a direção d^k é solução do sistema linear

$$F'(w^k)\, d = -F(w^k)$$

MÉTODOS DE PONTOS INTERIORES

Um ponto é <u>interior</u> quando todas as variáveis encontram-se estritamente dentro de seus limites ($x_i > 0$, $z_i > 0$ $\forall i$)

<u>Pergunta</u>: se x é primal viável e interior ($Ax = b$, $x > 0$), qual seria uma boa estimativa para y e z?

A "parte dual" das condições de otimalidade são

$$A^t y + z = c \quad , \quad XZe = 0$$

(vamos "esquecer" $z \geq 0$). Note que

$$Ze = z.$$

Uma boa estimativa para (y, z) é aquela que minimiza $\|XZe\|_2$:

$$\min_{z,y} \frac{1}{2}\|Xz\|_2^2 \quad s.a. \quad A^t y + z = c,$$

que equivale à

$$\min_{y} \tfrac{1}{2} \| X(c - A^t y) \|_2^2$$

Resolvendo:

$$\nabla \left\{ \tfrac{1}{2} \| X(c - A^t y) \|_2^2 \right\} = 0$$

$$\Rightarrow -AX^t X(c - A^t y) = 0$$

$$\Rightarrow AXXA^t y = AXXc \Rightarrow AX^2A^t y = AX^2 c$$

Como posto $A = m \leq n$ (A $m \times n$) e $x_i > 0$, a matriz $m\times m$ simétrica AX^2A^t é inversível. Logo

$$\boxed{y = (AX^2A^t)^{-1}AX^2c}.$$

e

$$\boxed{z = c - A^t y}$$

(estimativas de Dikin). Observe que <u>não</u> há garantia que $z > 0$...

Estamos supondo x primal viável e interior, isto é,

$$Ax = b, \quad x > 0.$$

Como atualizar x mantendo viabilidade primal?

Direção $d = -X^2 z$: $x \leftarrow x + d$.

(i) $A(x+d) = Ax + Ad = b - AX^2 z$

$$= b + AX^2A^t y - AX^2 c$$

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

$$\underbrace{= b. \quad \text{Ou seja,}}\quad A(x+d) = b.$$

(ii) $0 < \|X^{-1}d\|_2^2 = d^tX^{-2}d = z^tX^2X^{-2}X^2z$

$$= z^tX^2z = (c - A^ty)^tX^2z$$

$$= c^tX^2z - y^tAX^2z$$

$$= -c^td - y^tAX^2(c - A^ty)$$

$$= -c^td - y^t(\underbrace{AX^2c - AX^2A^ty}_{0}) = -c^td$$

Assim, $c^t(x+d) = c^t x + c^t d < c^t x$.

Ou seja, $d = -X^2 z$ é uma direção de <u>descida</u> e <u>primal viável</u>.

$\Rightarrow$ damos um passo $x^{k+1} = x^k + \alpha_k d^k$, $\alpha_k > 0$.

Queremos que $x^{k+1} > 0$ ...

- se $d_i^k \geq 0$ então $x_i^{k+1} = x_i^k + \alpha_k d_i^k > 0$

  $\forall \alpha_k > 0$ (lembre-se que $x_i^k > 0$).

- se $d_i^k < 0$ então

$$x_i^{k+1} > 0 \iff \alpha_k < -\frac{x_i^k}{d_i^k}.$$

Tomamos

$$\alpha_k = \tau \min_i \left\{ -\frac{x_i^k}{d_i^k} \, ; \, d_i^k < 0 \right\},$$

onde $\tau \in (0,1)$ é um parâmetro.

Obs: $\tau = 2/3$ garante convergência teórica, mas $\tau = 0{,}9$ ou $\tau = 0{,}99$ é melhor na prática.

Método primal afim escala (Dikin)

Dado x^0 primal viável interior.

para $k = 0, \ldots, maxit$

$$y^k = (A X_k^2 A^t)^{-1} A X_k^2 c$$

$$z^k = c - A^t y^k$$

$$d^k = - X_k^2 z^k$$

$$\alpha_k = \tau \min_i \left\{ -\frac{x_i^k}{d_i^k} ; d_i^k < 0 \right\}$$

$$x^{k+1} = x^k + \alpha_k d^k$$

até "convergir".

* O passo caro é resolver

$$(AX_k^2A^t)y = AX_k^2c.$$

$AX_k^2A^t$ é simétrica e definida positiva

$\Rightarrow$ usar Cholesky.

* Este método não é bom na prática.
Pois acumula muitos erros de arredondamento. Neste caso, podemos ter $Ax^* \neq b$.

Ponto inicial

Precisamos iniciar o método com x^0 tal que $Ax^0 = b$ e $x^0 > 0$... (FASE 1)

1) escolha qualquer $\tilde{x}^0 > 0$. Não há garantia que $A\tilde{x}^0 = b$...

2) $\rho = b - A\tilde{x}^0$.

3) Considere o problema

$$\begin{aligned} \min_{u,\tilde{x}} \quad & -u \\ \text{s.a.} \quad & A\tilde{x} + u\rho = b \\ & \tilde{x} \geq 0, \; u \geq 0 \end{aligned}$$

Observe que $(\tilde{x}, u) = (\tilde{x}^0, 1) > 0$ é viável! Portanto podemos aplicar o método primal afim escala.

4) Aplique o método, obtendo $(\tilde{x}^*, u^*)$:

- $u^* > 0 \Rightarrow$ P inviável
- u^* "$=$" 0: neste caso $A\tilde{x}^* = b$, $\tilde{x}^* \geq 0$.

"Acreditando" que $\tilde{x}^* > 0$, iniciamos o método no problema original com $x^0 = \tilde{x}^*$.

Convergência do método primal afim escala

Teorema: • (Vanderbei, 1986) Se P e D forem não degenerados, então para qualquer $\sigma < 1$ o método converge à um ótimo de P.

• (Tsuchiya e Muramatsu, 1992) Em geral, há convergência para $\sigma \leq 2/3$.

• (Mascarenhas, 1997) Existe um exemplo e um $\sigma < 1$ para os quais o método converge a um ponto não ótimo. (neste exemplo, $\sigma > 0{,}995$; para $2/3 < \sigma \leq 0{,}995$, ninguém sabe!).