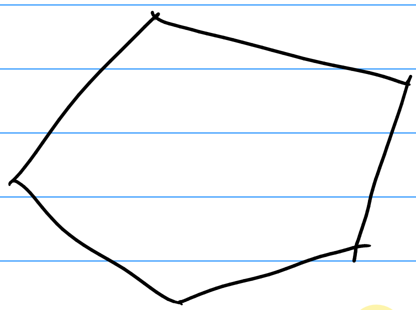
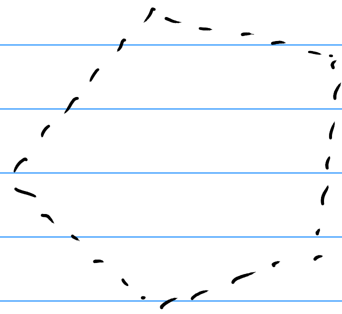
Penalização interna

- Martínez, J. M.; Santos, S. A. Métodos computacionais de otimização

P: $\min f(x)$ s.a. $g(x) \geq 0$, $x \in D$

$D \subset \mathbb{R}^n$ é compacto.





$\Omega = \{x \in D \,;\, g(x) \geq 0\}$

$\Omega^\circ = \{x \in D \,;\, g(x) > 0\}$

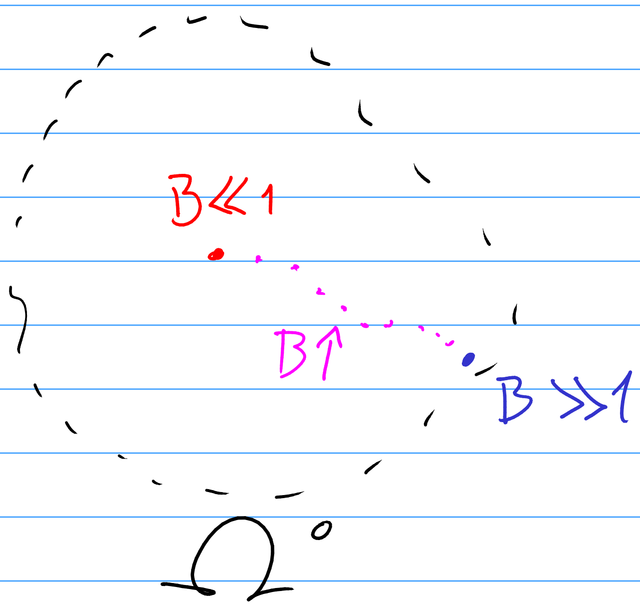
Hipótese: $\Omega^\circ \neq \emptyset$.

Ideia: resolver uma sequência de subproblemas onde a "viabilidade estrita" é penalizada.

Subproblema:

$$SP(t_k): \min_x f(x) + t_k B(x) \qquad (t_k > 0).$$
$$\text{s.a. } x \in \Omega^\circ$$

B: medida de "viabilidade estrita"



B usuais:

1) $B(x) = \sum_{i=1}^{m} \frac{1}{g_i(x)}$ (barreira inversa)

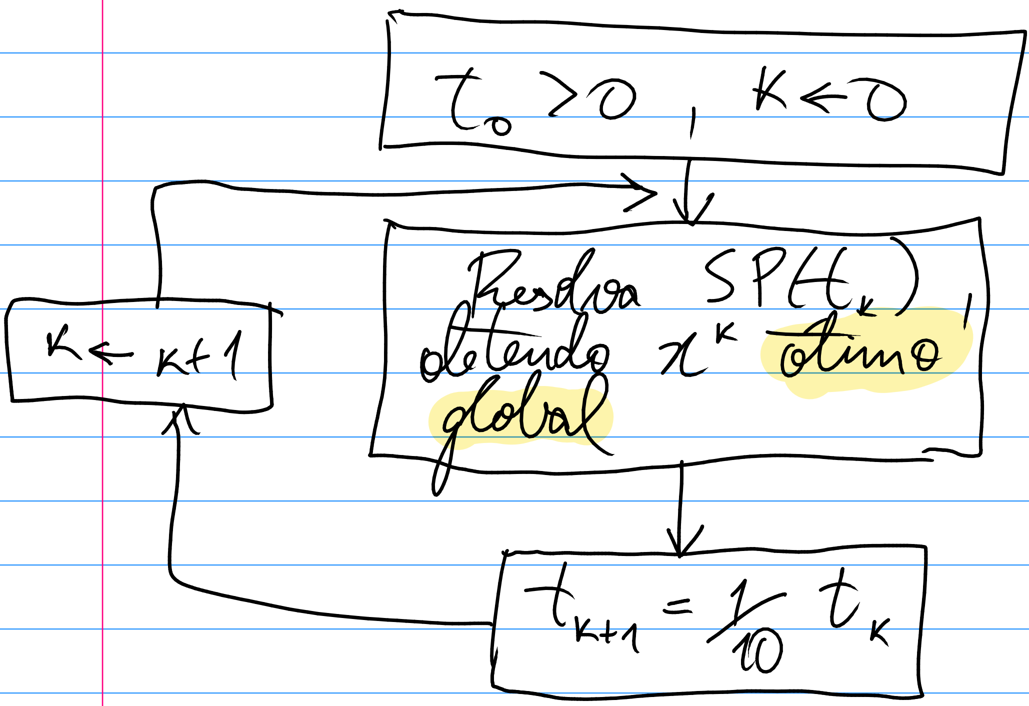
2) $B(x) = -\sum_{i=1}^{m} \log(g_i(x))$ (barreira logarítmica)

- Estão definidas em Ω° ($g(x) > 0$) mas não em Ω ($g(x) \geq 0$)
- $B(x) \geq 0$, $\forall x \in \Omega^{\circ}$.

↳ Para teoria, $B(x) \leftarrow -\sum_{i=1}^{m} \log(g_i(x)) + M$ onde $M \gg 1$. Assim, $B(x) \geq 0$, $\forall x \in \Omega^{\circ}$ pois $\Omega^{\circ} \subset D$ (limitado).

- Se $\lim_{x \to x^*} g_i(x) = 0$ para algum i, então $\lim_{x \to x^*} B(x) = \infty$.

O esquema:

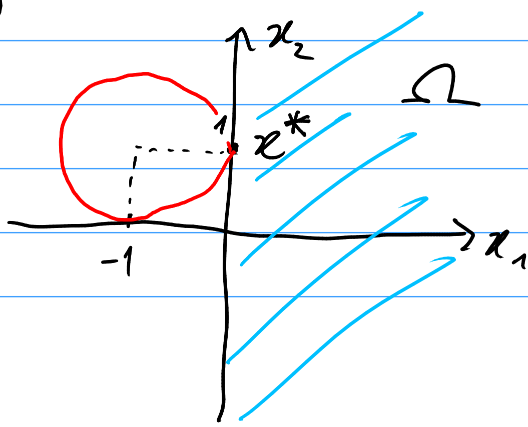


Note que $\lim_{k \to \infty} t_k = 0$.

Exemplo: $\min_x \; (x_1+1)^2 + (x_2-1)^2$

s.a. $x_1 \geq 0$

$x^* = (0,1)$.



Subproblema:

$$SP(t_k): \min_{x} (x_1+1)^2 + (x_2-1)^2 - t_k \ln(x_1)$$

$$s.a. \ x_1 > 0$$

Resolvendo: vamos desconsiderar por enquanto a restrição $x_1 > 0$.

$$\begin{bmatrix} 2(x_1^k+1) - t_k \frac{1}{x_1^k} \\ 2(x_2^k-1) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow x_2^k = 1 \text{ e}$$

$2(x_1^k+1) = \frac{t_k}{x_1^k}$. Daí, $2(x_1^k)^2 + 2x_1^k - t_k = 0$

$$\Rightarrow x_1^k = \frac{-2+\sqrt{4+8t_k}}{4} = \frac{1}{2}\left(\sqrt{1+2t_k} - 1\right) > 0$$

ou $x_1^k = \frac{-2-\sqrt{4+8t_k}}{4} < 0$ (não serve)

Dado que a F.O. de $SP(t_k)$ é convexa, sua solução é

$$x^k = \left(\frac{1}{2}\left(\sqrt{1+2t_k} - 1\right), 1\right).$$

Note que $x^k \to x^* = (0,1)$ //

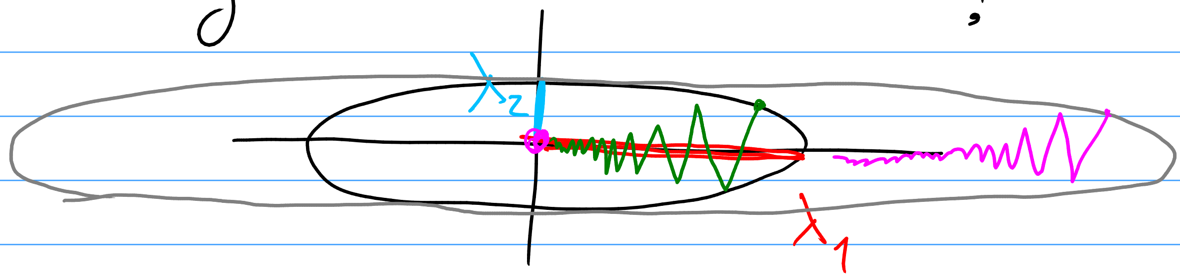
Obs:

1) Assim como a penalização externa, este esquema é instável numericamente.

$\hookrightarrow$ No exemplo acima, a hessiana da F.O. do $SP(t_k)$ é $\begin{bmatrix} 2 + t_k/(x_1^k)^2 & 0 \\ 0 & 2 \end{bmatrix}$,

cujos autovalores são $\lambda_1^k = 2 + \frac{t_k}{(x_1^k)^2}$ e $\lambda_2^k = 2$. Como $2(x_1^k + 1) - t_k/x_1^k \to 0$ e $x_1^k \to 0^+$, temos $\frac{t_k}{x_1^k} \to 2$, e logo $\frac{t_k}{(x_1^k)^2} = \frac{t_k}{x_1^k} \cdot \frac{1}{x_1^k} \to \infty$

$\Rightarrow \lambda_1^k \to \infty \Rightarrow$ Hessiana da F.O. de $SP(t_k)$ fica cada vez mais mal condicionada!



2) x^k, na prática, é apenas um ponto estacionário do subproblema, e não um ótimo global.

Teorema: Seja $\{x^k\}$ a sequência gerada pelo esquema de penalização interna. Então todo ponto de acumulação x^* de $\{x^k\}$ é minimizador global de P.

Prova: Como $0 < t_{k+1} < t_k$ e $B(x) \geq 0$, $\forall x \in \Omega^{\circ}$, então, pela minimalidade de x^{k+1} em $SP(t_{k+1})$,

$$f(x^{k+1}) + t_{k+1} B(x^{k+1}) \leq f(x^k) + t_{k+1} B(x^k)$$

$$\leq f(x^k) + t_k B(x^k), \quad \forall k. \qquad (1)$$

Definimos $b_k = f(x^k) + t_k B(x^k)$ o valor ótimo de $SP(t_k)$. Por (1), a sequência $\{b_k\}$ é monótona não crescente. Também, como $B(x^k) \geq 0$,

$$b_k \geq f(x^k) \geq f^*, \quad \forall k \qquad (2)$$

onde f^* é o valor ótimo de P. Ou seja, $\{b_k\}$ é limitada inferiormente. Assim, existe o limite

$\bar{b} = \lim_{k\to\infty} b_k$. Afirmamos que $\bar{b} = f^*$. Note que de (2) temos $\bar{b} \geq f^*$. Suponha então por absurdo que $\bar{b} > f^*$ e seja $\bar{x}$ um minimizador global de P (isto é, $\bar{x}$ é viável e $f(\bar{x}) = f^*$). Tomemos $x' \in V \cap \Omega^\circ$, onde V é uma vizinhança de $\bar{x}$, satisfazendo

$$f(x') < \frac{1}{2} f^* + \frac{1}{2} \bar{b} \qquad (3)$$

(lembre-se que estamos supondo $\bar{b} > f^*$ por absurdo). Como $t_k \to 0^+$, para este x' fixo temos

$$0 \leq t_k B(x') \to 0.$$

Assim, como $\bar{b} - f^* > 0$, temos

$$0 \leq t_k B(x') \leq \frac{1}{2}(\bar{b} - f^*), \quad \forall k \gg 1. \qquad (4)$$

Somando (3) e (4) obtemos

$$b_k \leq f(x') + t_k B(x') < \bar{b}, \quad \forall k \gg 1.$$

Assim,

$$\dots \leq b_{k+2} \leq b_{k+1} \leq b_k < \bar{b}, \quad \forall k \gg 1,$$

o que implica $\bar{b} > \lim_{k\to\infty} b_k$, um absurdo.

Concluímos portanto que $\bar{b} = f^*$.

Consideramos agora x^* ponto de acumulação da sequência viável $\{x^k\}$ gerada, digamos $\lim_{k\in K} x^k = x^*$.

Temos $0 = \lim_{k\in K} (b_k - \bar{b}) = \lim_{k\in K} [f(x^k) + t_k B(x^k) - f^*]$

$= \lim_{k\in K} [\underbrace{f(x^k) - f^*}_{\geq 0} + \underbrace{t_k B(x^k)}_{\geq 0}] \Rightarrow f(x^*) = \lim_{k\in K} f(x^*) = f^*$,

ou seja, x^* é minimizador global de P. ■

Pacote computacional: IPOPT.

(Interior Point OPTimizer)

- Penalização interna / métodos de barreiras
- Pontos interiores: quando no subproblema é aplicado método de Newton.

IPOPT: pontos interiores usando barreira logarítmica.

↳ combinação mais utilizada.

Caso particular: problemas lineares / quadráticos

Pontos interiores com barreira logarítmica são implementadas para PL's, p. ex., no CPLEX. ("barrier algorithm").

PL: $\min c^t x$

s.a. $Ax = b$

$x \geq 0$

Queremos aplicar a técnica de penalização interna à PL.

$$\begin{array}{l} \min c^t x \\ \text{s.a. } Ax = b \end{array} \overset{KKT}{\longleftrightarrow} \boxed{\begin{array}{c} c + A^t \lambda = 0 \\ Ax = b \end{array}} \quad (*)$$

(prob. convexo)

(KKT é um sistema linear)

(*) é suficiente para otimalidade

$$\begin{array}{ll} \min & c^t x \\ \text{s.a.} & Ax = b \\ & x \geq 0 \\ & (-x \leq 0) \end{array} \quad \overset{KKT}{\longleftrightarrow} \quad \boxed{\begin{array}{c} c + A^t\lambda - \mu = 0 \\ Ax = b, \quad x \geq 0 \\ \mu_i x_i = 0, \ \forall i \end{array}}$$

(KKT tem expressões não lineares)

<u>Conclusão</u>: quem atrapalha resolver PL's são as restrições de desigualdade!

↳ No nosso caso, $x \geq 0$.

Penalizamos <u>apenas</u> $x \geq 0$:

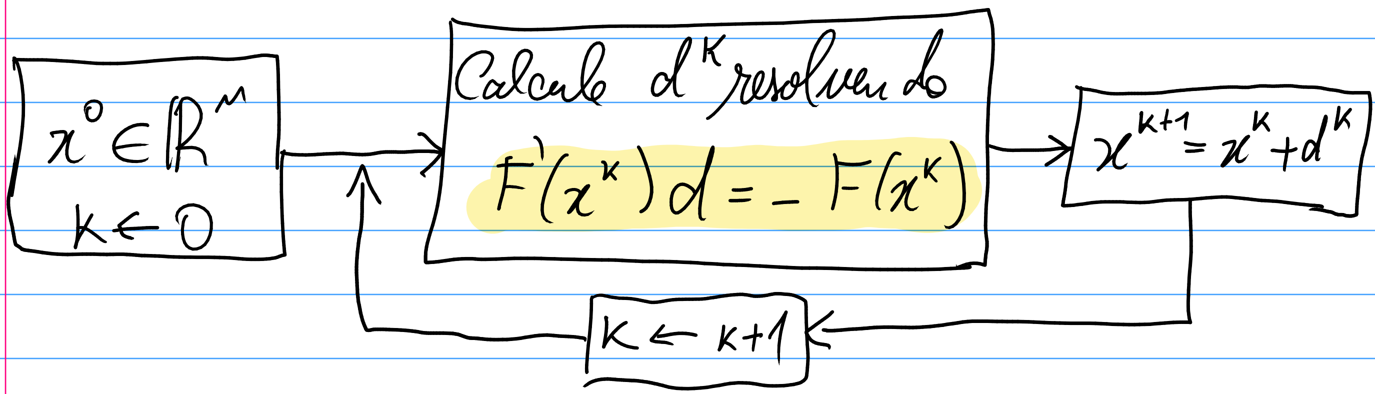
$$SP(t_k): \min_x \ c^t x - t_k \sum_{i=1}^{n} \ln(x_i)$$

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

<u>Obs</u>: em $SP(t_k)$, devemos ter $x_i > 0, \forall i$ pois ln só é definido em valores > 0.

Vamos aplicar passos do método de Newton à $SP(t_k)$.

Método de Newton para sistemas não lineares $F(x)=0$:



$$f: \mathbb{R}^n \to \mathbb{R}$$

$$f(x^k+d) = f(x^k) + \nabla f(x^k)^t d + r(d)$$

$\hookrightarrow \dfrac{r(d)}{\|d\|} \xrightarrow[d \to 0]{} 0$

Para $d \approx 0$,

$$f(x^k+d) \approx f(x^k) + \underbrace{\nabla f(x^k)^t}_{\text{derivada}} d.$$

Newton: $\nabla f(x^k)^t d + f(x^k) \underset{=}{\approx} 0.$

$$\boxed{F'(x^k)\, d = -F(x^k)}$$

$$SP(t_k): \quad \min_x \; c^t x - t_k \sum_{i=1}^{n} \ln(x_i) \qquad (x > 0)$$
$$\text{s.a.} \quad Ax = b$$

KKT:

$$c - t_k \begin{bmatrix} 1/x_1 \\ 1/x_2 \\ \vdots \\ 1/x_n \end{bmatrix} + A^t \lambda = 0$$

$$Ax - b = 0$$

$$F(x,\lambda) = \begin{bmatrix} c - t_k \begin{bmatrix} 1/x_1 \\ \vdots \\ 1/x_n \end{bmatrix} + A^t \lambda \\ Ax - b \end{bmatrix} = 0$$

Direção de Newton para $F(x,\lambda)$:

$$F'(x,\lambda)\, d = -F(x,\lambda)$$

$$F'(x,\lambda) = \begin{bmatrix} t_k \begin{bmatrix} 1/x_1^2 & 0 & \cdots & 0 \\ 0 & 1/x_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/x_n^2 \end{bmatrix} & A^t \\ A & 0 \end{bmatrix}$$

Vamos denotar

$$X = \operatorname{diag}(x_1, \dots, x_n) = \begin{bmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_n \end{bmatrix}$$

Se $x_i > 0, \forall i$, então

$$X^{-1} = \begin{bmatrix} 1/x_1 & & 0 \\ & \ddots & \\ 0 & & 1/x_n \end{bmatrix}$$

$$X^{-1} X^{-1} = X^{-2} = \begin{bmatrix} 1/x_1^2 & & 0 \\ & \ddots & \\ 0 & & 1/x_n^2 \end{bmatrix}.$$

Assim,

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

Sistema Newtoniano:

$$\begin{bmatrix} t_k X^{-2} & A^t \\ A & 0 \end{bmatrix} \begin{bmatrix} d_x \\ d_\lambda \end{bmatrix} = - \begin{bmatrix} c - t_k \begin{bmatrix} 1/x_1 \\ \vdots \\ 1/x_n \end{bmatrix} + A^t \lambda \\ Ax - b \end{bmatrix}$$

Poderíamos resolver este sistema, obtendo a direção Newtoniana d^k, e dar o passo $x^{k+1} = x^k + d^k$.

Problema: devemos garantir que $x^{k+1} > 0$!