Otimização: Quadrados Mínimos Não-Lineares e Otimização Global

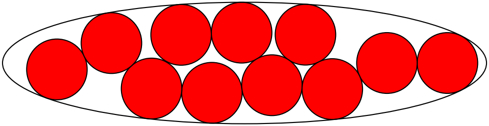
Hector Flores Callisaya

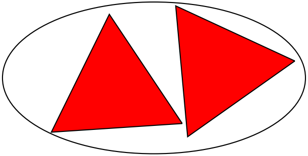
Motivação: Problemas de empacotamento

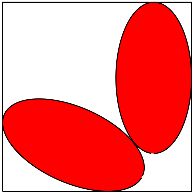
- Empacotamento de círculos em círculos,
- Empacotamento de politopos em regiões convexas,
- Empacotamento de elipses em politopos,
- Itens com bordas suave em politopos

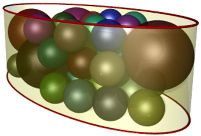
$$\begin{cases} h_i(x) = 0, & \forall i = 1, \ldots, m, \\ g_j(x) \leq 0, & \forall j = 1, \ldots, p, \\ x \in \Omega, \end{cases}$$

onde as funções $h_i, g_j : \mathbb{R}^n \to \mathbb{R}$ são suaves para todo $i = 1, \ldots, m$ e $j = 1, \ldots, p$.









$$\underset{x \in \mathbb{R}^n}{\text{Minimizar}} \ \sum_{i=1}^{m} h_i(x)^2 + \sum_{i=1}^{l} \max\{0, g_i(x)\}^2,$$
$$\text{sujeito a } x \in \Omega,$$

onde $m \geq n$

- Ω é um conjunto onde é fácil calcular a projeção P_Ω.
- O problema é limitado inferiormente por 0.

$$\sum_{i=1}^{m} h_i(x^*)^2 + \sum_{i=1}^{l} \max\{0, g_i(x^*)\}^2 = 0 \iff h_i(x^*) = 0, g_i(x^*) \leq 0,$$

$\boxed{\textbf{É um problema de Otimização Global}}$

$$\begin{aligned} \underset{x\in\mathbb{R}^n}{\text{Minimizar}} \quad & f(x) \\ \text{sujeito a} \quad & h_i(x) = 0,\ i = 1, \ldots, m. \\ & g_i(x) \leq 0,\ i = 1, \ldots, l. \\ & x \in \Omega. \end{aligned}$$

Consideramos o Lagrangiano Aumentado definido por Powell, Hesstenes e Rockafellar

$$L_\rho(x, \lambda, \mu) = f(x) + \frac{\rho}{2} \left\| h(x) + \frac{\lambda}{\rho} \right\|_2^2 + \left\| (g(x) + \frac{\mu}{\rho})_+ \right\|_2^2.$$

para $\rho,\ \lambda \in \mathbb{R}^m,\ \mu \in \mathbb{R}_+^p,\ x \in \Omega.$

$$\nabla_{(x,\lambda,\mu)} L_\rho(x, \lambda, \mu) = 0.$$

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{Minimizar}} & f(x) \\ \text{sujeito a} & h_i(x) = 0, \; i = 1, \ldots, m. \\ & x \geq 0. \end{array}$$

$$L_\mu(x, \lambda) = f(x) - \mu \sum_{s=1}^{n} \log(x_i) + \sum_{s=1}^{m} \lambda_i \, h_s(x),$$

para $\mu \in \mathbb{R}, \; \lambda \in \mathbb{R}^m$.

$$\nabla_{(x,\lambda)} L_\mu(x, \lambda) = 0.$$

Resto

$$r_i(x_1, \ldots, x_n) = 0, \quad i = 1, \ldots, m.$$

$$\min_{x \in \mathbb{R}^n} f(x), \quad f(x) = \frac{1}{2} \sum_{i=1}^{n} r_i^2(x)$$

$$\nabla f(x) = \sum_{i=1}^{m} r_i(x) \, \nabla r_i(x) = J(x)^T r(x),$$

$$\nabla^2 f(x) = J(x)^T \, J(x) + \sum_{i=1}^{m} r_i(x) \, \nabla^2 r_i(x),$$

$$\nabla^2 f(x) [x^{k+1} - x^k] = -\nabla f(x^k)$$

Métodos

- Método de Newton,

$$(J(x^k)^T J(x^k) + \sum_{j=1}^{m} r_j(x^k) \nabla^2 r_j(x^k))\, d_N^k = -J(x^k)^T r(x_k),$$

- Método de Gauss - Newton,

$$(J(x^k)^T J(x^k))\, d_{GN}^k = -J(x^k)^T r(x_k),$$

- Método de Levenberg - Marquardt

$$(J(x^k)^T J(x^k) + \mu_k I)\, d_{LM}^k = -J(x^k)^T r(x_k),$$

$$x^{k+1} = x^k + d^N, \quad x^{k+1} = x^k + d^{GN}, \quad x^{k+1} = x^k + d^{LM},$$

Estimativas para o parâmetro de Levenberg -Marquadt

- Yamashita, N., Fukushima, M. com a suposição adicional
$$\|r(x^k)\| \geq c_1\, d(x^k, X^*)$$
e fazendo
$$\mu_k = \|r(x_k)\|_2^2$$
mostram que é possível obter convergência quadrática sem ter não singularidade na solução do $J(x^*)$.
- Jin-yan and Ya-xiang Yuan (2004), mostram que é possível relaxar
$$\mu_k = \|r(x_k)\|_2^\delta$$
preservando convergência é também quadrática, onde $\delta \in [1, 2]$

Estratégias de globalização

- Métodos com busca linear

$$x^{k+1} = x^k + t_k\, d^k, \quad d_k = -Q_k \nabla f(x^k)$$

  t_k é calculado pela regra do Armijo

- Métodos de regiões de confiança

$$\min \Phi_k(x) \text{ sujeita a } x \in B(x^k, \delta_k) \tag{1}$$

Métodos Quase-Newton

Equação Secante

$$[H_{k+1}]_{n\times n}[s^k]_{n\times 1} = [y^k]_{m\times 1},$$

onde $s^k = x^{k+1} - x^k, \quad y^k = \nabla f(x^{k+1}) - \nabla^k f(x^k)$

$$x^{k+1} = x^k - H_{k+1}^{-1}\nabla f(x^k)$$

Método de Davidon-Fletcher-Powel (DFP)

$$H_{k+1} = H_{k+1} + \frac{y^k(y^k)^T}{\langle y^k, s^k\rangle} - \frac{H_k s^k (H_k s^k)^T}{\langle (H_k s^k), s^k\rangle}$$

Metodo de Broyden-Fletcher-Goldfarb-Shano

$$H_{k+1} = H_k + \frac{(y_k - H_k s_k)s_k + s_k\,(y_k - H_k s_k)^T}{s_k^T s_k} - \frac{(y_k - H_k s_k)^T s_k\; s_k s_k^T}{(s_k^T s_k)^2}$$

Barzilai and J.M. Borwein

Consideramos o problema

$$\min_{\alpha \in \mathbb{R}} \|\alpha I_n \, s^k - y^k\|_2^2$$

Tem como solução

$$\alpha_k = \frac{\langle s^k, y^k \rangle}{\langle s^k, s^k \rangle}$$

Globalização do método com busca linear tem a seguinte forma

$$x^{k+1} = x^k - t_k \, \alpha_k^{-1} \nabla f(x^k)$$

Otimização global

Problema principal de Otimização global

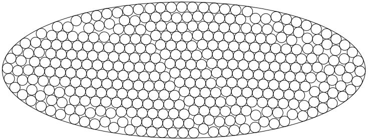
Dado um ponto estacionário (KKT), verificar quando é uma solução global.

- Não existe uma fórmula simples fácil de avaliar.

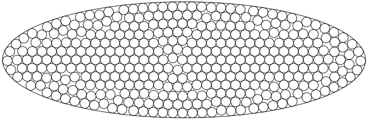
Porém, existem muitas heurísticas para encontrar o minimizador global, algumas são as seguintes:

- Multistart,
- Tunneling,
- αBB,
- ...

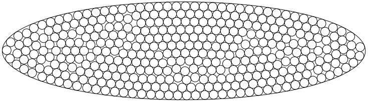
Empacotamento de círculos em elipse



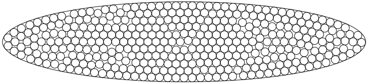
(a) s2i01



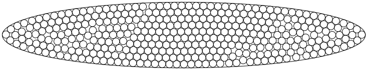
(b) s2i02



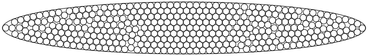
(c) s2i03



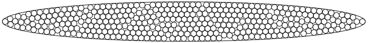
(d) s2i04



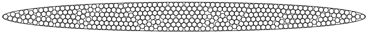
(e) s2i05



(f) s2i06



(g) s2i07



(h) s2i08

OBRIGADO!