Otimização não linear

Isabel Espírito Santo

Departamento de Produção e Sistemas

Escola de Engenharia

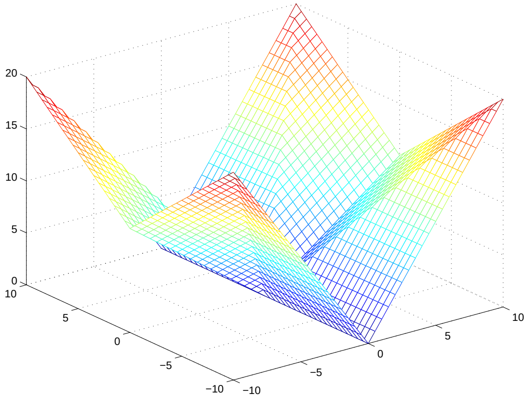
Universidade do Minho

iapinho@dps.uminho.pt

http://www.norg.uminho.pt/iapinho/

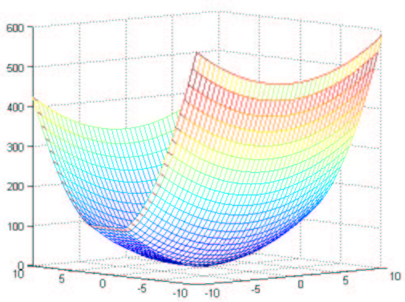
Problema não diferenciável

$$\min_{x \in \mathbb{R}^2} \quad f(x) \equiv \sum_{i=1}^{2} \min\{|x_i|, |x_1|\}$$



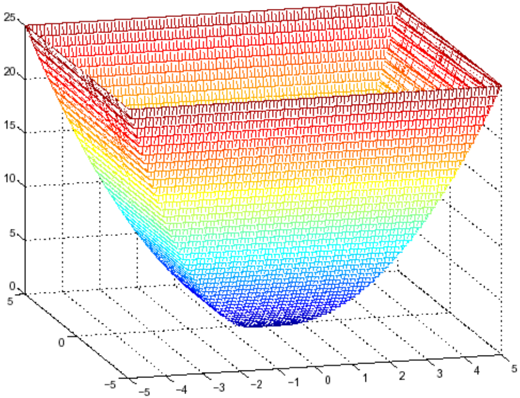
Problema não diferenciável

$$\min_{x \in \mathbb{R}^2} \; f(x) \equiv \max\{(x_1 - 1)^2, x_1^2 + 4(x_2 - 1)^2\}$$



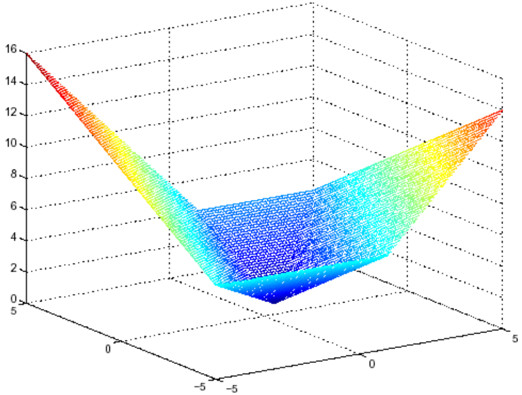
Problema não diferenciável

$$\min_{x \in \mathbb{R}^2} \; f(x) \equiv \max\{x_1^2, x_2^2\}$$



Problema não diferenciável

$$\min_{x \in \mathbb{R}^2} \; f(x) \equiv |x_1 - 1| + |x_1 - x_2|$$



Método de Nelder-Mead

O método de Nelder-Mead (NM) é iterativo e define em cada iteração um **simplex** - poliedro.

Em $\mathbb{R}^n$, sejam $x_1,\, x_2, \ldots,\, x_n,\, x_{n+1}$ – $n+1$ pontos – os vértices do simplex de dimensão n.

Notação:

$$S_k = \langle X_1, X_2, \ldots, X_n, X_{n+1} \rangle$$

representa o simplex da iteração k em que os vértices já estão ordenados por ordem crescente dos valores da função objectivo, isto é, $f(X_1) \leq f(X_2) \leq \cdots \leq f(X_n) \leq f(X_{n+1})$;

Método de Nelder-Mead

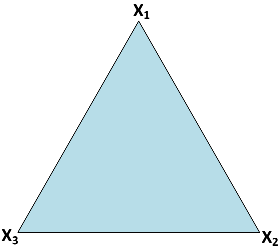
sendo

- X_1 - o melhor vértice
- X_n - o segundo pior vértice
- X_{n+1} - o pior vértice

Por exemplo, em $\mathbb{R}^2$, o simplex formado pelos $n+1(=3)$ pontos define um triângulo. Em $\mathbb{R}^3$, um simplex é um tetraedro.

Nota: Um simplex diz-se regular se as suas arestas são iguais. Em $\mathbb{R}^2$, um simplex regular é um triângulo equilátero.

Método de Nelder-Mead



Em cada iteração, definem-se pontos auxiliares - candidatos a vértices de um novo simplex - que serão aceites ou rejeitados comparando os correspondentes valores da função com $f(X_1)$, $f(X_n)$ e $f(X_{n+1})$.

Operações básicas na construção dos pontos auxiliares:

- refletir
- expandir
- contrair
- encolher

Método de Nelder-Mead

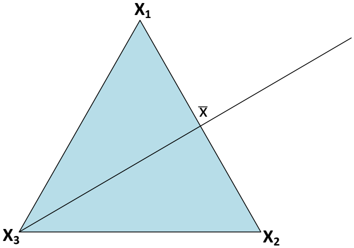
Seja

$$S_1 = \langle X_1, X_2, \ldots, X_{n+1} \rangle$$

o simplex inicial já ordenado ($k = 1$).

Em cada iteração, começa-se por calcular o **centróide** do simplex, que é o ponto médio do hiperplano definido por $X_1, X_2, \ldots, X_n$

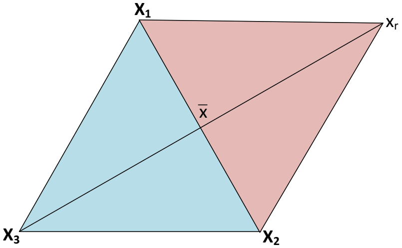
$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} X_i$$



Método de Nelder-Mead

De seguida, calcula-se o **vértice refletido** (com $\alpha = 1$)

$$\boxed{x_r = (1+\alpha)\,\bar{x} - \alpha X_{n+1}}$$



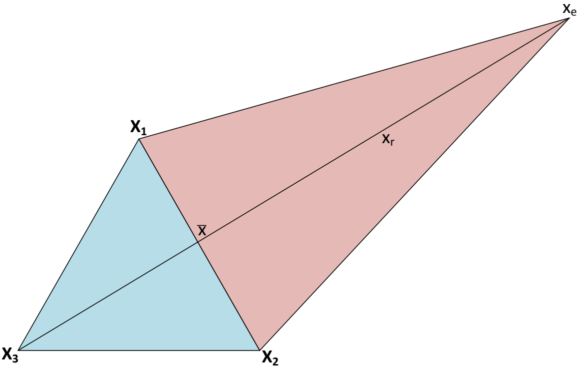
★ Se x_r for bom $(f(X_1) \leq f(x_r) < f(X_n))$ aceita-se x_r e $S_{k+1} = \langle X_1, X_2, \ldots, X_n, x_r \rangle$ é o simplex para a iteração seguinte.

Método de Nelder-Mead

★ Se x_r for muito bom $(f(x_r) < f(X_1))$ faz-se uma expansão do simplex:

- cálculo do **vértice expandido** (com $\gamma = 2$)

$$\boxed{x_e = \gamma x_r + (1 - \gamma)\,\bar{x}}$$



Método de Nelder-Mead

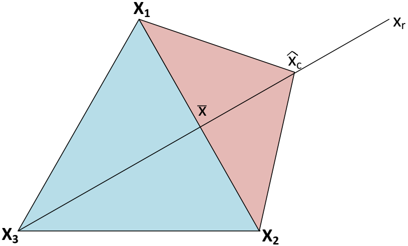
- Se x_e for muito bom $(f(x_e) < f(X_1))$ aceita-se x_e e $S_{k+1} = \langle X_1, X_2, \ldots, X_n, x_e \rangle$
- Senão aceita-se x_r e $S_{k+1} = \langle X_1, X_2, \ldots, X_n, x_r \rangle$

★ Se x_r for fraco $(f(X_n) \leq f(x_r) < f(X_{n+1}))$ faz-se uma contração para o exterior:

Método de Nelder-Mead

- cálculo do **vértice contraído para o exterior** (com $\beta = 1/2$)

$$\boxed{\hat{x}_c = \beta x_r + (1-\beta)\,\bar{x}}$$



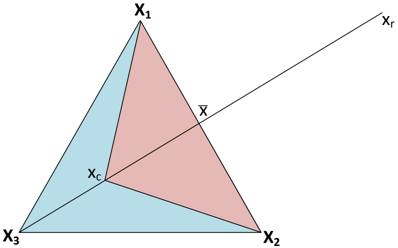
- Se $\hat{x}_c$ for bom $(f(\hat{x}_c) < f(X_n))$ aceita-se $\hat{x}_c$ e $S_{k+1} = \langle X_1, X_2, \ldots, X_n, \hat{x}_c \rangle$
- Senão encolhe-se o simplex.

Método de Nelder-Mead

★ Se x_r for muito fraco $(f(x_r) \geq f(X_{n+1}))$ faz-se uma **contração** para o interior:

- cálculo do **vértice contraído para o interior** (com $\beta = 1/2$)

$$\boxed{x_c = \beta X_{n+1} + (1-\beta)\,\bar{x}}$$



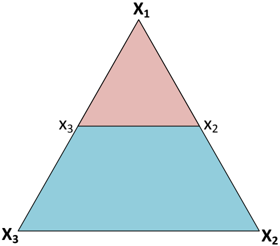
- Se x_c for bom $(f(x_c) < f(X_n))$ aceita-se x_c e $S_{k+1} = \langle X_1, X_2, \ldots, X_n, x_c \rangle$

Método de Nelder-Mead

- Senão encolhe-se o simplex.
  Encolher o simplex consiste em substituir cada um dos vértices $X_i,\ i = 2, ..., n + 1,$ pelo ponto médio do segmento que une esse X_i a X_1, i.e.,

$$\boxed{x_i = \frac{X_i + X_1}{2}}$$

  e $S_{k+1} = \langle X_1, x_2, \ldots, x_n, x_{n+1} \rangle$.



Critério de paragem - NM

O **critério de paragem** consiste em verificar se o tamanho relativo do simplex já é inferior ou igual a uma quantidade pequena, $\varepsilon > 0\,(\approx 0)$, i.e., o processo iterativo para se

$$\boxed{\frac{1}{\Delta} \max_{2 \leq i \leq n+1} \|X_i - X_1\|_2 \leq \varepsilon,}$$

com $\Delta = \max\left(1, \|X_1\|_2\right).$

Nota: Para verificar o critério de paragem é necessário que o simplex esteja **ordenado**.

Melhor aproximação - NM

Se o critério de paragem é verificado, o processo iterativo termina e o vértice do simplex com menor valor da função objetivo, X_1, é considerado como a melhor aproximação calculada à solução.

Senão, o processo iterativo continua.

Nota: Diferentes implementações do método podem ter condições diferentes no critério de paragem.