Otimização não linear
Otimização unidimensional: Método DSC
Otimização multidimensional: condições de otimalidade

Ana Maria A. C. Rocha

Departamento de Produção e Sistemas

Universidade do Minho

arocha@dps.uminho.pt

Otimização

A otimização:

- surge em processos de tomada de decisão quando se pretende atingir o melhor resultado possível;
- é um dos objetivos dos profissionais das áreas das **Ciências de Gestão** e **Engenharia** – embora também surja noutras áreas: ciências aplicadas, economia, finanças, medicina e estatística;

... está relacionada com a **maximização** ou **minimização** de modelos matemáticos - **função objetivo**

... em certos casos, as **variáveis de decisão** estão sujeitas a condições, designadas **restrições**

Classificação de problemas

De acordo com as características da função objetivo e das restrições, os problemas de otimização são divididos em Problemas de Otimização Linear e **Problemas de Otimização Não Linear**:

- Otimização Linear: se a função objetivo e as restrições são lineares;
- **Otimização Não Linear**: se a função objetivo e/ou as restrições contêm funções não lineares nas variáveis;
  - casos particulares mais fáceis de resolver:

$$\left\{\begin{array}{l}\text{problemas quadráticos}\\ \text{problemas convexos (funções convexas)}\\ \text{problemas sem restrições}\end{array}\right.$$

Classificação de problemas de otimização

Quanto ao número de restrições

- Problema sem restrições
- Problema com restrições

<table>
<tr><th>Função objetivo</th><th>Funções de restrição</th><th>Nome</th></tr>
<tr><td>linear</td><td>lineares</td><td>Programação Linear</td></tr>
<tr><td>quadrática</td><td>lineares</td><td>Programação Quadrática</td></tr>
<tr><td>polinomial (grau>2)</td><td>lineares</td><td rowspan="5">Programação
Não
Linear</td></tr>
<tr><td>genérica (envolve
funções trigonométricas,
exp, log, função inversa, ...)</td><td>lineares</td></tr>
<tr><td>linear</td><td>quadráticas
polinomiais
genéricas</td></tr>
<tr><td>quadrática</td><td>quadráticas
polinomiais
genéricas</td></tr>
<tr><td>genérica</td><td>quadráticas
polinomiais
genéricas</td></tr>
</table>

Outras classes de problemas

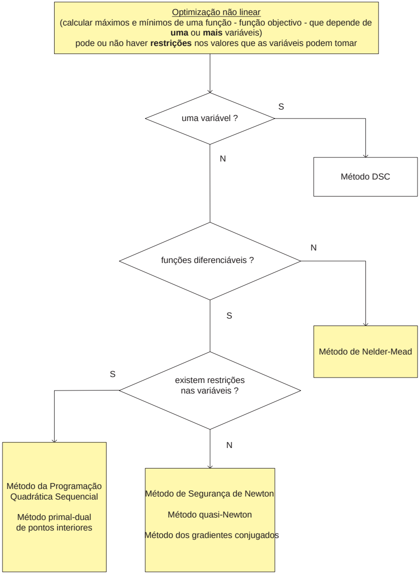
Outras diferenças que distinguem os problemas de otimização:

- diferenciável *vs* não diferenciável
- sem restrições *vs* com restrições
- unidimensional *vs* multidimensional
- convexo *vs* não convexo
- pequena/média dimensão *vs* grande dimensão
- contínuo *vs* discreto
- determinístico *vs* estocástico
- uniobjetivo *vs* multiobjetivo,

e

que obrigam a uma escolha adequada de métodos de resolução.

Características dos problemas



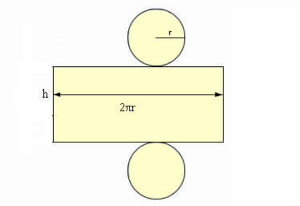
Exemplo 1

Tendo como objetivo fabricar latas cilíndricas com um volume de 1000 cm^3 e tapá-las em ambas as extremidades, qual deverá ser o raio da base e a altura da lata de modo a minimizar a quantidade de placa metálica, em termos de área superficial?

Exemplo 1 (cont.)

$$
\begin{aligned}
\text{Área Total} &= \text{Área}_{\text{retângulo}} + 2 \times \text{Área}_{\text{círculo}} \\
&= base \times h + 2\left(\pi r^2\right) \\
&= \text{Perímetro}_{\text{círculo}} \times h + 2\pi r^2 \\
&= 2\pi rh + 2\pi r^2
\end{aligned}
$$

$$
\begin{aligned}
\text{Volume} &= \pi r^2 \times h \\
1000 &= \pi r^2 \times h
\end{aligned}
$$



Formulação do problema:

$$
\begin{array}{ll}
\text{minimizar} & A(r,h) \equiv 2\pi rh + 2\pi r^2 \\
\text{sujeito a} & \pi r^2 h = 1000
\end{array}
$$

Problema com **2 variáveis** e **1 restrição**

Problema multidimensional com restrições

Exemplo 1 (cont.)

Este problema pode ser transformado num **problema sem restrições** e **uma variável**: $1000 = \pi r^2 \times h \Leftrightarrow h = \dfrac{1000}{\pi r^2}$

Substituindo em $A(r, h)$ vem

$$A(r) = 2\pi r \frac{1000}{\pi r^2} + 2\pi r^2 = \frac{2000}{r} + 2\pi r^2, r \neq 0$$

Formulação do problema:

$$\min_{r \in \mathbb{R}} \quad A(r) \equiv \frac{2000}{r} + 2\pi r^2, \quad r \neq 0$$

Problema com **1 variável** e **sem restrições**

<u>Problema unidimensional sem restrições</u>

Exemplo 2

O produto de três números positivos é igual a A (dado). Determine esses números por forma que a sua soma seja máxima.

Problema com 3 variáveis e com 1 restrição

$$\begin{array}{ll} \text{maximizar} & x_1 + x_2 + x_3 \\ \text{sujeito a} & x_1 x_2 x_3 = A \end{array}$$

Sendo $x_3 = \dfrac{A}{x_1 x_2}$ e substituindo $\Rightarrow$

Problema com 2 variáveis sem restrições

$$\max_{x_1, x_2} \quad x_1 + x_2 + \frac{A}{x_1 x_2}, \quad x_1, x_2 \neq 0$$

Exemplo 3

Uma empresa de sumos de frutas pretende lançar no mercado embalagens de sumos com capacidade de 1.5 litros e com a forma de um prisma quadrangular regular, como mostra a figura.

Calcule o comprimento da aresta da base, x, e a altura da embalagem, h, de forma que a área superficial total seja mínima.

Formulação do problema:

$$\begin{array}{ll} \text{minimizar} & 2x^2 + 4xh \\ \text{s.a} & x^2 h = 1.5 \end{array} \tag{1}$$

Formulação de um problema sem restrições

$$\min_{x \in \mathbb{R}^n} f(x) \tag{2}$$

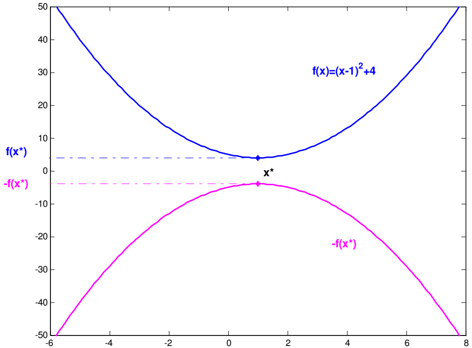
n representa o número de variáveis no problema.

- Se $n = 1 \Longrightarrow$ [problema diz-se **unidimensional**; x é escalar
- Se $n > 1 \Longrightarrow$ [problema multidimensional; $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ é vetor de dimensão n

Mínimos *vs* máximos

$$\max f(x) = -\min(-f(x))$$

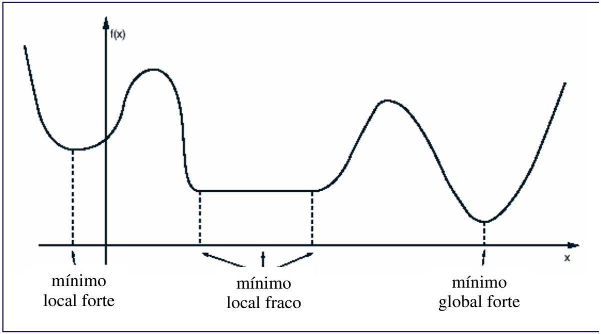
$$x^* = \underbrace{\arg\max\left(f\left(x\right)\right)}_{\text{maximizante}} = \underbrace{\arg\min\left(-f\left(x\right)\right)}_{\text{minimizante}}$$



Classificação de mínimos (e máximos)

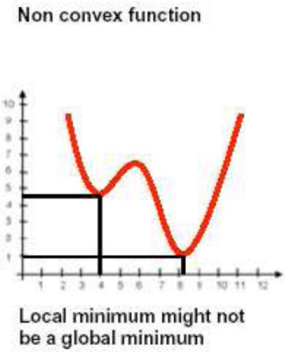
Seja $V(x, \delta)$ uma vizinhança (bola aberta) de x^* de raio δ $(\delta > 0)$.
x^* é minimizante **local** forte (fraco) se $\exists \delta > 0$:

- $f(x)$ é definida em $V(x^*, \delta)$
- $f(x^*) < (\leq) f(x), \ \forall x \in V(x^*, \delta); \ x \neq x^*$



Classificação de mínimos (e máximos)

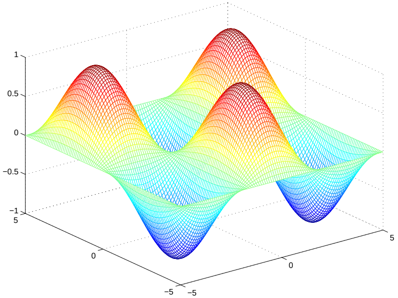
$$\begin{cases} \text{min(max)imizante global} \;\Rightarrow\; \text{min(max)imizante local} \\ \\ \text{min(max)imizante local} + f \; \underline{\text{convexa}} \; \text{(côncava)} \\ \;\Rightarrow\; \text{min(max)imizante global} \end{cases}$$



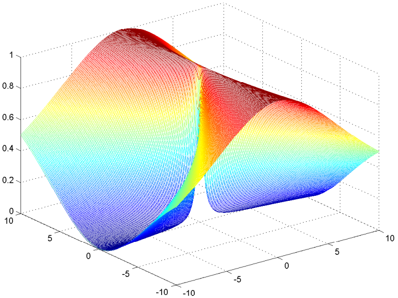
Classificação de mínimos (e máximos)

x^* é maximizante **local** forte (fraco) se $\exists \delta > 0$:

- $f(x)$ é definida em $V(x^*, \delta)$
- $f(x^*) > \ (\geq)\, f(x)\ \ \forall x \in V(x^*, \delta);\ \ x \neq x^*$



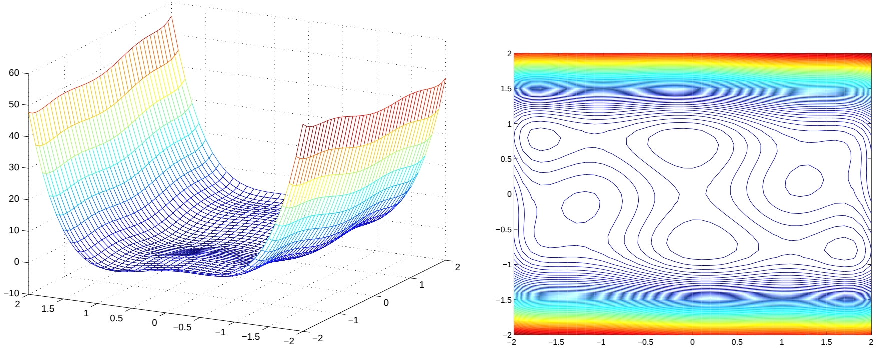
(máximos e mínimos fortes)



(máximos e mínimos fracos)

Classificação de mínimos (e máximos)

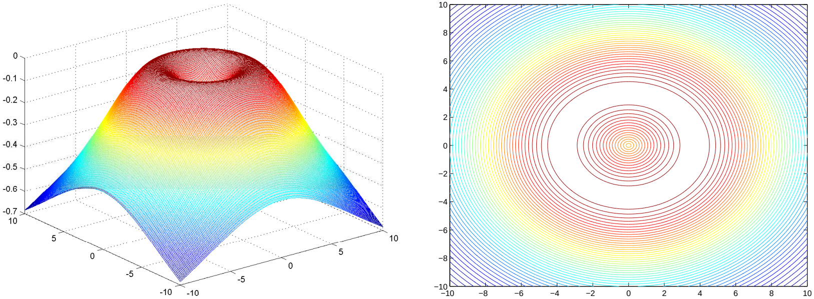
x^* é minimizante **global** forte (fraco) se $f(x^*) < (\leq) f(x)$, para todo o x que pertence ao domínio de $f(x)$ (onde a função é definida);



(2 mínimos globais e 4 mínimos locais)

Classificação de mínimos (e máximos)

x^* é maximizante **global** forte (fraco) se $f(x^*) > (\geq) f(x)$ para todo o x que pertence ao domínio de $f(x)$ (onde a função é definida);



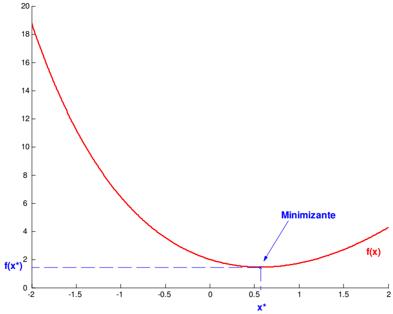
(máximos globais fracos)

Nota: Todo o ótimo global é local; no entanto, um ótimo local pode não ser global.

Problema unidimensional (n=1)

Exemplo 1

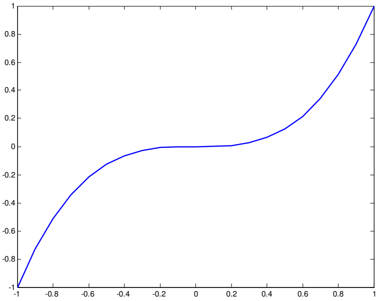
$$\min_{x \in \mathbb{R}} \; f(x) \equiv x^2 + 2e^{-x}$$



(tem 1 mínimo)

Exemplo 2

$$\min_{x \in \mathbb{R}} \; f(x) \equiv x^3$$



(não tem mínimos)

Condições de otimalidade

Assume-se que $f(x)$ é continuamente diferenciável até à 2^a ordem.

Condição necessária (e suficiente) de 1^a ordem:

Se x^* é uma solução do problema (2) ($n = 1$) então

- $f'(x^*) = 0$.

Nota: A condição $f'(x^*) = 0$ define os pontos estacionários de $f(x)$: minimizante (como no exemplo 1), maximizante, ou ponto de inflexão (como no exemplo 2).

Condição necessária de 2^a ordem (para ser mnimizante):

Se x^* é uma solução do problema (2) ($n = 1$) que satisfaz a condição de 1^a ordem, então

- $f''(x^*) \geq 0$.

Condições de otimalidade

Condição suficiente de 2^a ordem:

Se x^* é um ponto que verifica a condição de 1^a ordem e se

- $f''(x^*) > 0$

então x^* é um **minimizante** local forte de (2).

Nota: As condições necessárias e suficientes de 2^a ordem para um **maximizante** são respetivamente

- $f''(x^*) \leq 0$
- $f''(x^*) < 0$.

Métodos de resolução para problema unidimensional

$$\min_{x \in \mathbb{R}} f(x)$$

- Métodos de procura (ou pesquisa) direta (que **não** usam derivadas)
- Métodos de aproximação
- Métodos mistos

O método de Davies, Swann e Campey (DSC):

- é iterativo que só usa informação da função objetivo f
- é do tipo misto (tem uma fase de procura e uma fase de aproximação)
- a fase de aproximação usa interpolação quadrática
- é para problemas de otimização unidimensionais.

Método DSC

1. **Fase de procura**
   - procura no sentido positivo e/ou sentido negativo
   - constrói, em cada iteração, 3 pontos igualmente distanciados que definem um intervalo que contém o minimizante da função
   - compara apenas os valores da função objetivo em diversos pontos.

2. **Fase de aproximação**
   - seleciona 3 pontos igualmente distanciados que definem um intervalo que contém o minimizante da função objetivo
   - aproxima a função nesse intervalo por uma **quadrática** e usa o seu minimizante como aproximação ao minimizante da função.

3. **Critério de paragem (CP)**
   - é baseado na distância entre os pontos que foram usados na aproximação quadrática.
   - se o CP for verificado então o minimizante da quadrática é a melhor aproximação à solução.

Método de Davies, Swann e Campey (DSC)

Em qualquer iteração

a procura começa com uma aproximação inicial à solução, designada x_1, e uma perturbação $\delta > 0$.

1. procura no sentido positivo

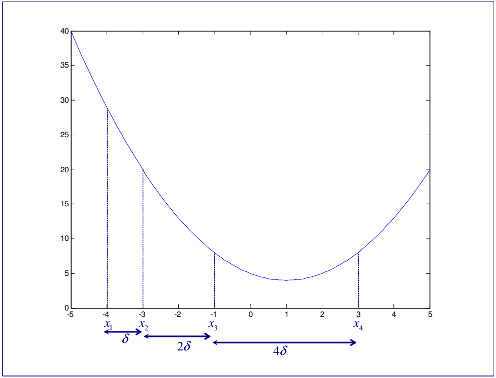
A partir do x_1 e no sentido positivo, calcula-se uma sequência de pontos $x_2, x_3, x_4, \ldots$

distanciados, respetivamente, de δ, 2δ, 4δ, 8δ, ...

$$
\begin{aligned}
& x_1 \\
& x_2 = x_1 + \delta \\
& x_3 = x_2 + 2\delta \\
& \ldots \\
& x_k = x_{k-1} + 2^{k-2}\delta
\end{aligned}
$$

até que no ponto x_k se tenha $f(x_k) > f(x_{k-1})$.

Método de Davies, Swann e Campey (DSC)



Nesta altura, tem-se

$$\cdots < x_{k-2} < x_{k-1} < x_k$$

em que $f(x_{k-2}) \geq f(x_{k-1})$ e $f(x_{k-1}) < f(x_k)$ e a distância entre x_k e x_{k-1} é duas vezes a distância entre x_{k-1} e x_{k-2}.

Método de Davies, Swann e Campey (DSC)

- Calcula-se o ponto médio do último intervalo:

$$x_m = \frac{x_k + x_{k-1}}{2}$$

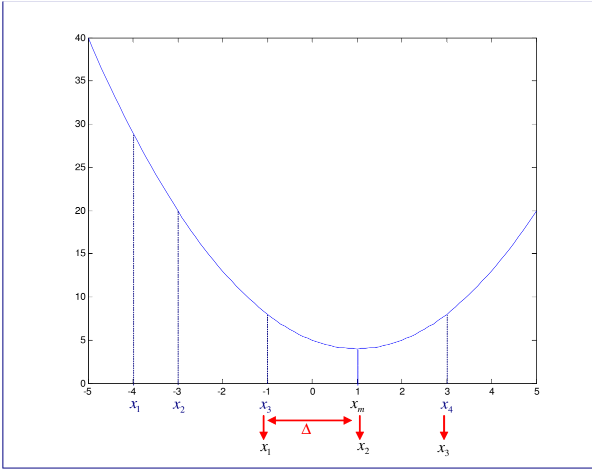
e fica-se com 4 pontos igualmente espaçados

$$x_{k-2} < x_{k-1} < x_m < x_k$$

- Para a aproximação quadrática, é necessário seleccionar três dos quatro pontos.
Comparam-se os valores de $f(x)$ nos dois pontos interiores do intervalo:

Método de Davies, Swann e Campey (DSC)

- <u>Se</u> $f(x_{k-1}) \leq f(x_m)$ <u>então</u> escolhem-se os pontos x_{k-2}, x_{k-1} e x_m <u>senão</u> $(f(x_{k-1}) > f(x_m))$ escolhem-se os pontos x_{k-1}, x_m e x_k



Fase de aproximação do método DSC

MIN q O minimizante da quadrática, $x^*(q)$, que passa por estes três pontos (nesta fase, redefinidos por $x_1 < x_2 < x_3$) determina-se por

$$x^*(q) = x_2 + \Delta \frac{f(x_1) - f(x_3)}{2\left(f(x_3) - 2f(x_2) + f(x_1)\right)}$$

com $\Delta = (x_2 - x_1) = (x_3 - x_2)$.

Critério de paragem:

verificar se a distância entre os pontos que foram usados para construir a quadrática não excede $\varepsilon > 0 \quad (\approx 0)$:

$$(x_2 - x_1) = (x_3 - x_2) = \Delta \leq \varepsilon$$

Paragem do método DSC ?

i) Se o critério de paragem for verificado, o processo iterativo termina, sendo $x^*(q)$ a melhor aproximação calculada à solução;

ii) Se o critério de paragem não se verificar, o processo repete-se e o minimizante da quadrática, $x^*(q)$, passa a ser o x_1 da nova iteração. A perturbação δ também deve ser reduzida através de: $\delta = M\delta$, com $M < 1$.

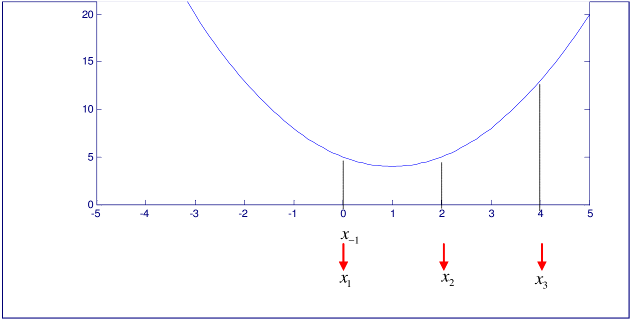
2. procura no sentido negativo

Quando, a partir de x_1, o valor de $f(x_2) > f(x_1)$ (para $x_2 = x_1 + \delta$) a procura deve voltar-se para o sentido negativo, a começar novamente por x_1.

O próximo ponto, na procura, é $x_{-1} = x_1 - \delta$.

Método de Davies, Swann e Campey (DSC)

2. i) Se $f(x_{-1}) > f(x_1)$, significa que o intervalo definido por $[x_{-1}, x_2]$, com x_1 como ponto médio, contém o minimizante desejado. Nesta altura, determina-se o minimizante da quadrática (que passa pelos **três pontos** agora calculados), $x^*(q)$, tal como está descrito no ponto **MIN q**.



Método de Davies, Swann e Campey (DSC)

2. ii) No entanto, se $f(x_{-1}) < f(x_1)$, significa que a procura deve continuar no sentido negativo até que $f(x_{-k}) > f(x_{-(k-1)})$, isto é, procede-se da seguinte forma:

$$
\begin{aligned}
&x_{-2} = x_{-1} - 2\delta \\
&\dots \\
&x_{-k} = x_{-(k-1)} - 2^{k-1}\delta
\end{aligned}
$$

até que no ponto x_{-k} se tenha $f(x_{-k}) > f\left(x_{-(k-1)}\right)$.
Nesta altura, tem-se

$$x_{-k} < x_{-(k-1)} < x_{-(k-2)} < \cdots$$

em que $f\left(x_{-(k-2)}\right) \geq f\left(x_{-(k-1)}\right)$ e $f\left(x_{-(k-1)}\right) < f(x_{-k})$

Método de Davies, Swann e Campey (DSC)

e a distância entre x_{-k} e $x_{-(k-1)}$ é duas vezes a distância entre $x_{-(k-1)}$ e $x_{-(k-2)}$.

Calcula-se o ponto médio do <u>último intervalo</u>:

$$x_m = \frac{x_{-k} + x_{-(k-1)}}{2}$$

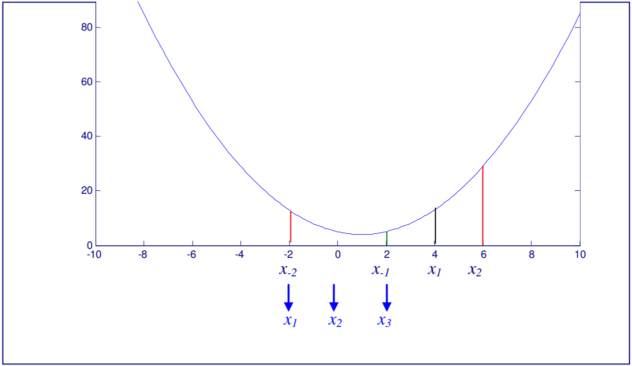
e fica-se com 4 pontos igualmente espaçados

$$x_{-k}, x_m, x_{-(k-1)}, x_{-(k-2)}$$

<u>Se</u> $f(x_m) < f\left(x_{-(k-1)}\right)$ <u>então</u> escolhem-se os pontos x_{-k}, x_m e $x_{-(k-1)}$
<u>senão</u> $\left(f(x_m) \geq f\left(x_{-(k-1)}\right)\right)$ escolhem-se os pontos $x_m, x_{-(k-1)}$ e $x_{-(k-2)}$

Método de Davies, Swann e Campey (DSC)

O processo entra na fase de aproximação - ponto **MIN q** - calcula-se o minimizante da quadrática que passa pelos 3 pontos agora selecionados - tal como foi descrito na procura no sentido positivo:

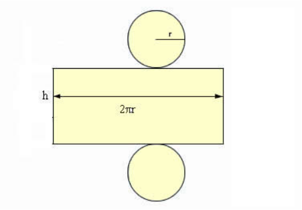


Exercício

Tendo como objetivo fabricar latas cilíndricas com um volume de $1000\ cm^3$ e tapá-las em ambas as extremidades, qual deverá ser o raio da base e a altura da lata de modo a minimizar a quantidade de placa metálica, em termos de área superficial? Utilize o algoritmo de DSC, baseado na interpolação quadrática, com o valor inicial $r_1 = 7$, $\delta = 0.5$, $\varepsilon = 0.3$ e $M = 0.5$. Use 4 casas decimais nos cálculos.

$$\boxed{\min_{r \in \mathbb{R}} \quad \underbrace{\frac{2000}{r} + 2\pi r^2}_{\text{função objetivo}}}$$

Resolução do Exercício



$$\begin{aligned}
\text{Área Total} &= \text{Área}_{\text{retângulo}} + 2 \times \text{Área}_{\text{círculo}} \\
&= base \times h + 2\left(\pi r^2\right) \\
&= \text{Perímetro}_{\text{círculo}} \times h + 2\pi r^2 \\
&= 2\pi rh + 2\pi r^2
\end{aligned}$$

$$\begin{aligned}
\text{Volume} &= \pi r^2 \times h \\
1000 &= \pi r^2 \times h
\end{aligned}$$

$$\min \quad \underbrace{2\pi rh + 2\pi r^2}_{A(r,h)}$$

$$\begin{aligned}
\text{Volume} &= \pi r^2 \times h \\
1000 &= \pi r^2 \times h \Longrightarrow h = \frac{1000}{\pi r^2}
\end{aligned}$$

Substituindo em $A(r,h)$ vem $A(r) = 2\pi r \dfrac{1000}{\pi r^2} + 2\pi r^2 = \dfrac{2000}{r} + 2\pi r^2$

Resolução do Exercício (cont.)

$$\boxed{\begin{array}{ll} \min & \dfrac{2000}{r} + 2\pi r^2 \\ \text{s.a} & r \in \mathbb{R} \end{array}}$$

Iniciar algoritmo DSC

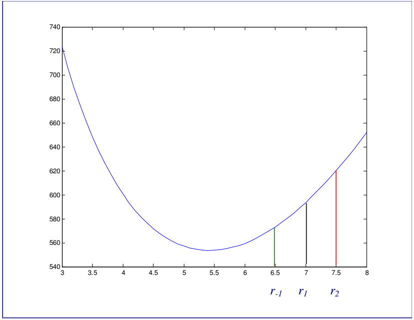
$r_1 = 7, \delta = 0.5, \varepsilon = 0.3, M = 0.5$

$$\begin{cases} r_1 = 7 \\ A\left(r_1\right) = 593.5904 \end{cases}$$

$$\begin{cases} r_2 = 7 + 0.5 = 7.5 \\ A\left(r_2\right) = 620.0958 \end{cases}$$

$A\left(r_2\right) > A\left(r_1\right) \Rightarrow$ sentido negativo

$$\begin{cases} r_{-1} = 7 - 0.5 = 6.5 \\ A\left(r_{-1}\right) = 573.1569 \end{cases}$$

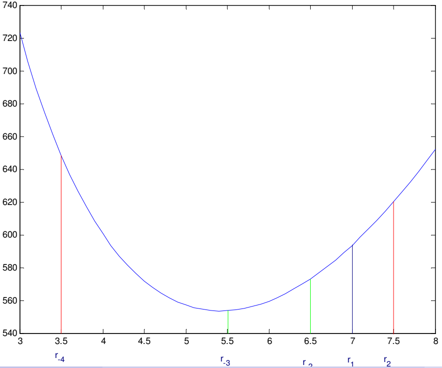


Resolução do Exercício (cont.)

$$A(r_{-1}) < A(r_1) \Rightarrow \begin{cases} r_{-2} = 6.5 - 2 \times 0.5 = 5.5 \\ A(r_{-2}) = 553.7027 \end{cases}$$

$$A(r_{-2}) < A(r_{-1}) \Rightarrow \begin{cases} r_{-3} = 5.5 - 4 \times 0.5 = 3.5 \\ A(r_{-3}) = 648.3976 \end{cases}$$

$A(r_{-3}) > A(r_{-2}) \Rightarrow$ PARAR.



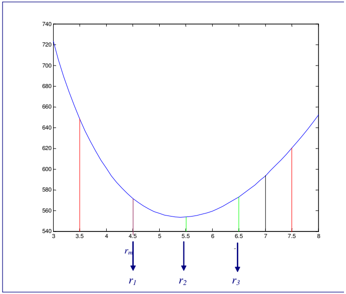
Resolução do Exercício (cont.)

Calcular ponto médio.

$$\begin{cases} r_m = (5.5 + 6.5)/2 = 4.5 \\ A(r_m) = 571.6789 \end{cases}$$

Escolher 3 pontos igualmente espaçados.

$$\begin{cases} r_1 = 4.5 & A(r_1) = 571.6789 \\ r_2 = 5.5 & A(r_2) = 553.7027 \\ r_3 = 6.5 & A(r_3) = 573.1569 \end{cases}$$



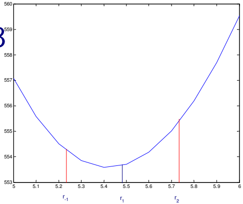
Minimizante da quadrática

$$\begin{cases} r_{\min} = 5.4803 \\ A(r_{\min}) = 553.6508 \end{cases}$$

Testar CP: $(\Delta = (r_2 - r_1) = 1) \leq 0.3$

Falso, por isso fazer nova iteração.

Resolução do Exercício (cont.)



Fazer $\delta = M\delta = 0.5 \times 0.5 = 0.25$ e $r_{min} \to r_1$

$$\begin{cases} r_1 = 5.4803 \\ A(r_1) = 553.6508 \end{cases} \quad \begin{cases} r_2 = 5.4803 + 0.25 = 5.7303 \\ A(r_2) = 555.3387 \end{cases}$$

Como $A(r_2) > A(r_1)$ então

$$\begin{cases} r_{-1} = 5.4803 - 0.25 = 5.2303 \\ A(r_{-1}) = 554.2703 \end{cases}$$

Como $A(r_{-1}) > A(r_1)$ então PARAR. Não é necessário calcular ponto médio.

Ordenar pontos. $\begin{cases} r_1 = 5.2303 & A(r_1) = 554.2703 \\ r_2 = 5.4803 & A(r_2) = 553.7027 \\ r_3 = 5.7303 & A(r_3) = 555.3387 \end{cases}$

Calcular minimizante da quadrática $\begin{cases} r_{\text{min}} = 5.4224 \\ A(r_{\text{min}}) = 553.5812 \end{cases}$

Testar CP: $((r_2 - r_1) = 0.25) \leq 0.3 \Leftrightarrow$ Verdadeiro

Solução$\Rightarrow \begin{cases} r^* = 5.4224 \text{ e a altura } h^* = \frac{1000}{\pi r^2} = 12.06 \\ A(r^*) = 553.5812 \end{cases}$

Formulação de um problema sem restrições

$$\min_{x \in \mathbb{R}^n} f(x) \tag{3}$$

n representa o número de variáveis no problema.

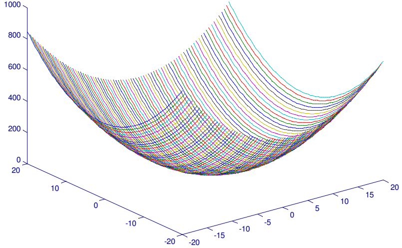
- Se $n = 1 \Longrightarrow \left[\begin{array}{l} \text{problema diz-se unidimensional} \\ x \text{ é escalar} \end{array} \right.$

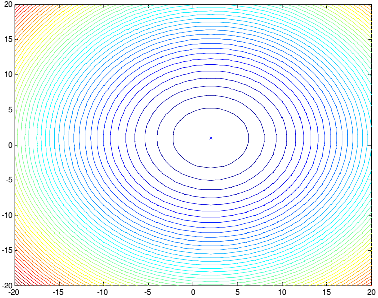
- Se $n > 1 \Longrightarrow \left[\begin{array}{l} \text{problema } \textbf{multidimensional} \\ x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ é vetor de dimensão } n \end{array} \right.$

Problema multidimensional sem restrições

Exemplo 1

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

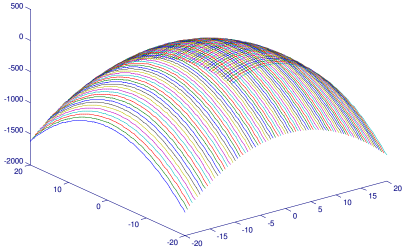


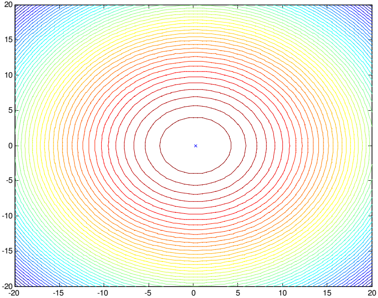


Problema multidimensional sem restrições

Exemplo 2

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

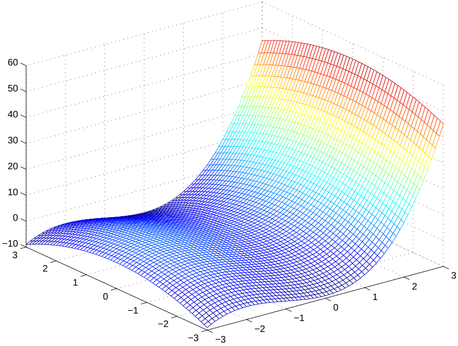


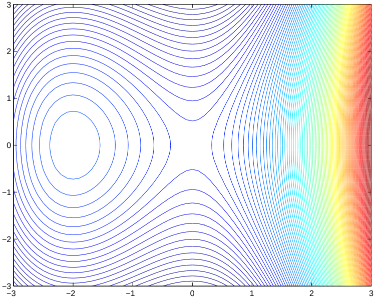


Problema multidimensional sem restrições

Exemplo 3

$$\min_{x \in \mathbb{R}^2} \; f(x) \equiv 3x_1^2 - x_2^2 + x_1^3$$



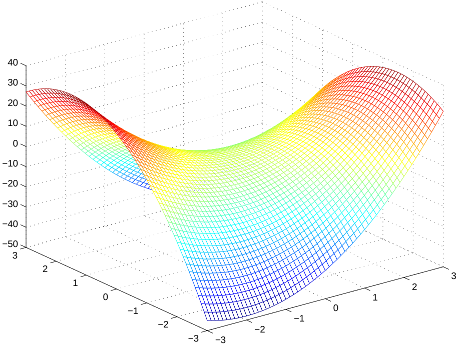


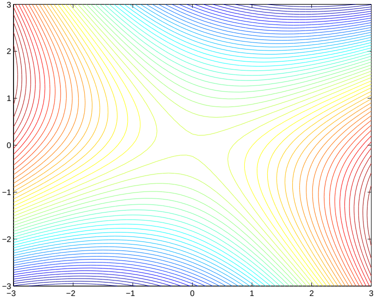
Ponto sela em $(0, 0)$

Problema multidimensional sem restrições

Exemplo 4

$$\min_{x \in \mathbb{R}^2} \; f(x) \equiv 3x_1^2 - 4x_1x_2 - 4x_2^2$$





Ponto sela em $(0, 0)$

Notação

Vetor <u>gradiente</u> da função $f(x)$ - $x \in \mathbb{R}^n$ -

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix} \text{ vetor de } \mathbb{R}^n$$

Matriz <u>Hessiana</u> da função $f(x)$

$$\nabla^2 f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \frac{\partial^2 f}{\partial x_2 \partial x_n} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix} \begin{array}{c} \text{matriz} \\ \text{simétrica} \\ \text{de } n \times n \end{array}$$

Condições de otimalidade

Assume-se $f(x)$ continuamente diferenciável até à 2^a ordem.

Condição necessária (e suficiente) de 1^a ordem:

Se x^* é uma solução do problema (3) então o vetor gradiente calculado em x^* anula-se, i.e.,

$$\nabla f(x^*) = 0$$

Os pontos estacionários da função f são os pontos que verificam

$$\nabla f(x) = \mathbf{0}$$

$$\left\{ \begin{array}{l} \text{minimizante (como no Exemplo 1)} \\ \text{maximizante (como no Exemplo 2)} \\ \text{ponto sela (como nos Exemplos 3 e 4)} \end{array} \right.$$

Condições de otimalidade

Condição necessária de 2^a ordem:

Se x^* é uma solução do problema (3) que satisfaz a condição de 1^a ordem, então $\nabla^2 f(x^*)$ é semi-definida positiva.

Condição suficiente de 2^a ordem:

Se x^* é um ponto que verifica a condição de 1^a ordem e se a matriz Hessiana calculada em x^*, $\nabla^2 f(x^*)$, é <u>definida positiva</u>, então x^* é um **minimizante** local forte de (3).

Condição suficiente de 2^a ordem:

Se x^* é um ponto que verifica a condição de 1^a ordem e se a matriz Hessiana calculada em x^*, $\nabla^2 f(x^*)$, é <u>definida negativa</u>, então x^* é um **maximizante** local forte de (3).

Condições de otimalidade

Assumindo $\nabla f(x^*) = 0$**:**

Condições de 2^a ordem para um minimizante

$\nabla^2 f(x^*)$ é semi-definida positiva (condição necessária)

$\nabla^2 f(x^*)$ é <u>definida positiva</u> (condição suficiente)

Condições de 2^a ordem para um maximizante

$\nabla^2 f(x^*)$ é semi-definida negativa (condição necessária)

$\nabla^2 f(x^*)$ é <u>definida negativa</u> (condição suficiente)

Se $\nabla^2 f(x^*)$ é <u>indefinida</u>, então x^* é ponto sela (ou de descanso).

Definições

- Uma matriz diz-se **definida positiva** se os determinantes das submatrizes principais são positivos;
- se pelo menos um dos determinantes das submatrizes principais é zero e os outros são positivos, a matriz diz-se **semi-definida positiva**.
- Uma matriz diz-se **definida negativa** se os determinantes das submatrizes principais têm sinais alternados, sendo o de ordem 1 negativo;
- se pelo menos um dos determinantes das submatrizes principais é zero e os outros têm sinais alternados, sendo o de ordem 1 negativo, a matriz diz-se **semi-definida negativa**.
- Uma matriz diz-se **indefinida** se os sinais dos determinantes das submatrizes principais não verificam nenhuma das 4 situações acima mencionadas.

NOTA: Para identificar uma matriz semi-definida positiva ou semi-definida negativa, a sequência dos determinantes tem de ter pelo menos um elemento não nulo. Se toda a sequência for constituída por elementos nulos nada se pode concluir.

Resumo

Seja x^* um ponto para o qual $\nabla f(x^*) = 0$ e $\nabla^2 f(x^*) \neq$ matriz nula:

- Se $\nabla^2 f(x^*)$ é definida positiva então x^* é minimizante
- Se $\nabla^2 f(x^*)$ é definida negativa então x^* é maximizante
- Se $\nabla^2 f(x^*)$ é semi-definida positiva então x^* é minimizante ou ponto sela
- Se $\nabla^2 f(x^*)$ é semi-definida negativa então x^* é maximizante ou ponto sela
- Se $\nabla^2 f(x^*)$ é indefinida então x^* é ponto sela.

Exercício

Dada a função $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ definida por

$$f(x_1, x_2, x_3) = 5x_1^2 + 2x_2^2 + x_3^4 - 32x_3 + 6x_1x_2 + 5x_2$$

verifique que ela tem apenas um ponto estacionário. Use a condição suficiente de 2^a ordem para o classificar.

Resolução do Exercício

Determinação do vetor gradiente e matriz Hessiana da função:

$$\nabla f(x) = \begin{pmatrix} 10x_1 + 6x_2 \\ 4x_2 + 6x_1 + 5 \\ 4x_3^3 - 32 \end{pmatrix} \text{ e } \nabla^2 f(x) = \begin{pmatrix} 10 & 6 & 0 \\ 6 & 4 & 0 \\ 0 & 0 & 12x_3^2 \end{pmatrix}$$

- Resolver o sistema

$$\nabla f(x) = 0$$

Resolução do Exercício (cont.)

Da **condição de 1^a ordem** (resolução do sistema $\nabla f(x) = 0$) temos que

$$x^* = (7.5, -12.5, 2)^T \text{ é ponto estacionário de } f(x_1, x_2, x_3).$$

Da **condição de 2^a ordem**

$$\nabla^2 f(x^*) = \begin{pmatrix} 10 & 6 & 0 \\ 6 & 4 & 0 \\ 0 & 0 & 48 \end{pmatrix} \Rightarrow \begin{cases} \det\left(\nabla^2 f_1\right) = 10 > 0 \\ \det\left(\nabla^2 f_2\right) = 4 > 0 \\ \det\left(\nabla^2 f_3\right) = 192 > 0 \end{cases}$$

Como $\nabla^2 f(x^*)$ é <u>definida positiva</u>, então

$$x^* = (7.5, -12.5, 2)^T \text{ é } \textbf{minimizante}.$$