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/

Otimização

Otimização:

- surge no processo de tomada de decisões para se atingir o melhor resultado possível;
- é um dos objetivos dos profissionais das áreas das **Ciências de Gestão** e **Engenharia**;
- 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**,
- também surge noutras áreas: ciências aplicadas, economia, finanças, medicina e estatística.

Classificação de problemas

Os problemas de otimização são divididos em Problemas de Otimização Linear e **Problemas de Otimização Não Linear** - de acordo com as características das funções objetivo e de restrição:

- Otimização Linear: se a função objetivo e as restrições são lineares;
- **Otimização Não Linear (ONL)**: se o objetivo e as restrições contêm funções não lineares nas variáveis;

  casos particulares:
  - problemas quadráticos
  - problemas convexos (funções convexas)
  - problemas sem restrições.

Classificação de problemas - exemplos

ONL:

$$\begin{array}{ll} \min & x_1^2 + 3x_2^2 \\ \text{s.a} & x_1 + 5x_2 - 1 \geq 0 \end{array}$$

Problema quadrático
rest. de desigualdade
funções diferenciáveis

ONL:

$$\begin{array}{ll} \min & 3x_1 - 4x_2 \\ \text{s.a} & (x_1 + 1)^2 + x_2^2 - 1 = 0 \\ & (x_1 - 1)^2 + x_2^2 - 1 = 0 \end{array}$$

rest. de igualdade
funções diferenciáveis

Classificação de problemas - exemplos

ONL:

$$\begin{array}{ll} \min & -x_1^2 - x_2^2 + x_3^2 \\ \text{s.a} & x_1 + x_2 = -1 \\ & x_1^2 + x_2^2 = 0.5 \end{array}$$

rest. de igualdade
funções diferenciáveis

ONL:

$$\min \quad (x_1 - 1)^2 + x_2^3 - x_1 x_2$$

sem restrições
função diferenciável

ONL:

$$\max \quad 2(-x_1^2 - x_2^2 + 1) + x_1$$

sem restrições
função diferenciável

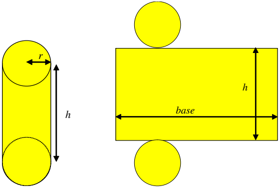
ONL:

$$\min \quad \max\{x_1, x_2\} + (|x_1| + |x_2|)$$

sem restrições
função não diferenciável

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.)

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

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

Formulação do problema:

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

Problema com 2 variáveis e 1 restrição

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 $A\left(r, h\right)$ vem

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

Problema unidimensional sem restrições

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

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

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

Formulação de um problema sem restrições

$$\min_{x \in \mathbb{R}^n} f(x) \tag{1}$$

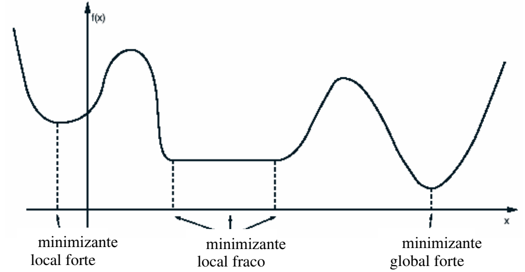
- Se $n = 1 \Longrightarrow \left[\begin{array}{l} \text{problema unidimensional} \\ x \text{ é escalar} \end{array} \right.$
- Se $n > 1 \Longrightarrow \left[\begin{array}{l} \text{problema multidimensional} \\ x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ é vetor de dimensão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

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 $\exists \delta > 0$:

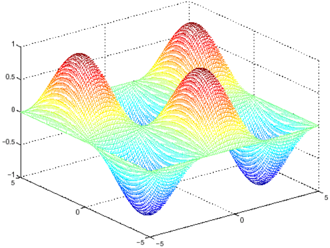
- $f(x)$ é definida em $V(x^*, \delta)$
- $f(x^*) < \ (\leq) \ f(x), \ \forall x \in V(x^*, \delta); \ x \neq x^*$



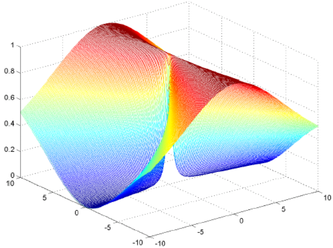
Classificação de mínimos (e máximos)

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

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



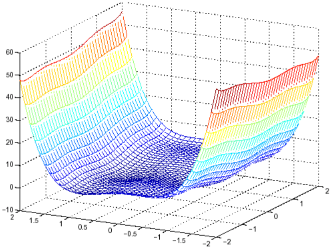
(máximos e mínimos fortes)

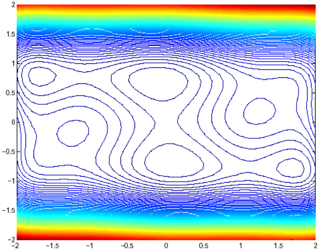


(máximos e mínimos fracos)

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);

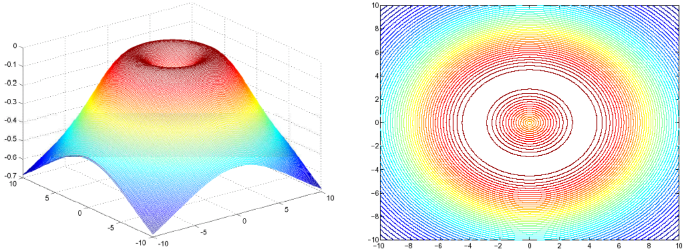




(2 mínimos globais e 4 mínimos locais)

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);



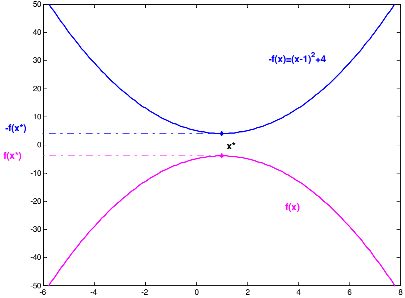
(máximos globais fracos)

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

Mínimos *vs* máximos

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

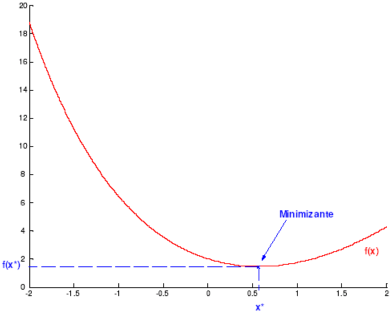
$$x^* = \underbrace{\arg\max\left(f\left(x\right)\right)}_{\text{maximizante}} = \underbrace{\arg\min\left(-f\left(x\right)\right)}_{\text{minimizante}}$$



Problema unidimensional ($n = 1$)

Exemplo 3

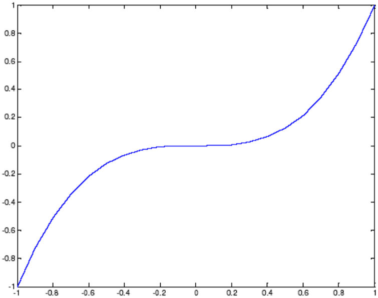
$$\min_{x \in \mathbb{R}} \; f(x) \equiv x^2 + 2e^{-x}$$



(tem 1 mínimo)

Exemplo 4

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



(não tem mínimos)

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 (1) ($n = 1$) então

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

Nota: A equação $f'(x) = 0$ define os pontos estacionários da função objetivo $f(x)$:

$$\begin{cases} \text{minimizante (exemplo 3)} \\ \text{maximizante} \\ \text{ponto de inflexão (exemplo 4).} \end{cases}$$

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

Pontos estacionários: $f'(x) \equiv 2x - 2e^{-x} = 0 \;\Leftrightarrow$
$2(x - e^{-x}) = 0 \;\Leftrightarrow$ as soluções **desta equação não linear** em x - resolução pelo método iterativo da secante ou Newton - são: (única) 0.567143

Condições de otimalidade

Condição necessária de 2^a ordem:

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

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

Condição suficiente de 2^a ordem:

- Se x^* é tal que $f'(x^*) = 0$ e se
$$f''(x^*) > 0$$
então x^* é um minimizante local forte de (1).
- Se x^* é tal que $f'(x^*) = 0$ e se
$$f''(x^*) < 0$$
então x^* é um maximizante local forte de (1).

Métodos de resolução ($n = 1$)

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

- Métodos de procura (ou pesquisa) direta
- Métodos de aproximação
- Métodos mistos

O método de Davies, Swann e Campey (DSC)

- é um método iterativo que só usa informação da função objetivo, f
- é um método misto constituído por uma fase de procura e uma fase de aproximação
- a fase de aproximação é baseada numa interpolação quadrática
- destina-se a problemas de otimização unidimensionais

Método misto de DSC

1. **Fase de procura**
   constrói, em cada iteração, 3 pontos igualmente distanciados que definem um intervalo que contém o minimizante da função, comparando apenas os valores da função em diversos pontos.
2. **Fase de aproximação**
   aproxima a função no intervalo por uma quadrática e usa o seu minimizante como aproximação ao minimizante da função.

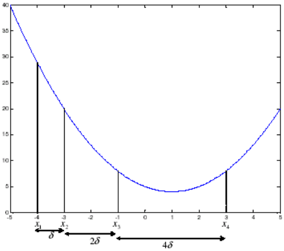
Método de Davies, Swann e Campey (DSC)

- A procura começa com 1 aproximação inicial x_1 e uma perturbação $\delta > 0$.
- A partir do x_1 e no sentido positivo, calcula-se uma sequência de pontos $x_2,\ x_3,\ x_4,\ \ldots$ distanciados uns dos outros de, respectivamente, $\delta, 2\delta, 4\delta, 8\delta, \ldots$

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

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)

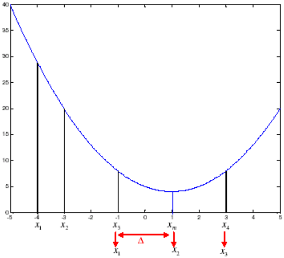
- Calcula-se o ponto médio do último intervalo:

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

e fica-se com 4 pontos igualmente espaçados

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

Método de Davies, Swann e Campey (DSC)



- Para a aproximação quadrática, é necessário seleccionar três dos quatro pontos. Para isso, comparam-se os valores de $f\left(x\right)$ nos dois pontos interiores do intervalo:
  <u>Se</u> $f\left(x_{k-1}\right) \leq f\left(x_m\right)$ <u>então</u> escolhem-se os pontos x_{k-2}, x_{k-1} e x_m
  <u>senão</u> $\left(f\left(x_{k-1}\right) > f\left(x_m\right)\right)$ 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 $\Delta = (x_2 - x_1) = (x_3 - x_2)$.

- De seguida, verifica-se o critério de paragem, que consiste em verificar se a distância entre os pontos que foram usados para construir a quadrática não excede uma certa quantidade:

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

em que $\varepsilon > 0 \quad (\approx 0)$.

Paragem do método DSC

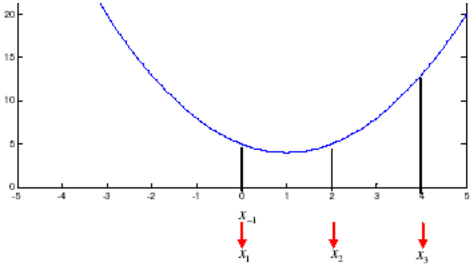
i) Se o critério de paragem for verificado, o processo iterativo termina, sendo $x^*(q)$ a melhor aproximação calculada à solução;

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 $M < 1$.

- 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 <u>sentido negativo</u>, a começar novamente por x_1. Neste caso, o próximo ponto, na procura, é $x_{-1} = x_1 - \delta$, e

Método de Davies, Swann e Campey (DSC)

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)

ii) No entanto, se $f\left(x_{-1}\right) < f\left(x_1\right)$, 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: 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 $f\left(x_{-(k-1)}\right) < f\left(x_{-k}\right)$

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 seleccionados - tal como foi descrito na procura no sentido positivo: