Optimização

Aula 25

Prof. Doutor Engº Jorge Nhambiu

Programação Não Linear

Aula 25: Programação Não-Linear
- Funções de Uma única variável

- Mínimo;
- Mínimo Global;
- Mínimo Local;
- Optimização Irrestrita;
- Condições Óptimas;
- Método da Bissecção;
- Método de Newton:
- Exemplos.



Programação Não Linear

A Programação Não Linear consiste em encontrar o vector de variáveis $x = (x_1, x_2, \ldots, x_n)$ que optimiza (minimiza ou maximiza) uma função objectivo sujeita ou não a restrições.



Características Principais:

- A função objectivo e/ou restrições não são lineares
- Pode envolver múltiplos mínimos locais e um global
- Não existe um único algoritmo geral - há várias classes de problemas (irrestritos, convexos, fraccionários, geométricos, etc.)

Programação Não Linear



Por que estudar Programação Não Linear?

- Permite resolver problemas reais de engenharia, economia e ciência dos materiais
- Usada em optimização de sistemas energéticos, calibração, desenho mecânico e controlo de processos
- Forma a base para métodos numéricos avançados (ex.: Bissecção, Newton, Gradiente)

Programação Não Linear



Em que consiste a Programação Não Linear?



A Programação Não Linear de uma forma geral consiste em encontrar um valor de $x = (x_1, x_2, \ldots x_n)$ de modo a:

$$\begin{aligned}
&Maximizar\ f(x),\\
&sujeito\ a\\
&h_i(x) = b_i\ para\ i = 1, 2, \ldots n,\\
&g_i(x) \leq b_i\ para\ i = 1, 2, \ldots n,\\
&e\\
&x \geq 0,
\end{aligned}$$

Em que f(x), h(x) e g(x) são funções dadas das n variáveis de decisão.

Mínimo

Optimizar é procurar numa região um valor provável de dar o menor valor da função custo. Durante a explanação passa-se a designar por x* a um valor particular do conjunto de restrições.

Mínimo Global

Uma função f(x) de n variáveis tem um mínimo global no ponto x* se para todos os valores de x na região provável:

$$f\left(x^*\right) \leq f\left(x\right)$$

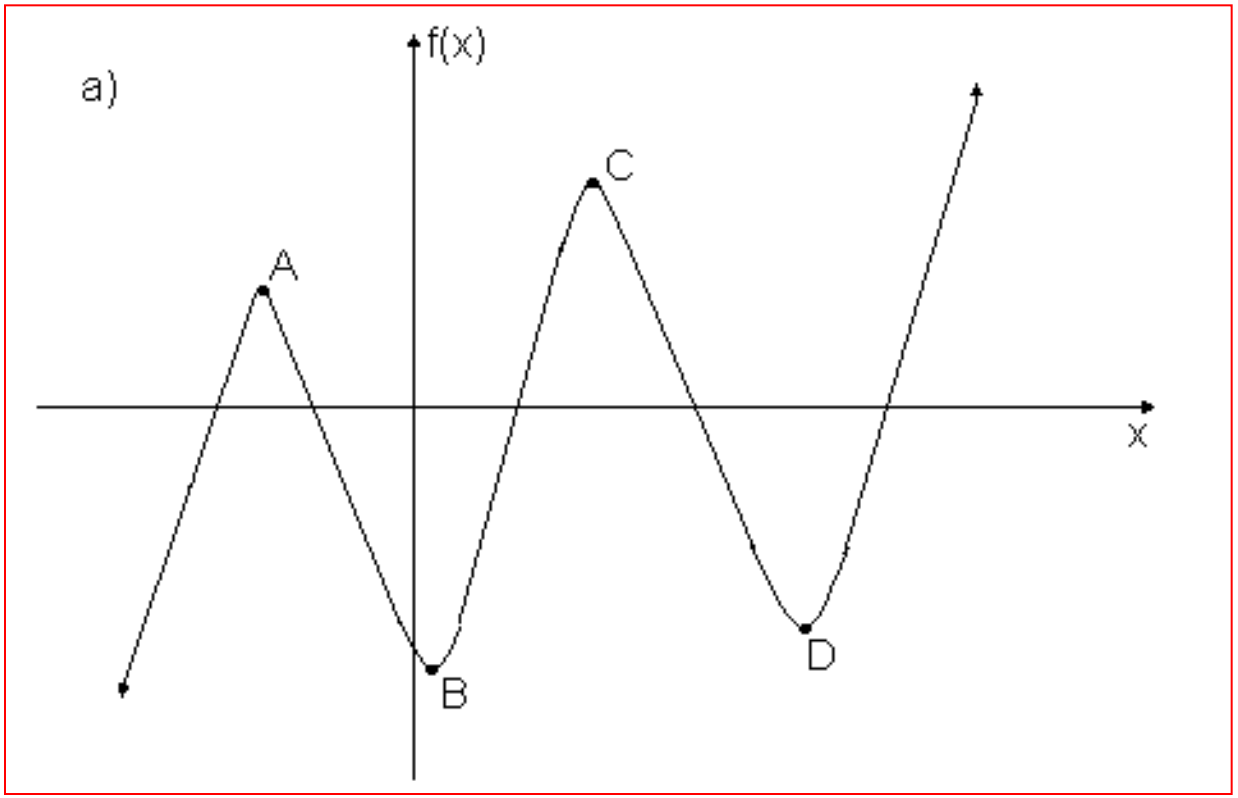
Se a desigualdade for satisfeita para todos os valores de x então considera-se x* mínimo global.

Mínimo Local

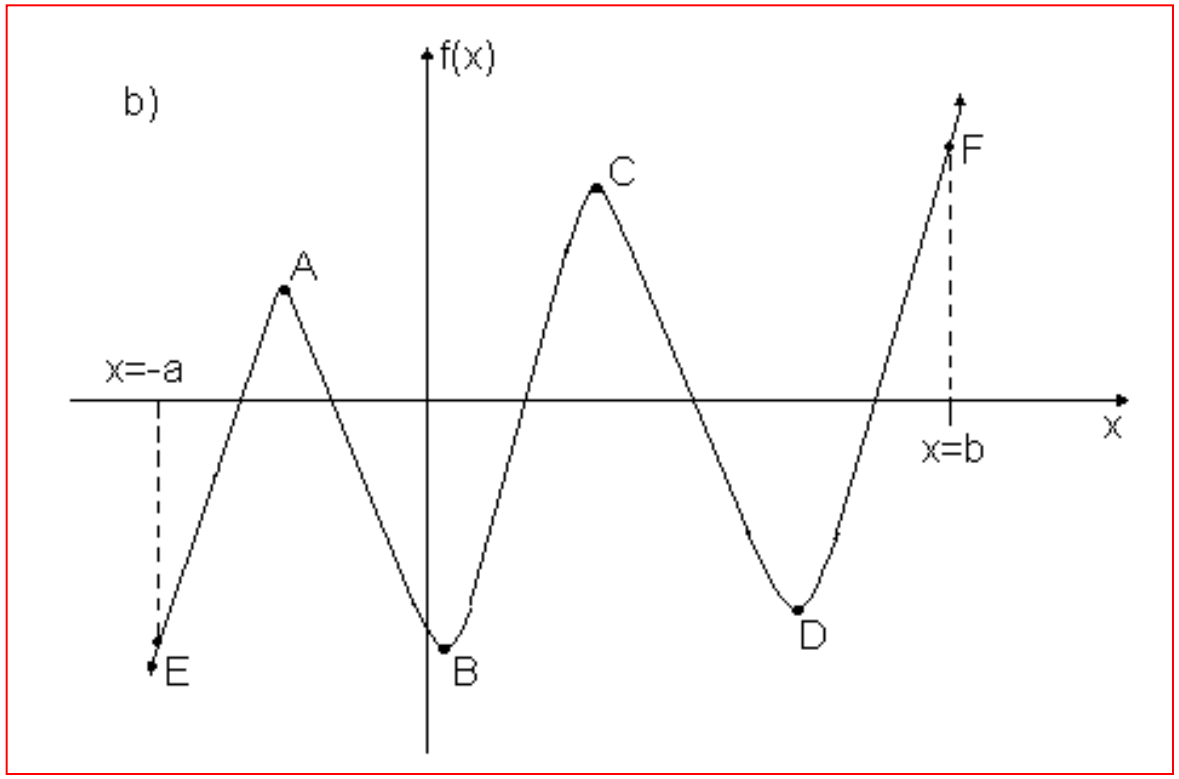
Uma função f(x) de n variáveis tem um mínimo local no ponto x* se a desigualdade:

$$f\left(x^*\right) \leq f\left(x\right)$$

for satisfeita para todos os valores de x só na vizinhança n de x* na região provável.



Domínio e função não limitados (óptimo não global)



Domínio e função limitados (existem um máximo e mínimo globais)

Programação Não Linear



Que tipos de problemas de Programação Não Linear existem?



Os problemas de programação não linear apresentam-se de muitas formas e formatos. Não existe um algoritmo único capaz de resolver todos estes tipos diferentes de problemas.



Programação Não Linear

Em vez de um algoritmo, foram desenvolvidos algoritmos para várias classes de problemas de programação não linear. As classes mais comuns são:

- Optimização Irrestrita;
- Optimização Linearmente Restrita;
- Programação Quadrática;
- Programação Convexa;
- Programação Separável;
- Programação Não-Convexa;
- Programação Geométrica;
- Programação Fraccionária.

Optimização Irrestrita



Quais são os problemas de Optimização Irrestrita?



Os problemas de Optimização irrestrita são aqueles que não possuem restrições de modo que o objectivo seja simplesmente :

$$Maximizar\ f(x),$$

Ao longo de todos os valores de $\boldsymbol{x} = (x_1, x_2, \ldots x_n)$. estes problemas não aparecem frequentemente nas aplicações práticas de engenharia. Contudo vai-se aqui referencia-los porque os problemas com restrições são uma extensão dos problemas deste tipo. **Também uma estratégia de solução de problemas com restrições é transforma-los numa sequência de problemas sem restrições.**

As condições óptimas

As condições óptimas para um problema com ou sem restrições podem ser utilizadas de duas formas:

1. As condições óptimas podem ser usadas para arranjar pontos candidatos a óptimo.
2. Para um ponto dado do projecto, através das condições óptimas, pode verificar-se se é ou não um candidato a óptimo.

Procedimentos para Estabelecer as Condições Óptimas (I)

As condições óptimas podem ser usadas para determinar os pontos mínimos de uma determinada função f(x).

O procedimento para estabelecer as condições para um mínimo local é assumir que está-se num mínimo, no ponto x* e daí examinar a vizinhança para estudar as propriedades da função e das suas derivadas. Se examinar-se só uma pequena vizinhança do ponto, as condições obtidas são chamadas locais.

Procedimentos para Estabelecer as Condições Óptimas (II)

Seja x* um ponto mínimo local da função f(x). Para investigar a sua vizinhança, seja x um ponto perto de x*. Defina-se o incremento d e Δf em x* e em f(x*) respectivamente, como:

$$d = x - x^* \qquad \text{e} \qquad \Delta f = f(x) - f(x^*)$$

Desde que f(x) seja um mínimo local no ponto x*, a alteração da função na vizinhança de x* não pode ser negativa, isto é, o valor da função pode permanecer constante ou aumentar, mas nunca diminuir.

Condições Óptimas para Funções de uma Única Variável

Condições Necessárias de Primeira Ordem (I):

A série de Taylor de f(x) no ponto x* é dada por:

$$f(x) = f(x^*) + f'(x^*)d + \tfrac{1}{2} f''(x^*)d^2 + R$$

Como: $\Delta f = f(x) - f(x^*) \geq 0$

Δf tem de ser não negativa para que x* seja um mínimo local, então:

$$\Delta f = f'(x^*)d + \tfrac{1}{2} f''(x^*)d^2 + R$$

Desde que d seja pequeno, o termo de primeira ordem domina os outros.

Condições Óptimas para Funções de uma Única Variável

Condições Necessárias de Primeira Ordem (II):

A única hipótese de Δf não ser negativo para todos os d na vizinhança de x* é quando:

$$f'\left(x^*\right) = 0$$

Esta equação, são as condições de primeira ordem, necessárias para a ocorrência de um mínimo local de f(x) no ponto x*. São chamadas de primeira ordem porque só envolvem primeiras derivadas da função.

Condições Óptimas para Funções de uma Única Variável

Condições Suficientes (I)

Desde que um ponto estacionário satisfaça a condição necessária, as alterações na função Δf transformam-na em:

$$\Delta f = \frac{1}{2} f''\left(x^*\right) d^2 + R$$

Como o termo de segunda ordem domina todos os outros viremos todas as atenções para ele. Note-se que o termo será positivo para todos os d≠0, se:

$$f''\left(x^*\right) > 0$$

Condições Óptimas para Funções de uma Única Variável

Condições Suficientes (II)

Os pontos estacionários que satisfazem a desigualdade

$$f'\left(x^*\right) = 0$$

devem ser um mínimo local porque satisfazem a desigualdade

$$f''\left(x^*\right) > 0$$

isto é, a função tem uma curvatura positiva no ponto mínimo.

A desigualdade que se segue, é então suficiente para que x* seja um mínimo local.

$$f''\left(x^*\right) > 0$$



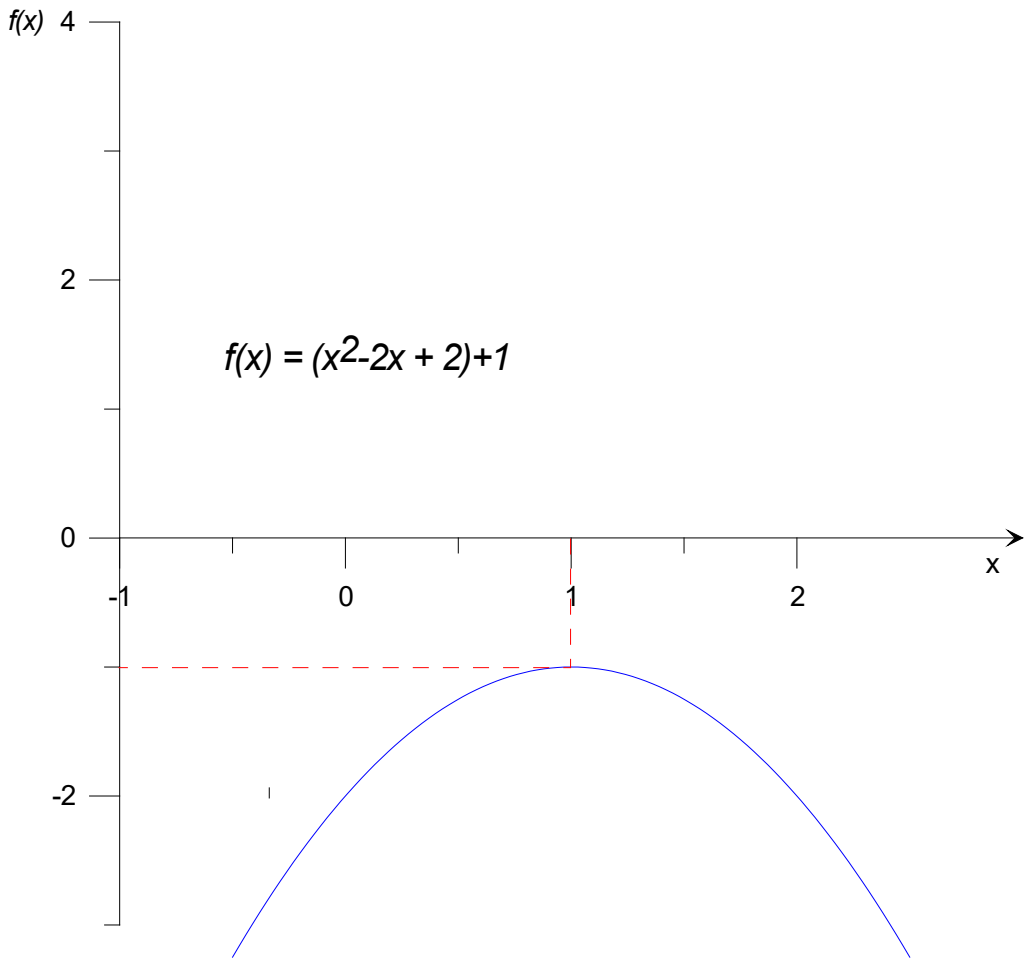
Método da Bissecção

Quando se aplica o Método da Bissecção?



Este procedimento de busca sempre pode ser aplicado quando f(x) for côncava (de modo que a segunda derivada seja negativa ou zero para todo o x) conforme se mostra na figura. Ele também pode ser usado para algumas outra funções. Em particular se x* representar a solução óptima, tudo o que é necessário é que:

Condições Óptimas para Funções de uma Única Variável sem restrições



Método da Bissecção

$$\frac{df(x)}{dx} > 0 \; se \; x < x^*$$

$$\frac{df(x)}{dx} = 0 \; se \; x = x^*$$

$$\frac{df(x)}{dx} < 0 \; se \; x > x^*$$

Estas condições são automaticamente satisfeitas quando f(x) for côncava, mas elas também podem ser satisfeitas quando a segunda derivada for positiva para alguns (mas não todos) valores de x.

Método da Bissecção

A ideia por trás do método de Bissecção é muito intuitiva, isto é, seja a inclinação (derivada) positiva ou negativa em uma solução experimental indica definitivamente se a melhoria encontra-se imediatamente à direita ou à esquerda, respectivamente. Logo se a derivada calculada em dado valor de x for positiva, então x* deve ser maior que esse e portanto esse x se torna um limite inferior para as soluções experimentais que precisam de ser consideradas de ai em diante.

Método da Bissecção

Ao contrário se a derivada for negativa, então x* tem de ser menor que esse x e, por isso, se tornaria um limite superior. Por essa razão após ambos os tipos de limites serem identificados, cada nova solução experimental seleccionada entre os limites actuais fornece novo limite mais apertado de um tipo e assim, imitando mais a busca.

Método da Bissecção

- $x' \rightarrow$, *designa* ***Solução Experimental Actual.***
- $\underline{x} \rightarrow$, *designa o* ***Limite Inferior actual sobre x*.***
- $\overline{x} \rightarrow$, *designa o* ***Limite Superior actual sobre x****
- $\varepsilon \rightarrow$, *designa a* ***Tolerância de Erro para x*.***

Método da Bissecção

Inicialização: Seleccione a tolerância, encontre um limite inferior actual e um limite superior actual. Seleccione uma solução experimental inicial

$$x' = \frac{\underline{x} + \overline{x}}{2}$$

$$Iteração:$$

$$1.\ calcular\ \frac{df(x)}{dx} em\ x = x'$$

$$2.\ se \frac{df(x)}{dx} \geq 0\ reinicialize\ \underline{x} = x'$$

Método da Bissecção

3. $se \dfrac{df(x)}{dx} \leq 0$ *reinicialize* $\overline{x} = x'$

4. *Selecione uma nova* $x' = \dfrac{\underline{x} + \overline{x}}{2}$

Regra de Paragem: *Se* $\overline{x} - \underline{x} \leq 2\varepsilon$,

De modo que o novo **x´** deva estar dentro da tolerância **ε** de **x***, pare. Caso contrário, executa-se nova iteração.

Método da Bissecção

Maximizar a função $f(x)=12x-3x^4-2x^6$

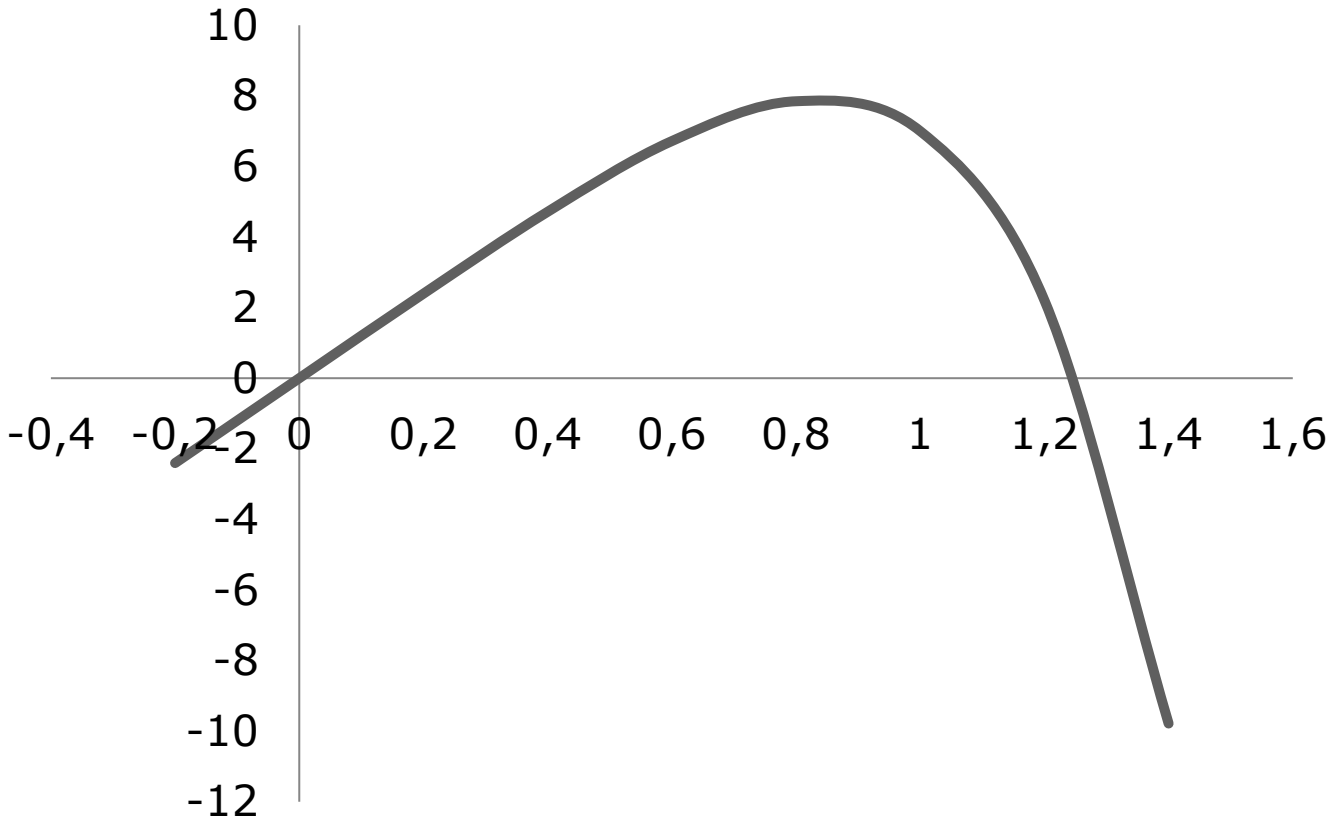
As duas primeiras derivadas são dadas por:

$$\frac{df(x)}{dx} = 12\left(1 - x^3 - x^5\right)$$

$$\frac{df^2(x)}{dx} = 12\left(-3x^2 - 5x^4\right)$$

Pelo facto da segunda derivada ser não positiva em qualquer ponto, **f(x)** é uma função côncava e portanto o método da bissecção pode ser aplicado tranquilamente para encontrar o seu máximo global (supondo-se que exista um máximo global).

Método da Bissecção



Método da Bissecção

Iteração	df(x)/dx	$\underline{x}$	$\overline{x}$	Novo x´	f(x´)
0		0	2	1	7,000
1	-12	0	1	0,5	5,7812
2	+10,12	0,5	1	0,75	7,6948
3	+4,09	0,75	1	0,875	7,8439
4	-2,19	0,75	0,875	0,8125	7,8672
5	+1,31	0,8125	0,875	0,84375	7,8829
6	-0,34	0,8125	0,84375	0,828125	7,8815
7	+0,51	0,828125	0,84375	0,8359375	7,8839
Parar					

Método de Newton

Embora o método de Bissecção seja um procedimento intuitivo e simples, ele apresenta a desvantagem de convergir de forma relativamente lenta para uma soluções óptima. Cada iteração diminui apenas pela metade a diferença entre os limites.

A razão básica para essa convergência lenta é o facto da única informação sobre a $f(x)$ que está sendo empregue ser o valor da primeira derivada $f'(x)$ nos respectivos valores experimentais de x. Informações úteis adicionais podem ser obtidas considerando a segunda derivada $f''(x)$, que é isso que o Método de Newton faz.

Método de Newton

O conceito básico do Método de Newton é aproximar f(x) nas vizinhanças da solução experimental actual por uma função quadrática e, em seguida, maximizar (ou minimizar) a função aproximada, exactamente para obter a nova solução experimental para iniciar a iteração seguinte. Essa ideia de trabalhar com uma aproximação quadrática da função objectivo tornou-se, a partir de então um recurso fundamental de vários algoritmos para tipos de problemas de programação não-linear mais genéricos.

Método de Newton

Essa função quadrática aproximada obtém-se truncando a série de Taylor após o termo de segunda derivada. Particularmente, fazendo-se xi+1 a solução experimental gerada na iteração i para iniciar a iteração i+1 (de modo que xi seja a solução experimental inicial fornecida pelo utilizador para começar a iteração 1), a série de Taylor truncada fica:

$$f\left(x_{i+1}\right) \approx f\left(x_i\right) + f'\left(x_i\right)\left(x_{i+1} - x_i\right) + \tfrac{1}{2} f''\left(x^*\right)\left(x_{i+1} - x_i\right)^2$$

Método de Newton

Esta função quadrática pode ser maximizada da maneira usual fazendo com que a sua primeira derivada seja igual a zero e executando a resolução para (x_{i+1}),

$$\frac{f(x_{i+1})-f(x_i)}{(x_{i+1}-x_i)}=\frac{f'(x_i)(x_{i+1}-x_i)}{(x_{i+1}-x_i)}+\frac{1}{2}\frac{f''(x^*)(x_{i+1}-x_i)^2}{(x_{i+1}-x_i)}$$

$$f'(x_{i+1})=f'(x_i)+f''(x^*)(x_{i+1}-x_i)$$

Método de Newton

Já que x_i, $f(x_i)$, $f'(x_i)$ e $f''(x_i)$ são constantes, fazer com que a primeira derivada a direita seja igual a zero leva a que:

$$f'(x_{i+1}) + f''(x_i)(x_{i+1} - x_i) = 0$$

O que conduz directa e algebricamente à solução:

$$x_{i+1} = x_i - \frac{f'(x_i)}{f''(x_i)}$$

Esta é a formula usada a cada iteração i para calcular a solução experimental seguinte, x_{i+1}.

Método de Newton

Inicialização: Seleccione uma solução inicial experimental x_i, por inspecção. Faça que i=1.

Iteração i :

1. *calcular* $\frac{df(x_i)}{dx}$ *e* $\frac{d^2 f(x_i)}{dx^2}$ *calcular* $f(x_i)$ *é opcional*

2. *configure* $x_{i+1} = x_i - \frac{f'(x_i)}{f''(x_i)}$

Regra da Paragem: se $|x_{i+1}-x_i| \leq \varepsilon$ **pare** x_{i+1} é essencialmente a solução óptima. Caso contrário, reinicialize $i=i+1$ e execute uma outra iteração.

Método de Newton

Voltemos ao exemplo usado no Método de Bissecção Maximizar a função $f(x)=12x-3x^4-2x^6$

As duas primeiras derivadas são dadas por:

$$\frac{df(x)}{dx} = 12\left(1-x^3-x^5\right)$$

$$\frac{df^2(x)}{dx} = -12\left(3x^2+5x^4\right)$$

Portanto, a fórmula para calcular a nova solução experimental (x_{i+1}) a partir da actual x_i é:

$$x_{i+1} = x_i - \frac{f'(x_i)}{f''(x_i)} = x_i - \frac{12\left(1-x^3-x^5\right)}{-12\left(3x^2+5x^4\right)} = x_i + \frac{1-x^3-x^5}{3x^2+5x^4}$$

Método de Newton

Após seleccionar-se $\varepsilon=0{,}00001$, e escolher-se $x_i=1$ como solução experimental inicial pode-se ver na tabela seguinte as soluções restantes:

Iteração i	x_i	$f(x_i)$	$f'(x_i)$	$f''(x_i)$	x_{i+1}
1	1	7	-12	-96	0,875
2	0,875	7,8439	-2,1940	-62,733	0,84003
3	0,84003	7,8838	-0,1325	-55,279	0,83763
4	0,83763	7,8839	-0,0006	-54,790	0,83762

Uma comparação dos Métodos de Bissecção e de Newton mostram como o método de Newton converge muito mais rapidamente que o de Bissecção. Seriam necessárias 20 iterações para o método de Bissecção convergir com o mesmo grau de precisão alcançado com o de Newton.

Aplicações à Engenharia Mecânica

Os conceitos desta aula (condições óptimas, Método da Bissecção e Método de Newton) aplicam-se directamente a problemas de projecto mecânico com uma única variável de decisão:

1. **Viga de máxima resistência**
   Secção rectangular serrada de um tronco circular
   Condições óptimas (analítico)

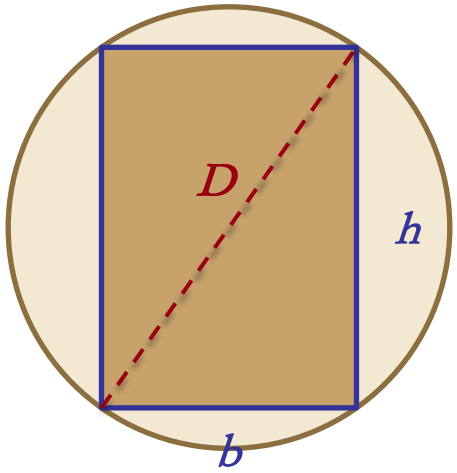
2. **Espessura de isolamento**
   Custo do isolamento vs. custo da energia perdida
   Método de Newton

3. **Ciclo Brayton**
   Razão de pressões de máximo trabalho líquido
   Método da Bissecção

4. **Velocidade de corte**
   Custo mínimo por peça em torneamento (Taylor)
   Método de Newton

Exemplo 1: Viga de Máxima Resistência

De um tronco de diâmetro **D = 300 mm** pretende-se serrar uma viga de secção rectangular b × h com o **maior módulo de resistência à flexão** W (máxima carga para a tensão admissível).



Maximizar $W = b \cdot h^2/6$ sujeito a $b^2 + h^2 = D^2$

Com $h^2 = D^2 - b^2$: $W(b) = b(D^2 - b^2)/6$

1.ª ordem: $dW/db = (D^2 - 3b^2)/6 = 0 \Rightarrow b^* = D/\sqrt{3} =$ **173,2 mm**

Suficiente: $d^2W/db^2 = -b^* < 0 \Rightarrow$ máximo

$h^* = \sqrt{2} \cdot b^* =$ **244,9 mm** $(h/b = \sqrt{2})$; $W_{max} = 1{,}732 \times 10^6$ mm³

Exemplo 2: Isolamento Térmico

Uma conduta de vapor a 175 °C, num ambiente a 25 °C ($\Delta T = 150$ K), é tratada como parede plana de 1 m². Sem isolamento, $R_0 = 0{,}1$ m²·K/W. Isolamento: $k = 0{,}04$ W/(m·K), espessura t (m).

Custo anual do isolamento: $1500 \cdot t$ u.m./ano

Custo anual da energia: 0,05 u.m./kWh × 8000 h = 0,4 u.m./(W·ano) × $q(t)$, $q = \Delta T/(R_0 + t/k)$

Minimizar **$C(t) = 1500\,t + 60/(0{,}1 + 25\,t)$**

$C'(t) = 1500 - 1500/(0{,}1 + 25\,t)^2$

$C''(t) = 75\,000/(0{,}1 + 25\,t)^3 > 0 \Rightarrow$ $C(t)$ é convexa (mínimo único)

Newton: $t_{i+1} = t_i - C'(t_i) / C''(t_i)$

Exemplo 2: Método de Newton

Com $t_1 = 0{,}01$ m (10 mm) e $\varepsilon = 0{,}00001$ m obtém-se:

i	t_i (m)	$C(t_i)$	$C'(t_i)$	$C''(t_i)$	t_{i+1} (m)
1	0,01000	186,43	−10 744,9	1 749 271	0,01614
2	0,01614	143,37	−4 415,4	587 356	0,02366
3	0,02366	122,26	−1 637,0	226 823	0,03088
4	0,03088	115,13	−473,0	113 144	0,03506
5	0,03506	114,03	−73,2	80 560	0,03597
6	0,03597	114,00	−2,48	75 186	0,03600
7	0,03600	114,00	−0,003	75 000	0,03600

t* = 36 mm; C_{min} = 114 u.m./ano; q = 150 W/m² (1500 W/m² sem isolamento). Mais espessura já não se paga: o custo do material cresce mais do que a economia de energia.

Exemplo 3: Ciclo Brayton

Turbina a gás (ciclo Brayton ideal, ar-padrão): $T_1 = 300$ K à entrada do compressor e $T_3 = 1200$ K à entrada da turbina. Qual a razão de pressões r_p que **maximiza o trabalho líquido** específico?

$$\mathbf{w(r_p) = c_p T_1 \left[\tau(1 - r_p^{-a}) - (r_p^{a} - 1) \right]}$$

$\tau = T_3/T_1 = 4$; $a = (k-1)/k = 0{,}2857$; $c_p = 1{,}005$ kJ/(kg·K)

$$dw/dr_p = c_p T_1 \cdot a \cdot [\, \tau \cdot r_p^{-a-1} - r_p^{a-1} \,]$$

$d^2w/dr_p^2 < 0$ para $r_p < 31{,}6 \Rightarrow$ w é côncava em [1; 31] $\Rightarrow$ Bissecção aplicável

Verificação analítica: $dw/dr_p = 0 \Rightarrow r_p^* = \tau^{1/(2a)} = 4^{1,75} = 11{,}31$

Exemplo 3: Método da Bissecção

Limites iniciais $r_p = 1$ e $r_p = 31$; tolerância $\varepsilon = 0{,}25$:

Iter.	dw/dr_p	Lim. inf.	Lim. sup.	Novo r_p'	$w(r_p')$ kJ/kg
0		1	31	16	295,58
1	−2,14	1	16	8,5	297,47
2	+3,32	8,5	16	12,25	301,19
3	−0,64	8,5	12,25	10,375	301,13
4	+0,82	10,375	12,25	11,3125	301,50
5	+0,001	11,3125	12,25	11,78125	301,42
6	−0,34	11,3125	11,78125	11,546875	301,48
Parar					—

$r_p^* \approx 11{,}5$ (exacto 11,31); $w_{max} \approx 301{,}5$ kJ/kg. A curva é plana junto ao óptimo: entre $r_p = 8$ e 16 o trabalho cai apenas ~2 %.

Exemplo 4: Velocidade de Corte

Torneamento de uma peça com comprimento de corte equivalente K = 300 m (tempo de maquinagem $t_m = K/V$). Vida da ferramenta pela **equação de Taylor**: $V \cdot T^n = C$, com n = 0,25 e C = 400 m/min.

Máquina + operador c_0 = 2 u.m./min; troca de ferramenta t_c = 2 min; custo por aresta c_t = 5 u.m.

Custo por peça = maquinagem + (troca + aresta) × n.º de trocas por peça:

$$c(V) = c_0 \cdot K/V + (c_0 t_c + c_t) \cdot (K/V) \cdot (V/C)^{1/n}$$

$$\mathbf{c(V) = 600/V + 2700 \cdot V^3/400^4}$$

V baixo → maquinagem lenta; V alto → ferramenta gasta-se depressa

Exemplo 4: Método de Newton

$c'(V) = -600/V^2 + 8100 \cdot V^2/400^4$

$c''(V) = 1200/V^3 + 16\,200 \cdot V/400^4 > 0$

Com $V_1 = 100$ m/min e $\varepsilon = 0{,}01$ m/min:

i	V_i	$c(V_i)$	$c'(V_i)$	$c''(V_i)$	V_{i+1}
1	100,000	6,1055	$-5{,}684\times10^{-2}$	$1{,}263\times10^{-3}$	144,991
2	144,991	4,4597	$-2{,}189\times10^{-2}$	$4{,}854\times10^{-4}$	190,082
3	190,082	3,8809	$-5{,}174\times10^{-3}$	$2{,}950\times10^{-4}$	207,620
4	207,620	3,8338	$-2{,}800\times10^{-4}$	$2{,}655\times10^{-4}$	208,675
5	208,675	3,8337	$-7{,}186\times10^{-7}$	$2{,}641\times10^{-4}$	208,678

V* = 208,7 m/min; vida da ferramenta $T^* = (C/V^*)^4 = 13{,}5$ min; $c_{min} = 3{,}83$ u.m./peça.

Verificação: $V^* = C \cdot [n \cdot c_0 / ((1-n)(c_0 t_c + c_t))]^n = 208{,}7$ m/min

Síntese dos Exemplos

Exemplo	Variável	Objectivo	Método	Óptimo
1. Viga	b (mm)	máx. W	Cond. óptimas	173,2 mm
2. Isolamento	t (m)	mín. custo anual	Newton (7 it.)	36 mm
3. Brayton	r_p	máx. trabalho	Bissecção (6 it.)	≈ 11,5
4. Corte	V (m/min)	mín. custo/peça	Newton (5 it.)	208,7 m/min

Em todos os casos há um **compromisso** entre dois efeitos opostos — é isso que cria um óptimo interior.

Exercício proposto: resolva o Exemplo 2 pelo Método da Bissecção (limites 0 e 0,1 m, $\varepsilon = 0{,}0005$ m) e compare o número de iterações com o do Método de Newton.