



Optimização

Aula 27



Programação Não Linear

Aula 27: Programação Não-Linear - Funções de Várias variáveis

Tópicos:

- Compreender os fundamentos da optimização sem restrições
- Introduzir ferramentas matemáticas: Gradiente, Hessiana e Convexidade
- Aplicar métodos analíticos e numéricos para encontrar pontos óptimos

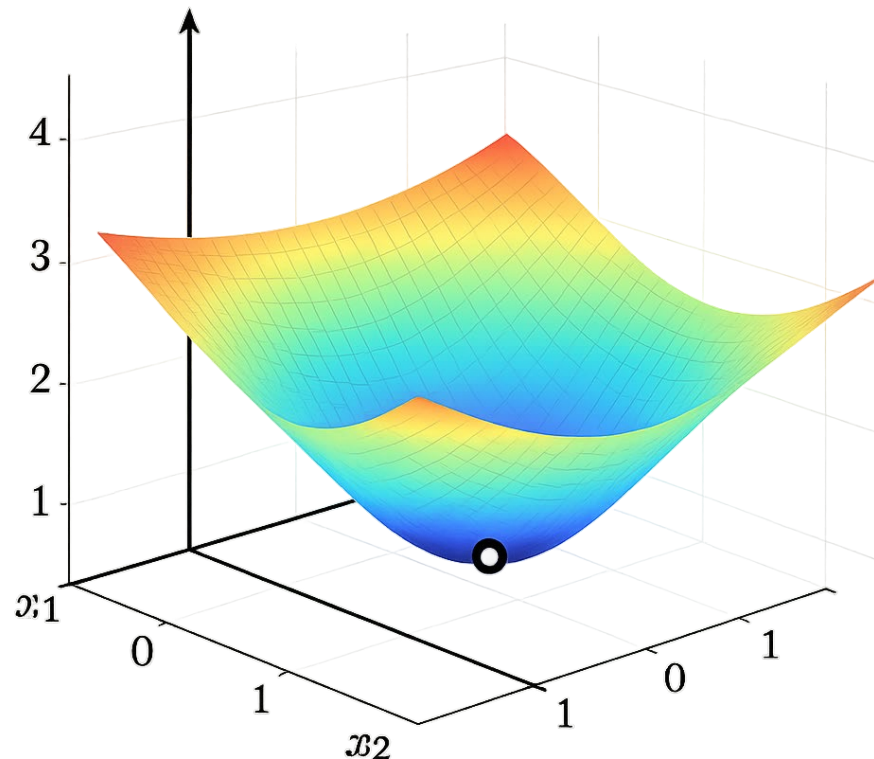


Programação Não Linear



Aplicações:

- Engenharia de sistemas energéticos
- Modelação matemática de processos complexos
- Optimização multidimensional de recursos





Programação Não Linear

Por que estudar Programação Não Linear?

Motivação:

- Problemas reais frequentemente não seguem relações lineares
- Funções com múltiplas variáveis exigem técnicas específicas

Exemplos típicos:

- Maximizar eficiência energética
- Minimizar perdas industriais
- Ajustar modelos de previsão

Desafio central:

- Como identificar e validar pontos óptimos sem restrições explícitas?



Programação Não Linear

Quais as Ferramentas Matemáticas Essenciais?

Vector Gradiente (∇f):

- Direcção de maior crescimento da função
- Usado para localizar pontos estacionários

Matriz Hessiana (H):

- Avalia curvatura da função
- Crucial para testar convexidade e natureza dos óptimos

Convexidade:

- Determina unicidade e existência de mínimos
- Garante convergência dos métodos de optimização

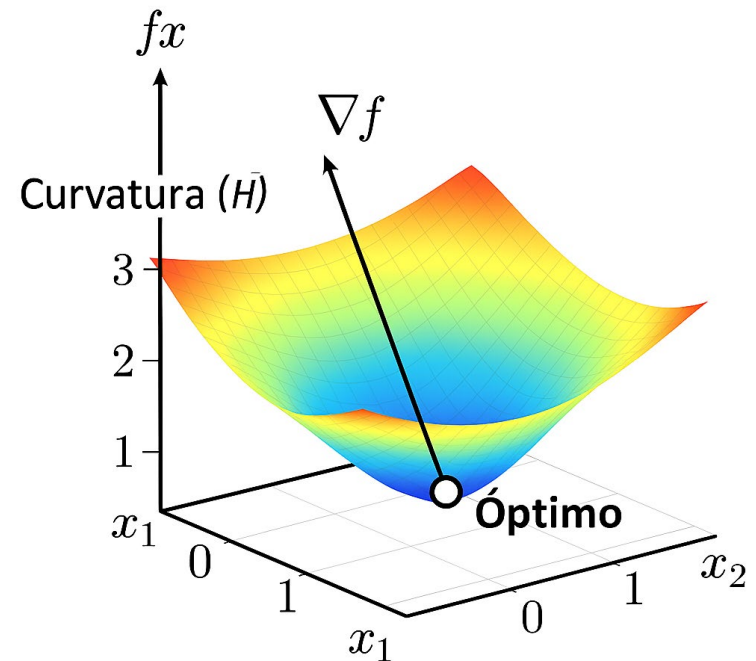


Programação Não Linear



A direcção de maior crescimento
da função (**Gradiente**)

A curvatura da superfície (**via
Hessiana**)





Programação Não Linear



Considere-se o problema de maximização de uma função $f(\mathbf{x})$ com variáveis múltiplas $\mathbf{x} = (x_1, x_2, \dots, x_n)$ quando não existe nenhuma restrição sobre os valores viáveis. A condição de optimalidade consegue-se igualando as primeira e segunda derivada a zero.

Como a função agora tem várias variáveis, a primeira e segunda derivadas alcança-se introduzindo o conceito de vector Gradiente e matriz Hessiana



Vector Gradiente

O que é o vector gradiente?

Chama-se gradiente de f no ponto x a um vector cujas componentes são as derivadas parciais de f em x . Pode representar-se por ***grad* $f(x)$** , ou por $\nabla f(x)$, em que ∇ (lê-se *nabla*) e é o operador diferencial definido por:

$$\nabla f(x^*) = \begin{bmatrix} \frac{\partial f(x^*)}{\partial x_1} \\ \frac{\partial f(x^*)}{\partial x_2} \\ \vdots \\ \frac{\partial f(x^*)}{\partial x_n} \end{bmatrix} = \left[\frac{\partial f(x^*)}{\partial x_1} \frac{\partial f(x^*)}{\partial x_2} \dots \frac{\partial f(x^*)}{\partial x_n} \right]^T$$



Matriz Hessiana

O que é a matriz Hessiana?

Chama-se Matriz Hessiana de uma função de **n** variáveis a matriz quadrada $n \times n$ das derivadas parciais de segunda ordem da função:

$$\frac{\partial^2 f}{\partial x \partial x} = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$



Convexidade



O que é a Convexidade de uma função com uma única variável?

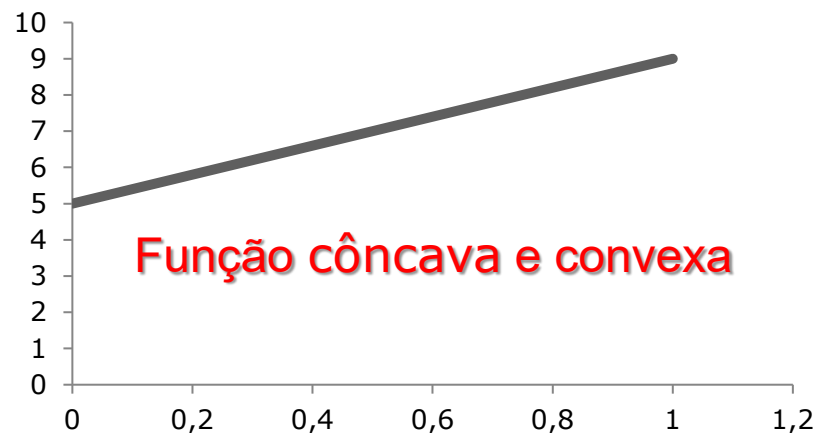
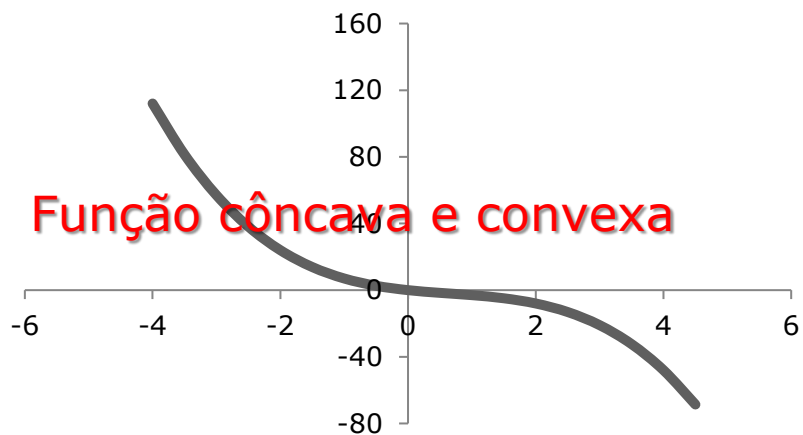
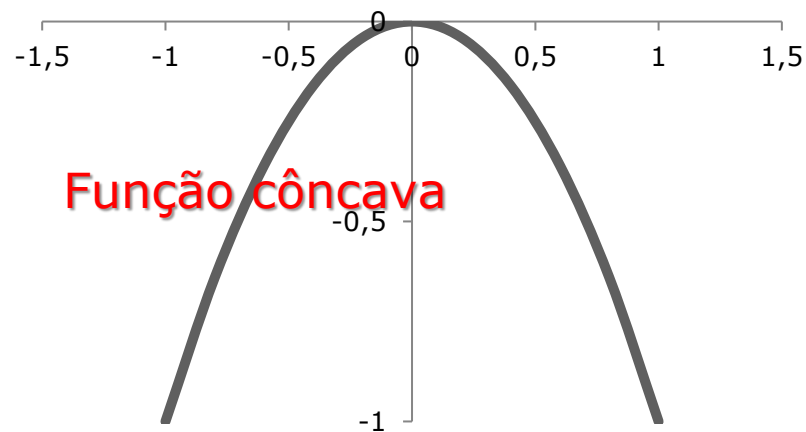
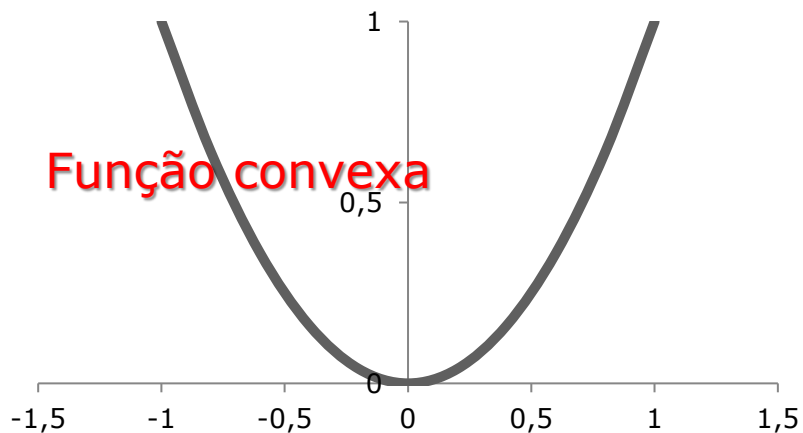
Uma função com uma única variável $f(x)$ é uma função convexa se para cada um dos pares de valores de x' e x'' ($x' < x''$)

$$f[\lambda x'' + (1 - \lambda)x'] \leq \lambda f(x'') + (1 - \lambda)f(x')$$

para todos os valores de λ tais que $0 < \lambda < 1$. Ela será uma função estritamente convexa se $\leq$ puder ser substituído por $<$. Ela é uma função côncava (ou uma função estritamente côncava) se essa afirmação continuar válida quando $\leq$ for substituído por $\geq$ (ou por $>$)



Convexidade





Convexidade



Como se faz o teste de Convexidade de uma função com uma única variável?

Considere-se uma função com uma única variável $f(x)$ que possui uma segunda derivada em todos os possíveis valores de x . Então $f(x)$ é:



Convexa se e somente se para todos os x

$$\frac{d^2 f}{dx^2} \geq 0$$



Estritamente Convexa se e somente se para todos os x

$$\frac{d^2 f}{dx^2} > 0$$



Convexidade



Côncava se e somente se para todos os x

$$\frac{d^2 f}{dx^2} \leq 0$$



Estritamente Côncava se e somente se para todos os x

$$\frac{d^2 f}{dx^2} < 0$$



É de notar que uma função estritamente convexa também é convexa, mas uma função convexa não é estritamente convexa se a segunda derivada for igual a zero para alguns valores de x . de forma similar uma função estritamente côncava é côncava porém o inverso não é necessariamente verdadeiro.



Convexidade

O que é a Convexidade de uma função com várias variáveis?

Uma função com várias variáveis $f(x_1, x_2, \dots, x_n)$ é uma função **convexa** se para cada pares de pontos na curva $f(x_1, x_2, \dots, x_n)$, o segmento de recta que une esses dois pontos cai inteiramente em cima ou sobre a curva de $f(x_1, x_2, \dots, x_n)$. Ela é uma função **estritamente convexa** se esse segmento de recta realmente cair inteiramente acima dessa curva excepto nos pontos extremos do segmento de recta. As funções **côncavas** e **estritamente côncavas** são definidas da mesma forma, excepto pelo facto de acima ser substituído por abaixo.



Convexidade

Como se faz o teste de Convexidade de uma função com várias variáveis?

Uma função de n variáveis $f(x_1, x_2, \dots, x_n)$ definida num conjunto convexo S é:

Convexa se e somente se a matriz Hessiana da função for positiva semi-definida em todos os pontos do conjunto S .

Estritamente Convexa se e somente se a matriz Hessiana da função for positiva definida em todos os pontos do conjunto S .



Convexidade



Concava se e somente se a matriz Hessiana da função for negativa semi-definida em todos os pontos do conjunto S .

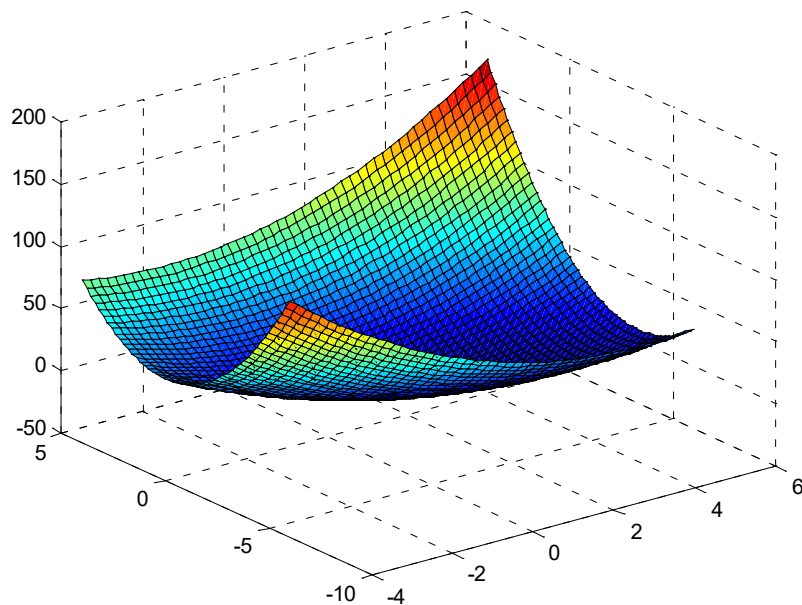


Estritamente Concava se e somente se a matriz Hessiana da função for negativa definida em todos os pontos do conjunto S .

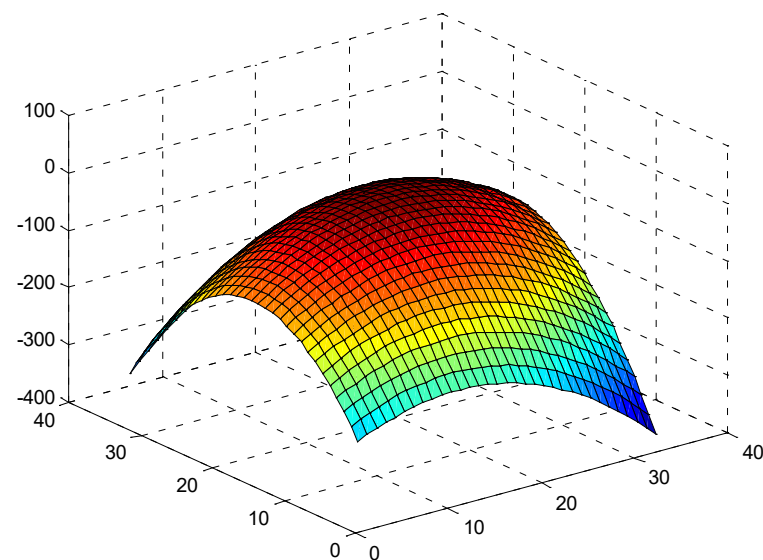


Convexidade

Função convexa



Função côncava





Convexidade

O que um conjunto convexo?

Um conjunto convexo é um conjunto de pontos tal que, para cada par de pontos no conjunto, todo o segmento de recta que une esses dois pontos também se encontra no conjunto.

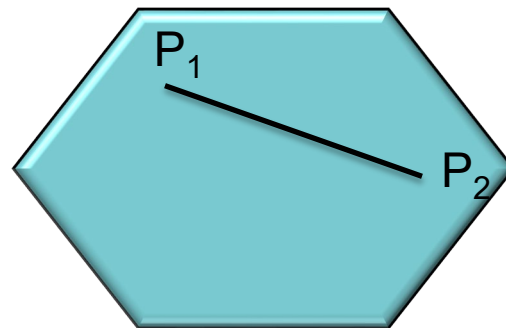
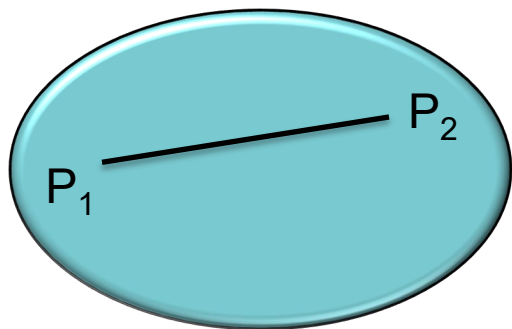
As figuras mostram alguns casos de conjuntos convexos e não convexos.



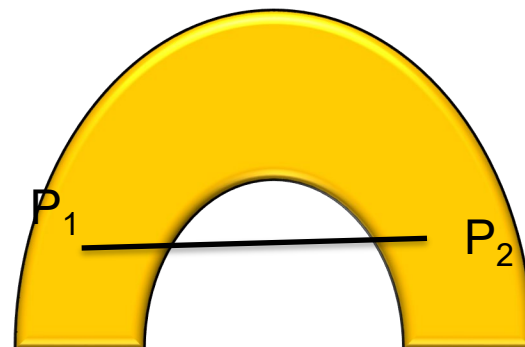
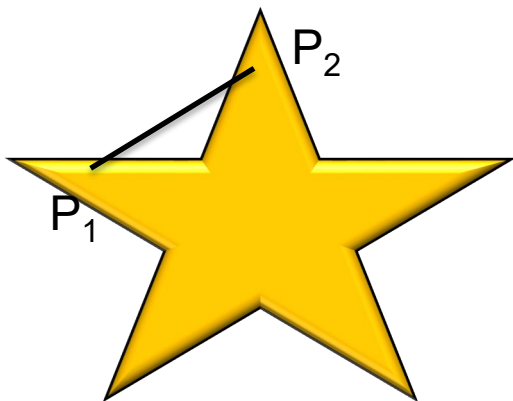


Convexidade

Conjuntos Convexos



Conjuntos não convexos





Condições Óptimas para Funções de várias Variáveis sem restrições

Seja $\mathbf{x}^*$ um ponto mínimo local da função $f(\mathbf{x})$. Para investigar a sua vizinhança, seja $\mathbf{x}$ um ponto perto de $\mathbf{x}^*$. Escrevendo a Série de Taylor em notação matricial obtém-se:

$$f(\mathbf{x}) = f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^T d + \frac{1}{2} d^T H(\mathbf{x}^*) d + R$$

Defina-se o incremento Δf em $f(\mathbf{x}^*)$ como:

$$\Delta f = \nabla f(\mathbf{x}^*)^T d + \frac{1}{2} d^T H(\mathbf{x}^*) d + R$$



Condições Óptimas para Funções de várias Variáveis sem restrições

Se assumir-se um mínimo local em x^* daí Δf não deve ser negativo, isto é $\Delta f \geq 0$.
Concentrando-se só nos termos de primeira ordem observa-se como antes que Δf pode ser não negativo para todos os d possíveis quando:

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

Isto é o gradiente da função no ponto x^* tem de ser zero. Esta condição pode ser escrita na forma:

$$\frac{\partial f(x^*)}{\partial x_i} = 0 \quad \text{para } i = 1 \dots n$$



Condições Óptimas para Funções de várias Variáveis sem restrições

Os pontos que satisfazem a equação anterior são chamados pontos estacionários. Considerando o segundo termo na equação de Taylor avaliado para um ponto estacionário, a positividade de Δf está assegurada se:

$$d^T H(x^*) d > 0$$

para todos os $d \neq 0$. Isso será verdadeiro se a Hessiana $H(x^*)$ for uma matriz positiva definida que será a condição suficiente para um mínimo local de $f(x)$ no ponto x^* .



Condições Óptimas para Funções de várias Variáveis sem restrições

Teorema: Condições necessárias e suficientes para que haja um mínimo local. Se $f(\mathbf{x})$ for um mínimo local no ponto $\mathbf{x}^*$ então:

$$\frac{\partial f(\mathbf{x}^*)}{\partial x_i} = 0 \quad \text{para } i = 1 \dots n$$

Condições necessárias de segunda ordem.

Se $f(\mathbf{x})$ for um mínimo local no ponto $\mathbf{x}^*$, daí a matriz Hessiana:

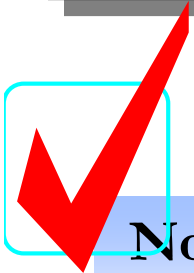
$$H(\mathbf{x}^*) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j} \right]_{(n \times n)}$$

terá de ser semi-definida positiva ou definida positiva
no ponto $\mathbf{x}^*$



Condições Óptimas para Funções de várias Variáveis sem restrições

Condições suficientes de segunda ordem: Se a matriz $H(\mathbf{x}^*)$ for definida positiva no ponto estacionário $\mathbf{x}^*$, daí $\mathbf{x}^*$ é um ponto mínimo local para a função $f(\mathbf{x})$.



Nota: As condições envolvem a derivada de $f(\mathbf{x})$ e não o valor da função.



Forma da matriz (I)

Teorema: Forma da matriz através de autovalores

Seja λ_i , $i = 1$ até n , autovalores de uma matriz simétrica H $n \times n$, associada à forma quadrática pode-se determinar a partir dos seguintes resultados da forma quadrática ou da matriz A .



Forma da matriz (II)



$F(x)$ é positiva definida se e somente se, todos os autovalores de H forem positivos, isto é: $\lambda_i > 0$, $i = 1$ até n .



$F(x)$ é positiva semi-definida se e só se os autovalores de H forem não negativos, isto é: $\lambda_i \geq 0$, $i = 1$ até n .



$F(x)$ é negativa definida se e somente se todos os autovalores de H forem negativos isto é: $\lambda_i < 0$.



$F(x)$ é negativa semi-definida se e somente se todos os autovalores de H forem não positivos, isto é: $\lambda_i \leq 0$.



$F(x)$ é indefinida se alguns valores $\lambda_i < 0$ e outros valores $\lambda_i > 0$.



Condições para que uma matriz seja Definida Positiva



Uma matriz é **definida positiva** se e só se puder ser reduzida a uma matriz triangular superior, utilizando somente as operações elementares sobre linhas (adição a qualquer linha elemento por elemento de uma outra linha multiplicada por um escalar).



Uma matriz é **definida positiva** se e só se todos os seus menores principais forem positivos.



Uma matriz é **definida positiva** se e só se os seu valores próprios (autovalores) forem positivos.



Autovalores e autovectores

Dada uma matriz A de $n \times n$, qualquer vector x de valores diferentes de zero satisfaz a equação:

$$Ax = \lambda x$$

Onde λ é um factor de escala, chamado autovector e os escalares λ são chamados autovalores. Desde que $x \neq 0$, vê-se que λ dá as raízes da equação característica:

$$|A - \lambda I| = 0$$

A última equação dá-nos o grau do polinómio de λ



Autovalores e autovectores

Depois da determinação dos autovalores, os autovectores podem ser determinados pela equação:

$$Ax = \lambda x$$

Os coeficientes da matriz A tanto podem ser simétricos como também assimétricos. Em muitas das aplicações A é uma matriz simétrica.



- Os autovalores e autovectores de uma matriz real são reais. Eles podem ser imaginários para matrizes reais não simétricas
- Os autovectores correspondentes a distintos autovalores de matrizes reais simétricas, são ortogonais



Autovalores e autovectores

Exemplo 27.1



Calcule os autovalores e os autovectores da matriz:

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

Os autovalores do problema são definidos como:

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

A característica polinomial é dada por:

$$|A - \lambda I| = 0$$



Autovalores e autovectores

Exemplo 27.1

Então:

$$\begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix} = 0 \quad \text{ou}$$

$$\lambda^2 - 4\lambda + 3 = 0$$

As raízes deste polinómio são: $\lambda_1 = 3$, $\lambda_2 = 1$, daí os autovalores são 3 e 1

Os autovectores para $\lambda_1 = 3$, dados pela primeira equação são:

$$\begin{bmatrix} (2-3) & 1 \\ 1 & (2-3) \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{Daí:}$$

$$x^{(1)} = \left(1/\sqrt{2}\right) \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$



Autovalores e autovectores

Exemplo 27.1

Os autovectores para $\lambda_2 = 1$, dados pela primeira equação são:

$$\begin{bmatrix} (2-1) & 1 \\ 1 & (2-1) \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Daí:

$$x^{(2)} = \left(1/\sqrt{2}\right) \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$



Método da matriz triangular Exemplo 27.2



Dada a matriz A de dimensão 3×3 determine a sua forma.

$$A = \begin{bmatrix} 6 & 2 & -2 \\ 2 & 6 & -2 \\ -2 & -2 & 10 \end{bmatrix}$$



Método da matriz triangular

Exemplo 27.2

$$A = \begin{bmatrix} 6 & 2 & -2 \\ 0 & 16/3 & -4/3 \\ -2 & -2 & 10 \end{bmatrix}$$

Adicionando a primeira linha multiplicada por $-1/3$ à segunda

$$A = \begin{bmatrix} 6 & 2 & -2 \\ 0 & 16/3 & -4/3 \\ 0 & -4/3 & 28/3 \end{bmatrix}$$

Adicionando a primeira linha multiplicada por $1/3$ à terceira

$$A = \begin{bmatrix} 6 & 2 & -2 \\ 0 & 16/3 & -4/3 \\ 0 & 0 & 27/3 \end{bmatrix}$$

Adicionando a segunda linha multiplicada por $1/4$ à terceira

Os *pivots* **6**, **16/3** e **27/3** são todos positivos a matriz é definida positiva



Método dos determinantes

Exemplo 27.2

Os menores principais de A são:

$$\det[6] = 6$$

$$\begin{vmatrix} 6 & 2 \\ 2 & 6 \end{vmatrix} = 36 - 4 = 32$$

$$\begin{vmatrix} 6 & 2 & -2 \\ 2 & 6 & -2 \\ -2 & -2 & 10 \end{vmatrix} = 288$$

Como todos os três menores principais são positivos, a matriz é definida positiva



Método da matriz triangular Exemplo 27.3



Dada a matriz A de dimensão 3×3 determine a sua forma:

$$A = \begin{bmatrix} 2 & 10 & -2 \\ 10 & 5 & 8 \\ -2 & 8 & 11 \end{bmatrix}$$



Método da matriz triangular

Exemplo 27.3

Adicionando a primeira linha multiplicada por (-5) à segunda linha

$$A = \begin{bmatrix} 2 & 10 & -2 \\ 0 & -45 & 18 \\ -2 & 8 & 11 \end{bmatrix}$$

Como o segundo pivot, -45 é negativo, A não é definida positiva nem semi-definida positiva, pode-se pôr de parte que A seja definida negativa ou semi-definida negativa, porque o primeiro pivot 2 não é negativo.



Método da matriz triangular Exemplo 27.4



Dada a matriz A de dimensão 4×4 determine a sua forma:

$$A = \begin{bmatrix} 11 & -3 & 5 & 8 \\ -3 & 11 & -5 & -8 \\ 5 & -5 & 19 & 0 \\ -8 & -8 & 0 & 16 \end{bmatrix}$$



Método da matriz triangular Exemplo 27.4

Utilizando somente operações elementares sobre linhas:

$$A = \begin{bmatrix} 11 & -3 & 5 & -8 \\ 0 & 112/11 & -40/11 & -112/11 \\ 0 & 0 & 108/7 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Como nem todos os pivots, 11, 112/11, 108/7 e 0 são positivos, a matriz não é definida positiva. No entanto estes pivots são não negativos; assim A é semi-definida positiva.



Condições Óptimas para Funções de várias Variáveis sem restrições. Exemplo 27.5



Procurar o mínimo local para a seguinte função de duas variáveis:

$$f(x) = x_1^2 + 2x_1x_2 + 2x_2^2 - 2x_1 + x_2 + 8$$

Solução: As condições necessárias do problema dão:

$$\frac{\partial f}{\partial x} \equiv \begin{bmatrix} (2x_1 + 2x_2 - 2) \\ (2x_1 + 4x_2 + 1) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Esta equação é linear em relação as variáveis x_1 e x_2 . Se este sistema tiver solução, esta será única. Resolvendo o sistema em simultâneo obtém-se os pontos estacionários :

$$x^* = \left(\frac{5}{2}, -\frac{3}{2} \right)$$



Condições Óptimas para Funções de várias Variáveis sem restrições. Exemplo 27.5

Para saber se o ponto estacionário é um mínimo local é necessário
Avaliar a Hessiana no ponto x^*

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix}$$

H é positiva definida no ponto estacionário x^* , daí é um ponto mínimo local
com $f(x^*) = 4.75$



Método de Busca por Gradiente

Inicialização: Seleccione a tolerância ϵ e uma solução experimental inicial x' qualquer. Vá primeiramente para a regra de paragem.

Iteração:

1. Expresse $f(x' + t\nabla f(x'))$ em função de t fazendo que :

$$x_j = x'_j + t \left(\frac{\partial f}{\partial x_j} \right)_{x=x'} \quad \text{para } j = 1, 2, \dots, n.$$

e depois substitua essa expressão em $f(x)$



Método de Busca por Gradiente

2. Use um procedimento de busca para optimização irrestrita de uma variável (ou cálculo) para encontrar $t = t^*$ que maximize $f(x' + t \nabla f(x'))$ ao longo de $t \geq 0$.

3. Reinicialize : $x' = x' + t^* \nabla f(x')$. A seguir vá para a regra de paragem

Regra de Paragem : calcule $\nabla f(x')$ em $x = x'$. Verifique se

$$\left| \frac{\partial f}{\partial x_j} \right| \leq \varepsilon \quad \text{para todo o } j = 1, 2, \dots, n$$

Caso positivo pare no x' e aceite - o como solução, caso contrário realize mais uma iteração.



Método de Busca por Gradiente

Maximizar a função $f(x) = 2x_1x_2 + 2x_2 - x_1^2 - 2x_2^2$

As derivadas em função de x_1 e x_2 são dadas por:

$$\frac{df(x)}{dx_1} = 2x_2 - 2x_1$$

$$\frac{df(x)}{dx_2} = 2x_1 + 2 - 4x_2$$

Para iniciar o método de busca por gradiente, depois de escolher um valor adequadamente pequeno para ε (normalmente bem abaixo de 0,1) suponha que $x = 0$ seja seleccionado como solução experimental inicial.



Método de Busca por Gradiente

Pelo facto das respectivas derivadas parciais serem 0 e 2 nesse ponto o gradiente fica:

$$\nabla f(0,0) = (0,2)$$

Como $\varepsilon \leq 2$, a regra de paragem diz para realizar mais uma iteração.

Iteração 1: Com valores iguais a 0 e 2 para as respectivas derivadas parciais a primeira iteração começa fazendo com que:

$$x_1 = 0 + t(0) = 0,$$

$$x_2 = 0 + t(2) = 2t$$



Método de Busca por Gradiente

E depois substituindo essas expressões em $f(x)$ para obter:

$$\begin{aligned}f(x' + t\nabla f(x')) &= f(0, 2t) \\&= 2(0)(2t) + 2(2t) - 0^2 - 2(2t)^2 \\&= 4t - 8t^2\end{aligned}$$

como:

$$f(0, 2t^*) = \max_{t \geq 0} f(0, 2t) = \max_{t \geq 0} (4t - 8t^2)$$

$$e \frac{d}{dt}(4t - 8t^2) = 4 - 16t = 0 \Rightarrow t^* = \frac{1}{4}$$

$$\text{Reinicialize } x' = (0, 0) + \frac{1}{4}(0, 2) = \left(0, \frac{1}{2}\right)$$



Método de Busca por Gradiente

Para essa nova solução experimental o gradiente é:

$$\nabla f\left(0, \frac{1}{2}\right) = (1, 0)$$

Como $\varepsilon \leq 1$, a regra de paragem diz para realizar mais uma iteração.

Iteração 2: Com valores iguais a 1 e 0 para as respectivas derivadas parciais a primeira iteração começa fazendo que

$$x_1 = 0 + t(1) = t,$$

$$x_2 = \frac{1}{2} + t(0) = \frac{1}{2}$$



Método de Busca por Gradiente

E depois substituindo essas expressões em $f(x)$ para obter:

$$\begin{aligned}f(x' + t\nabla f(x')) &= f\left(0 + t, \frac{1}{2} + 0t\right) \\&= 2(t)\left(\frac{1}{2}\right) + 2\left(\frac{1}{2}\right) - t^2 - 2\left(\frac{1}{2}\right)^2 \\&= t - t^2 + \frac{1}{2}\end{aligned}$$

como:

$$f\left(t^*, \frac{1}{2}\right) = \max_{t \geq 0} f\left(0, \frac{1}{2}\right) = \max_{t \geq 0} \left(t - t^2 + \frac{1}{2}\right)$$

$$e \frac{d}{dt} \left(t - t^2 + \frac{1}{2}\right) = 1 - 2t = 0 \Rightarrow t^* = \frac{1}{2}$$

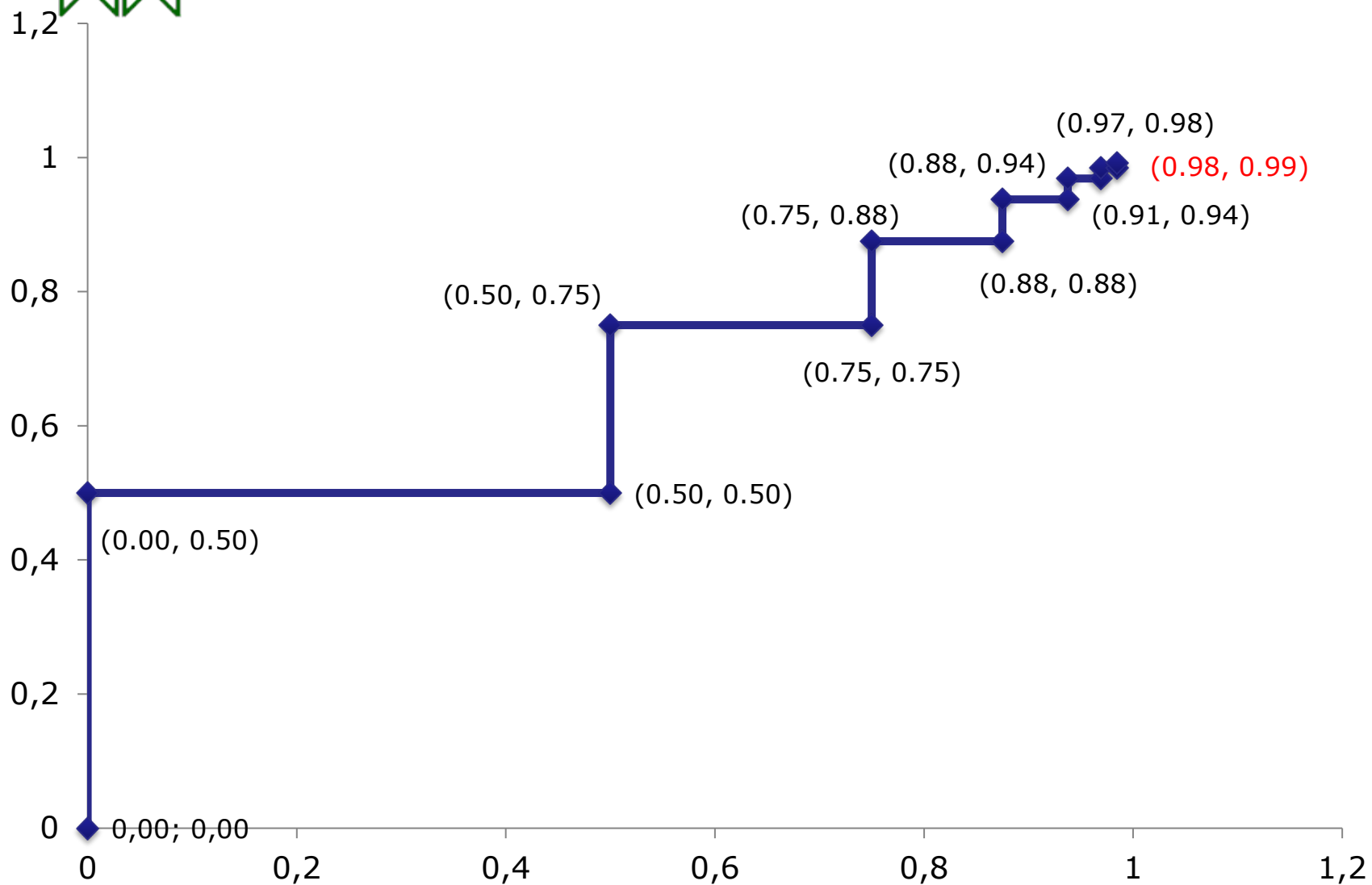
$$\text{Reinicialize } x' = \left(0, \frac{1}{2}\right) + \frac{1}{2}(1, 0) = \left(\frac{1}{2}, \frac{1}{2}\right)$$

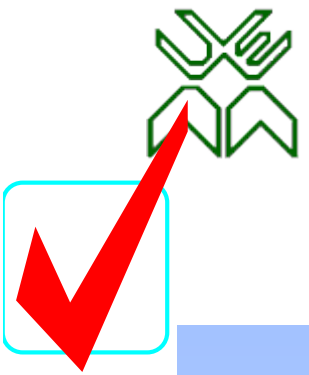


Método de Busca por Gradiente

Uma maneira interessante de se organizar o trabalho é como se apresenta na tabela a seguir:

Iteração	x_1	x_2	df/dx_1	df/dx_2	$f(x_1, x_2)$	t^*	x_1'	x_2'
1	0	0	0	2	0	0,25	0	0,5
2	0	0,5	1	0	0,5	0,5	0,5	0,5
3	0,5	0,5	0	1	0,75	0,25	0,5	0,75
4	0,5	0,75	0,5	0	0,875	0,5	0,75	0,75
5	0,75	0,75	0	0,5	0,9375	0,25	0,75	0,875
6	0,75	0,875	0,25	0	0,96875	0,5	0,875	0,875
7	0,875	0,875	0	0,25	0,984375	0,25	0,875	0,9375
8	0,875	0,9375	0,125	0	0,992188	0,5	0,9375	0,9375
9	0,9375	0,9375	0	0,125	0,996094	0,25	0,9375	0,96875
10	0,9375	0,96875	0,0625	0	0,998047	0,5	0,96875	0,96875
11	0,96875	0,96875	0	0,0625	0,999023	0,25	0,96875	0,984375
12	0,96875	0,984375	0,03125	0	0,999512	0,5	0,984375	0,984375
13	0,984375	0,984375	0	0,03125	0,999756	0,25	0,984375	0,992188
14	0,984375	0,9921875	0,015625	0	0,999878	0,5	0,992188	0,992188





Método de Busca por Gradiente

Se $f(x)$ não fosse uma função côncava o método de busca por gradiente ainda convergiria para um máximo local. A única alteração na descrição do procedimento para esse caso é que t^* agora corresponderia ao primeiro máximo local de $f(x' + t \nabla f(x'))$ à medida que t for incrementado a partir de 0.

Se, ao contrário, o objectivo fosse minimizar $f(x)$, uma alteração no procedimento seria mover-se na direcção oposta do gradiente a cada iteração. Em outras palavras, a regra para obter o ponto seguinte seria:

Método de Busca por Gradiente



Reinicialize $x' = x' - t * \nabla f(x')$. A seguir vá para a regra, a única alteração é que t^* seria o valor não negativo de t que minimiza $f(x' - t \nabla f(x'))$ isto é:

$$f(x' - t * \nabla f(x')) = \min_{t \geq 0} f(x' - t \nabla f(x'))$$