

## Capítulo 3 - Mínimos Quadrados Lineares

Carlos Balsa

balsa@ipb.pt

Departamento de Matemática  
Escola Superior de Tecnologia e Gestão de Bragança

2º Ano - Eng. Civil, Química e Gestão Industrial



# Outline

- 1 Ajuste de Dados pelos Mínimos Quadrados
- 2 Existência, Unicidade e Condicionamento
  - Existência e Unicidade
  - Condicionamento
- 3 Resolução de Problemas de Mínimos Quadrados
  - Método da Equação Normal

## Ajuste de Dados

- Dados  $m$  pontos  $(t_i, y_i)$  pretende-se encontrar um vector  $x$  de dimensão  $n$  cujos parâmetros ajustam melhor a função modelo  $f(t, x)$ ,

$$\min_x \sum_{i=1}^m (y_i - f(t_i, x))^2$$

- Problema **linear** se a função  $f$  for linear nas componentes de  $x$

$$f(t, x) = x_1\phi_1 + x_2\phi_2 + \dots + x_n\phi_n$$

em que  $\phi_j$  depende apenas de  $t$

- Problema pode ser escrito na forma matricial como  $Ax \cong b$ , com  $a_{ij} = \phi_j(t_i)$  e  $b_i = y_i$

## Ajuste de Dados

- Problema pode ser escrito na forma matricial como

$$\begin{bmatrix} \phi_1(t_1) & \phi_2(t_1) & \cdots & \phi_n(t_1) \\ \phi_1(t_2) & \phi_2(t_2) & \cdots & \phi_n(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(t_m) & \phi_2(t_m) & \cdots & \phi_n(t_m) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

- Fazendo  $a_{ij} = \phi_j(t_i)$  e  $b_i = y_i$  podemos escrever

$$Ax = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = b$$

- Como nestes problemas  $m > n$ , o sistema resultante é **sobredeterminado** (mais equações do que incógnitas)

## Ajuste de Dados

- Sistema é mais correctamente representado por  $Ax \cong b$  porque a igualdade não é geralmente satisfeita
- Solução  $x$  é o vector que minimiza o quadrado dos desvios anteriormente apresentado que, na notação matricial, coincide com o quadrado da norma-2 do vector resíduo  $b - Ax$

$$\min_x \|r\|_2^2 = \min_x \|b - Ax\|_2^2$$

## Exemplo: Ajuste de Dados

- Encontrar o polinómio do segundo grau  $P(t) = x_1 + x_2 t + x_3 t^2$  que melhor ajusta os seguintes dados no sentido dos mínimos quadrados

|     |      |      |     |     |     |
|-----|------|------|-----|-----|-----|
| $t$ | -1.0 | -0.5 | 0.0 | 0.5 | 1.0 |
| $y$ | 1.0  | 0.5  | 0.0 | 0.5 | 2.0 |

- Ajustando um polinómio de segundo grau aos cinco pontos origina o problema de mínimos quadrados lineares

$$Ax = \begin{bmatrix} 1 & t_1 & t_1^2 \\ 1 & t_2 & t_2^2 \\ 1 & t_3 & t_3^2 \\ 1 & t_4 & t_4^2 \\ 1 & t_5 & t_5^2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \cong \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} = b$$

- Matrizes cujas colunas (ou linhas) são potencias sucessivas da variável independente é chamada **matriz de Vandermonde**

## Exemplo, continuação

- Introduzindo os respectivos valores

$$Ax = \begin{bmatrix} 1 & -1.0 & 1.0 \\ 1 & -0.5 & 0.25 \\ 1 & 0.0 & 0.0 \\ 1 & 0.5 & 0.25 \\ 1 & 1.0 & 1.0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \cong \begin{bmatrix} 1.0 \\ 0.5 \\ 0.0 \\ 0.5 \\ 2.0 \end{bmatrix} = b$$

- Solução deste sistema, que aprenderemos mais tarde a calcular, é

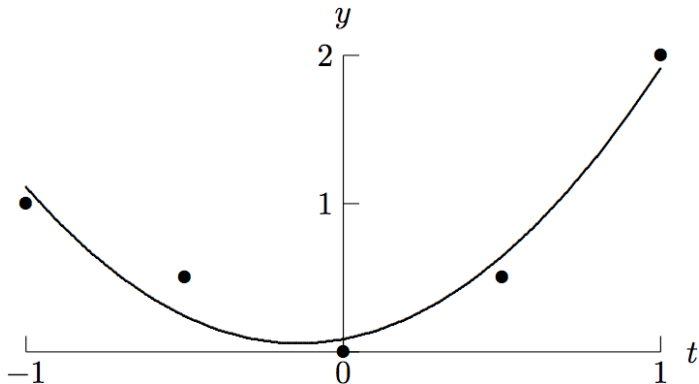
$$x = [0.086 \quad 0.40 \quad 1.4]^T$$

pelo que o polinómio é

$$P(t) = 0.086 + 0.40t + 1.4t^2$$

## Exemplo, continuação

- Curva resultante e pontos dados são mostrados no gráfico



## Existência e Unicidade

- Problema de mínimos quadrados linear  $Ax \cong b$  tem **sempre** solução
- Solução é única se, e só se, as colunas de  $A$  são **linearmente independentes**, i.e., se  $r(A) = n$  (característica de  $A$  é  $n$ ), sendo  $A$  uma matriz  $m \times n$
- Se  $r(A) < n$  o problema de mínimos quadrados linear não é única
- Vamos considerar apenas o caso em que  $A$  tem todas as suas colunas linearmente independentes  $r(A) = n$

## Ortogonalidade

- Vectors  $v_1$  e  $v_2$  são **orthogonais** se o seu produto interno (escalar) é nulo,  $v_1^T v_2 = 0$
- Espaço gerado pelas colunas da matriz  $A$ ,  $\text{span}(A) = \{Ax : x \in \mathbb{R}^n\}$ , tem dimensão não superior a  $n$
- Se  $m > n$ , geralmente  $b \notin \text{span}(A)$ , pelo que não existe solução exacta de  $Ax = b$
- Vector  $y = Ax$  em  $\text{span}(A)$  mais próximo de  $b$  em norma-2 ocorre quando o resíduo  $r = b - Ax$  é orthogonal ao  $\text{span}(A)$ ,

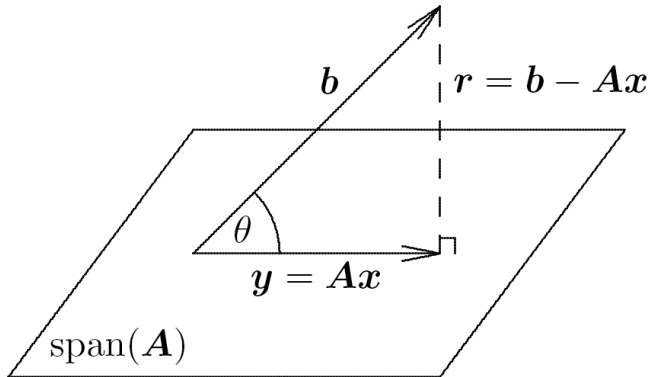
$$A^T r = A^T (b - Ax) = 0,$$

originado a **equação normal**

$$A^T Ax = A^T b$$

## Ortogonalidade, continuação

- Relação geométrica entre  $b$ ,  $r$  e  $\text{span}(\mathbf{A})$ :



## Condicionamento

- Matriz não quadrada  $A$ ,  $m \times n$ , não admite inversa no sentido usual
- Se  $r(A) = n$ , a pseudoinversa de  $A$  é definida por

$$A^+ = (A^T A)^{-1} A^T$$

e o número de condição por

$$\text{cond}(A) = \|A\|_2 \cdot \|A^+\|_2$$

- Por convenção,  $\text{cond}(A) = \infty$  se  $r(A) < n$
- Tal como o número de condição de matrizes quadradas mede a proximidade da singularidade, o número de condição de uma matriz rectangular mede a proximidade de ter uma característica incompleta ( $r(A) < n$ )
- Solução do problema de mínimos quadrados linear  $Ax \cong b$  é dada por  $x = A^+ b$  que é a solução da equação normal  $A^T A x = A^T b$

## Sensibilidade e Condicionamento

- Sensibilidade da solução do problema de mínimos quadrados  $Ax \cong b$  depende de  $b$  assim como de  $A$
- Definindo o ângulo  $\theta$  entre  $b$  e  $y = Ax$  por

$$\cos(\theta) = \frac{\|y\|_2}{\|b\|_2} = \frac{\|Ax\|_2}{\|b\|_2}$$

- Limite da perturbação  $\Delta x$  na solução  $x$  devido à perturbação  $\Delta b$  em  $b$  é dada por

$$\frac{\|\Delta x\|_2}{\|x\|_2} \leq \text{cond}(A) \frac{1}{\cos(\theta)} \frac{\|\Delta b\|_2}{\|b\|_2}$$

- Da mesma forma para uma perturbação  $E$  na matriz  $A$ ,

$$\frac{\|\Delta x\|_2}{\|x\|_2} \leq \left( [\text{cond}(A)]^2 \tan(\theta) + \text{cond}(A) \right) \frac{\|E\|_2}{\|A\|_2}$$

## Método da Equação Normal

- Se  $A$ ,  $m \times n$ , tem característica  $n$ , a matriz  $A^T A$ , de dimensão  $n \times n$ , é simétrica e positiva definida, pelo que a sua factorização de Cholesky

$$A^T A = LL^T$$

pode ser usada para obter a solução  $x$  da equação normal

$$A^T Ax = A^T b$$

que tem a mesma solução que o problema de mínimos quadrados linear  $Ax \cong b$

- Método da equação normal envolve a transformação  
rectangular  $\longrightarrow$  quadrada  $\longrightarrow$  triangular

## Exemplo: Método da Equação Normal

- Utilize o método da equação normal para o ajuste polinomial do exemplo anterior
- Vimos que o problema de mínimos quadrados linear consiste em determinar o vector  $x$  que minimiza o resíduo do sistema sobredeterminado

$$Ax = \begin{bmatrix} 1 & -1.0 & 1.0 \\ 1 & -0.5 & 0.25 \\ 1 & 0.0 & 0.0 \\ 1 & 0.5 & 0.25 \\ 1 & 1.0 & 1.0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \cong \begin{bmatrix} 1.0 \\ 0.5 \\ 0.0 \\ 0.5 \\ 2.0 \end{bmatrix} = b$$

- Solução  $x$  obtida através da resolução da equação normal  
 $A^T Ax = A^T b$

## Exemplo, continuação

- Matriz dos coeficientes e termo independente

$$\begin{aligned}
 A^T A &= \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -1.0 & -0.5 & 0.0 & 0.5 & 1.0 \\ 1.0 & 0.25 & 0.0 & 0.25 & 1.0 \end{bmatrix} \begin{bmatrix} 1 & -1.0 & 1.0 \\ 1 & -0.5 & 0.25 \\ 1 & 0.0 & 0.0 \\ 1 & 0.5 & 0.25 \\ 1 & 1.0 & 1.0 \end{bmatrix} \\
 &= \begin{bmatrix} 5.0 & 0.0 & 2.5 \\ 0.0 & 2.5 & 0.0 \\ 2.5 & 0.0 & 2.125 \end{bmatrix}, \\
 A^T b &= \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -1.0 & -0.5 & 0.0 & 0.5 & 1.0 \\ 1.0 & 0.25 & 0.0 & 0.25 & 1.0 \end{bmatrix} \begin{bmatrix} 1.0 \\ 0.5 \\ 0.0 \\ 0.5 \\ 2.0 \end{bmatrix} = \begin{bmatrix} 4.0 \\ 1.0 \\ 3.25 \end{bmatrix}
 \end{aligned}$$

## Exemplo, continuação

- Factorização de Cholesky da matriz simétrica e positiva definida  $A^T A$  resulta em

$$\begin{aligned} A^T A &= \begin{bmatrix} 5.0 & 0.0 & 2.5 \\ 0.0 & 2.5 & 0.0 \\ 2.5 & 0.0 & 2.125 \end{bmatrix} \\ &= \begin{bmatrix} 2.236 & 0 & 0 \\ 0 & 1.581 & 0 \\ 1.118 & 0 & 0.935 \end{bmatrix} \begin{bmatrix} 2.236 & 0 & 1.118 \\ 0 & 1.581 & 0 \\ 0 & 0 & 0.935 \end{bmatrix} = L^T L \end{aligned}$$

- Resolvendo o sistema triangular inferior  $Lz = A^T b$  por substituição regressiva obtemos  $z = [1.789 \ 0.632 \ 1.336]^T$
- Resolvendo o sistema triangular superior  $L^T x = z$  por substituição progressiva obtemos  $x = [0.086 \ 0.400 \ 1.429]^T$