Mínimos Quadrados

Cálculo Numérico

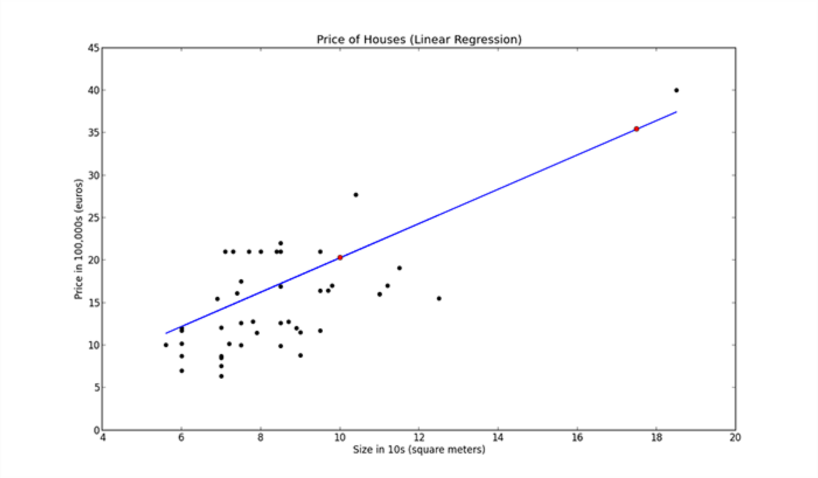
Bóris Marin

UFABC

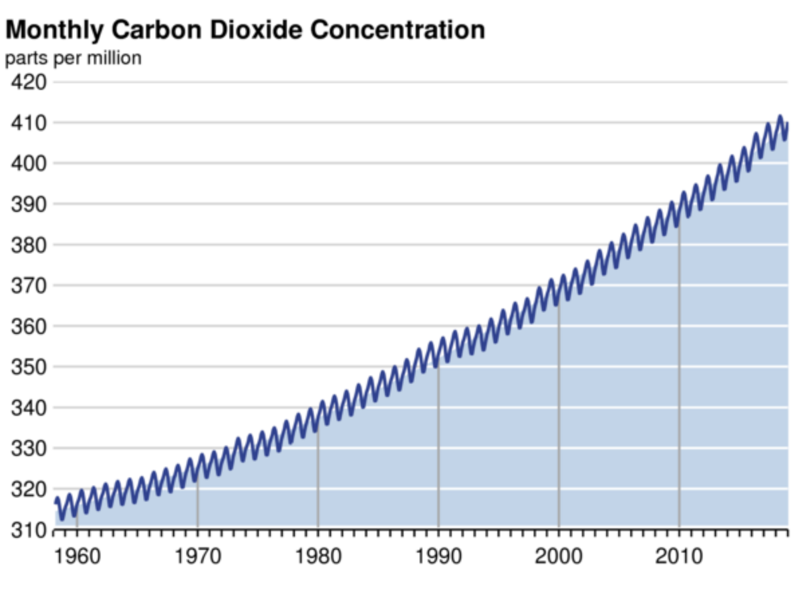
Motivação

Regressão: Desemprego / Crescimento do PIB

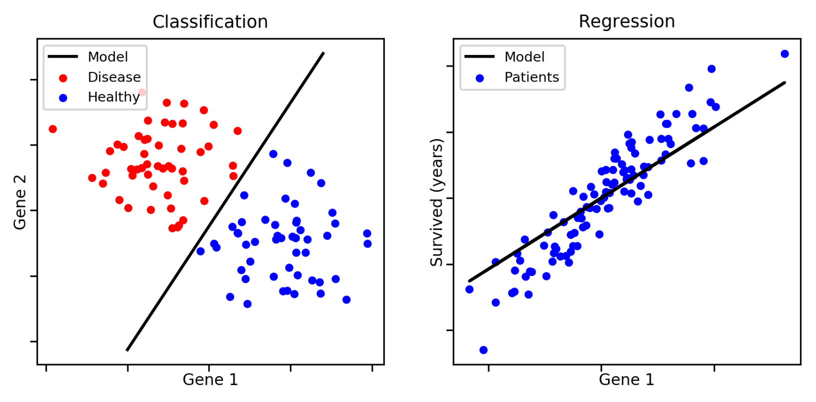
Regressão: Mercado Imobiliário



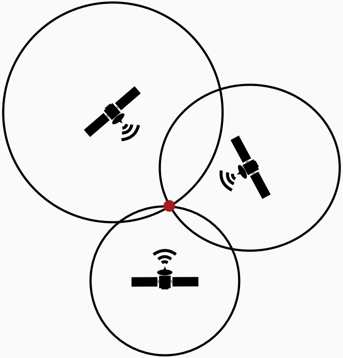
Modelos Matemáticos: Gás Carbônico

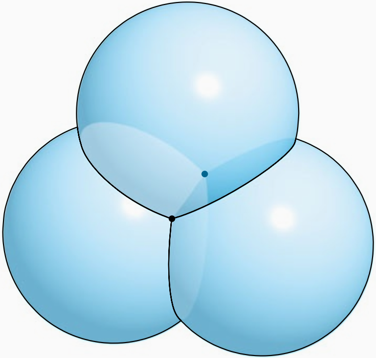


Aprendizado de Máquina — Predição / Classificação



Sistema Sobredeterminado: GPS





Sistemas inconsistentes

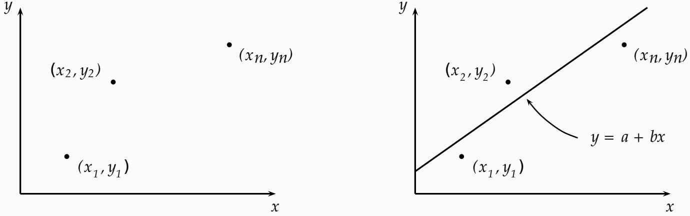
- Até agora, lidamos sistemas lineares com soluções (e únicas)

Sistemas inconsistentes

- Até agora, lidamos sistemas lineares com soluções (e únicas)
- O que acontece se o sistema *não tiver* soluções?

Sistemas inconsistentes

- Até agora, lidamos sistemas lineares com soluções (e únicas)
- O que acontece se o sistema *não tiver* soluções?
- Exemplo: Sistema inconsistente (comum quando o número de equações é maior do que o de incógnitas). Temos o que fazer?



$$\begin{cases} a + x_1 b = y_1 \\ a + x_2 b = y_2 \\ \vdots \\ a + x_n b = y_n \end{cases} \iff \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \\ 1 & x_n \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \iff \mathbf{A}$$

Sistemas inconsistentes

- Exemplo: como satisfazer à 1a. e 3a. equações abaixo?

$$\begin{cases} x_1 + x_2 = 2 \\ x_1 - x_2 = 1 \\ x_1 + x_2 = 3 \end{cases}$$

Sistemas inconsistentes

- Exemplo: como satisfazer à 1a. e 3a. equações abaixo?

$$\begin{cases} x_1 + x_2 = 2 \\ x_1 - x_2 = 1 \\ x_1 + x_2 = 3 \end{cases}$$

- Não há solução: sistema **inconsistente**.

Sistemas inconsistentes

- Exemplo: como satisfazer à 1a. e 3a. equações abaixo?

$$\begin{cases} x_1 + x_2 = 2 \\ x_1 - x_2 = 1 \\ x_1 + x_2 = 3 \end{cases}$$

- Não há solução: sistema **inconsistente**.
    - Coeficientes imprecisos

Sistemas inconsistentes

- Exemplo: como satisfazer à 1a. e 3a. equações abaixo?

$$\begin{cases} x_1 + x_2 = 2 \\ x_1 - x_2 = 1 \\ x_1 + x_2 = 3 \end{cases}$$

- Não há solução: sistema **inconsistente**.
  - Coeficientes imprecisos
  - Mais equações do que variáveis

Sistemas inconsistentes

- Exemplo: como satisfazer à 1a. e 3a. equações abaixo?

$$\begin{cases} x_1 + x_2 = 2 \\ x_1 - x_2 = 1 \\ x_1 + x_2 = 3 \end{cases}$$

- Não há solução: sistema **inconsistente**.
  - Coeficientes imprecisos
  - Mais equações do que variáveis
- Eliminação Gaussiana não nos dará uma solução. Mas há algum $\bar{x}$ que esteja 'perto' de ser solução?

Sistemas inconsistentes

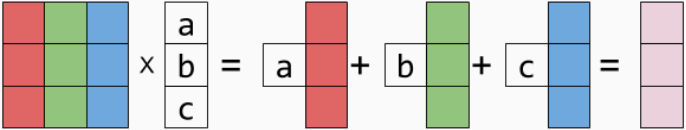
- Exemplo: como satisfazer à 1a. e 3a. equações abaixo?

$$\begin{cases} x_1 + x_2 = 2 \\ x_1 - x_2 = 1 \\ x_1 + x_2 = 3 \end{cases}$$

- Não há solução: sistema **inconsistente**.
  - Coeficientes imprecisos
  - Mais equações do que variáveis
- Eliminação Gaussiana não nos dará uma solução. Mas há algum $\bar{x}$ que esteja 'perto' de ser solução?
- Se 'perto' significa distância Euclideana, o **Método dos Mínimos Quadrados** permite encontrar $\bar{x}$

Parênteses — Multiplicando Matriz por vetor

$$\begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} ax_1 + by_1 + cz_1 \\ ax_2 + by_2 + cz_2 \\ ax_3 + by_3 + cz_3 \end{bmatrix}$$



Qualquer sistema $Ax + b$ (m equações, n incógnitas) pode ser escrito desta forma (combinações lineares das m colunas v_i de A, com coeficientes $x_1, \ldots, x_n$):

$$x_1 v_1 + x_2 v_2 + \ldots + x_n v_n = b$$

Mínimos Quadrados

- A forma matricial $Ax = b$ do sistema anterior é

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

Mínimos Quadrados

- A forma matricial $Ax = b$ do sistema anterior é

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

- Representando b como combinação linear das colunas de A:

$$x_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

Mínimos Quadrados

- A forma matricial $Ax = b$ do sistema anterior é

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

- Representando b como combinação linear das colunas de A:

$$x_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

- Ou seja, queremos obter **b** como uma combinação linear de dois vetores tridimensionais

Mínimos Quadrados

- A forma matricial $Ax = b$ do sistema anterior é

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

- Representando b como combinação linear das colunas de A:

$$x_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

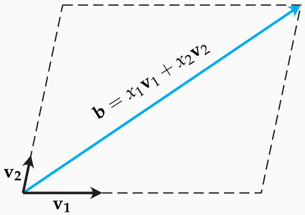
- Ou seja, queremos obter **b** como uma combinação linear de dois vetores tridimensionais
- Como todas as combinações possíveis destes vetores formam um plano em $\mathbb{R}^3$, o sistema só terá solução se **b** estiver neste plano (o que não é o caso).

Mínimos Quadrados

- Nos casos em que **b** não está no *espaço coluna* de A, o sistema é **inconsistente**. Mas podemos achar uma solução “próxima”:

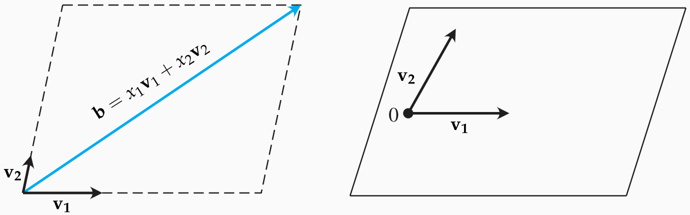
Mínimos Quadrados

- Nos casos em que $\mathbf{b}$ não está no *espaço coluna* de A, o sistema é **inconsistente**. Mas podemos achar uma solução "próxima":



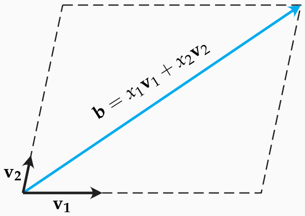
Mínimos Quadrados

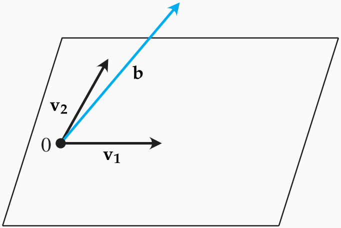
- Nos casos em que $\mathbf{b}$ não está no *espaço coluna* de A, o sistema é **inconsistente**. Mas podemos achar uma solução "próxima":



Mínimos Quadrados

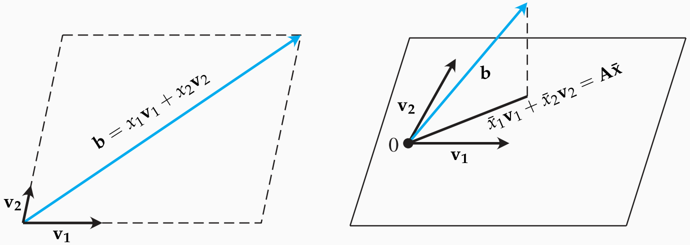
- Nos casos em que **b** não está no *espaço coluna* de A, o sistema é **inconsistente**. Mas podemos achar uma solução "próxima":





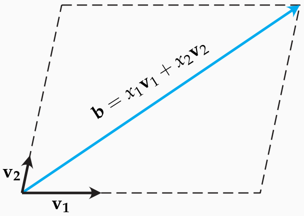
Mínimos Quadrados

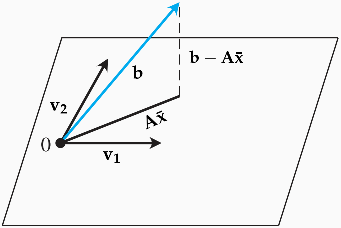
- Nos casos em que **b** não está no *espaço coluna* de A, o sistema é **inconsistente**. Mas podemos achar uma solução "próxima":



Mínimos Quadrados

- Nos casos em que **b** não está no *espaço coluna* de A, o sistema é **inconsistente**. Mas podemos achar uma solução "próxima":





Mínimos Quadrados

- Não há um par (x_1, x_2) que resolva o sistema, mas há um ponto no plano $\mathbf{Ax}$ que é o mais próximo de $\mathbf{b}$

Mínimos Quadrados

- Não há um par (x_1, x_2) que resolva o sistema, mas há um ponto no plano $\mathbf{Ax}$ que é o mais próximo de $\mathbf{b}$
- Este ponto é dado por $\mathbf{A\bar{x}}$ na figura anterior.

Mínimos Quadrados

- Não há um par (x_1, x_2) que resolva o sistema, mas há um ponto no plano $\mathbf{Ax}$ que é o mais próximo de $\mathbf{b}$
- Este ponto é dado por $\mathbf{A\bar{x}}$ na figura anterior.
- O vetor $\bar{\mathbf{x}}$ é chamado de **solução de mínimos quadrados**.

Mínimos Quadrados

- Não há um par (x_1, x_2) que resolva o sistema, mas há um ponto no plano $\mathbf{Ax}$ que é o mais próximo de $\mathbf{b}$
- Este ponto é dado por $\mathbf{A\bar{x}}$ na figura anterior.
- O vetor $\bar{\mathbf{x}}$ é chamado de **solução de mínimos quadrados**.
- O que o **resíduo** (vetor $\mathbf{b} - \mathbf{A\bar{x}}$) tem de especial neste caso?

Parênteses — Álgebra Matricial

- Seja A uma uma matriz $m \times n$

Parênteses — Álgebra Matricial

- Seja A uma uma matriz $m \times n$
- A **matriz transposta** A^T tem como *colunas* as *linhas* de A, e como *linhas* as *colunas* de A

Parênteses — Álgebra Matricial

- Seja A uma uma matriz $m \times n$
- A **matriz transposta** A^T tem como *colunas* as *linhas* de A, e como *linhas* as *colunas* de A
- A transposta de uma soma é a soma das transpostas:

$$(A + B)^T = A^T + B^T$$

Parênteses — Álgebra Matricial

- Seja A uma uma matriz $m \times n$
- A **matriz transposta** A^T tem como *colunas* as *linhas* de A, e como *linhas* as *colunas* de A
- A transposta de uma soma é a soma das transpostas:

$$(A + B)^T = A^T + B^T$$

- A transposta de um produto é o produto das transpostas *em ordem invertida*:

$$(AB)^T = B^T A^T$$

Parênteses — Álgebra Matricial

- O *produto escalar* de dois vetores coluna de m dimensões u, v é dado por uma multiplicação de matrizes:

$$u \cdot v \equiv u^T v = \begin{bmatrix} u_1, \ldots, u_m \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix}$$

Parênteses — Álgebra Matricial

- O *produto escalar* de dois vetores coluna de m dimensões u, v é dado por uma multiplicação de matrizes:

$$u \cdot v \equiv u^T v = \begin{bmatrix} u_1, \ldots, u_m \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix}$$

- A norma ℓ^2 (ou Euclideana) de um vetor é dada por

$$\|u\|_2 = \sqrt{u_1^2 + u_2^2 + \ldots + u_n^2} = \sqrt{u \cdot u} = \sqrt{u^T u}$$

Parênteses — Álgebra Matricial

- O *produto escalar* de dois vetores coluna de m dimensões u, v é dado por uma multiplicação de matrizes:

$$u \cdot v \equiv u^T v = \begin{bmatrix} u_1, \ldots, u_m \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix}$$

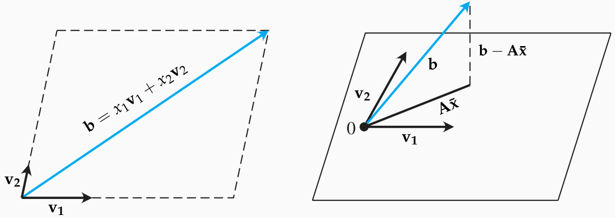
- A norma ℓ^2 (ou Euclideana) de um vetor é dada por

$$\|u\|_2 = \sqrt{u_1^2 + u_2^2 + \ldots + u_n^2} = \sqrt{u \cdot u} = \sqrt{u^T u}$$

- Dois vetores u, v são *perpendiculares* ou *ortogonais* se o seu produto escalar for zero.

Mínimos Quadrados

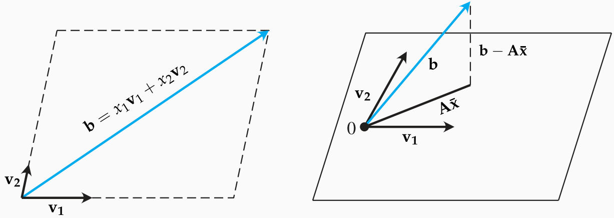
- Determinamos geometricamente que a solução de MQ envolve um resíduo $b - A\bar{x}$ perpendicular ao espaço-coluna de A



$$(b - A\bar{x}) \perp \{Ax | x \in \mathbb{R}^n\}$$

Mínimos Quadrados

- Determinamos geometricamente que a solução de MQ envolve um resíduo $b - A\bar{x}$ perpendicular ao espaço-coluna de A



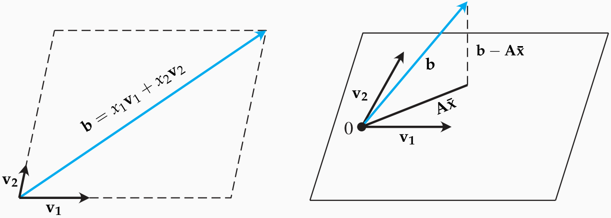
$$(b - A\bar{x}) \perp \{Ax | x \in \mathbb{R}^n\}$$

- Vetores ortogonais $\implies$ produto escalar nulo:

$$(Ax)^T(b - A\bar{x}) = 0 \qquad \text{para todos } x \text{ em } \mathbb{R}^n$$

Mínimos Quadrados

- Determinamos geometricamente que a solução de MQ envolve um resíduo $b - A\bar{x}$ perpendicular ao espaço-coluna de A



$$(b - A\bar{x}) \perp \{Ax | x \in \mathbb{R}^n\}$$

- Vetores ortogonais $\implies$ produto escalar nulo:

$$(Ax)^T(b - A\bar{x}) = 0 \qquad \text{para todos } x \text{ em } \mathbb{R}^n$$

- Abrindo a transposta do produto no produto das transpostas:

$$x^T A^T (b - A\bar{x}) = 0 \qquad \text{para todos } x \text{ em } \mathbb{R}^n$$

Mínimos Quadrados

- Chegamos em

$$x^T A^T (b - A\bar{x}) = 0 \qquad \text{para todos } x \text{ em } \mathbb{R}^n$$

Mínimos Quadrados

- Chegamos em

$$x^T A^T (b - A\bar{x}) = 0 \qquad \text{para todos } x \text{ em } \mathbb{R}^n$$

- A parte colorida é um produto escalar: $x \cdot A^T(b - A\bar{x}) = 0$

Mínimos Quadrados

- Chegamos em

$$x^T A^T (b - A\bar{x}) = 0 \qquad \text{para todos } x \text{ em } \mathbb{R}^n$$

- A parte colorida é um produto escalar: $x \cdot A^T(b - A\bar{x}) = 0$
- Isso quer dizer que o vetor n-dimensional $A^T(b - A\bar{x})$ é perpendicular a todos os vetores x em $\mathbb{R}^n$, *incluindo ele mesmo*.

Mínimos Quadrados

- Chegamos em

$$x^T A^T (b - A\bar{x}) = 0 \qquad \text{para todos } x \text{ em } \mathbb{R}^n$$

- A parte colorida é um produto escalar: $x \cdot A^T(b - A\bar{x}) = 0$
- Isso quer dizer que o vetor n-dimensional $A^T(b - A\bar{x})$ é perpendicular a todos os vetores x em $\mathbb{R}^n$, *incluindo ele mesmo*.
- Isto só ocorre se $A^T(b - Ax)$ for o vetor nulo:

$$A^T(b - A\bar{x}) = 0$$

Mínimos Quadrados

- Chegamos em

$$x^T A^T(b - A\bar{x}) = 0 \qquad \text{para todos } x \text{ em } \mathbb{R}^n$$

- A parte colorida é um produto escalar: $x \cdot A^T(b - A\bar{x}) = 0$
- Isso quer dizer que o vetor n-dimensional $A^T(b - A\bar{x})$ é perpendicular a todos os vetores x em $\mathbb{R}^n$, *incluindo ele mesmo*.
- Isto só ocorre se $A^T(b - Ax)$ for o vetor nulo:

$$A^T(b - A\bar{x}) = 0$$

- Temos então um sistema de equações para $\bar{x}$

$$\boxed{A^T A\bar{x} = A^T b}$$

Equações Normais

- Resumindo: dado um sistema indeterminado $Ax = b$, obtemos as **equações normais**

$$\boxed{A^T A\bar{x} = A^T b}$$

Equações Normais

- Resumindo: dado um sistema indeterminado $Ax = b$, obtemos as **equações normais**

$$\boxed{A^T A\bar{x} = A^T b}$$

- A solução $\bar{x}$ deste sistema é o vetor que minimiza a norma Euclideana do resíduo $r = b - Ax$

Equações Normais

- Resumindo: dado um sistema indeterminado $Ax = b$, obtemos as **equações normais**

$$\boxed{A^T A\bar{x} = A^T b}$$

- A solução $\bar{x}$ deste sistema é o vetor que minimiza a norma Euclideana do resíduo $r = b - Ax$
- O vetor $\bar{x}$ é denominado **solução de mínimos quadrados** para $Ax = b$.

MMQ — Exemplo

- Vamos agora usar o MMQ para resolver o sistema indeterminado
$$\begin{cases} x_1 + x_2 = 2 \\ x_1 - x_2 = 1 \\ x_1 + x_2 = 3 \end{cases}$$

MMQ — Exemplo

- Vamos agora usar o MMQ para resolver o sistema indeterminado
$$\begin{cases} x_1 + x_2 = 2 \\ x_1 - x_2 = 1 \\ x_1 + x_2 = 3 \end{cases}$$
- Escrevendo este sistema em forma matricial, obtemos
$$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}, \ b = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

MMQ — Exemplo

- As componentes das equações normais são

$$A^T A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$$

MMQ — Exemplo

- As componentes das equações normais são

$$A^T A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$$

- Assim como

$$A^T b = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

MMQ — Exemplo

- As equações normais ficam

$$\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

MMQ — Exemplo

- As equações normais ficam

$$\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

- Aplicando eliminação Gaussiana:

$$\left[\begin{array}{cc|c} 3 & 1 & 6 \\ 1 & 3 & 4 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 3 & 1 & 6 \\ 0 & 8/3 & 2 \end{array}\right]$$

MMQ — Exemplo

- As equações normais ficam

$$\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

- Aplicando eliminação Gaussiana:

$$\left[\begin{array}{cc|c} 3 & 1 & 6 \\ 1 & 3 & 4 \end{array}\right] \to \left[\begin{array}{cc|c} 3 & 1 & 6 \\ 0 & 8/3 & 2 \end{array}\right]$$

- O que nos dá a solução

$$\bar{x} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} = \begin{bmatrix} \frac{7}{4} \\ \frac{3}{4} \end{bmatrix}$$

MMQ — Exemplo

- Substituindo $\bar{x}$ no problema original, temos

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{7}{4} \\ \frac{3}{4} \end{bmatrix} = \begin{bmatrix} 2.5 \\ 1 \\ 2.5 \end{bmatrix} \neq \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

MMQ — Exemplo

- Substituindo $\bar{x}$ no problema original, temos

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{7}{4} \\ \frac{3}{4} \end{bmatrix} = \begin{bmatrix} 2.5 \\ 1 \\ 2.5 \end{bmatrix} \neq \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

- Para avaliar a qualidade da nossa solução, calculamos o *resíduo* da solução de mínimos quadrados $\bar{x}$

$$r \equiv b - A\bar{x} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} 2.5 \\ 1 \\ 2.5 \end{bmatrix} = \begin{bmatrix} -0.5 \\ 0 \\ 0.5 \end{bmatrix}$$

MMQ — Exemplo

- Substituindo $\bar{x}$ no problema original, temos

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{7}{4} \\ \frac{3}{4} \end{bmatrix} = \begin{bmatrix} 2.5 \\ 1 \\ 2.5 \end{bmatrix} \neq \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

- Para avaliar a qualidade da nossa solução, calculamos o *resíduo* da solução de mínimos quadrados $\bar{x}$

$$r \equiv b - A\bar{x} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} 2.5 \\ 1 \\ 2.5 \end{bmatrix} = \begin{bmatrix} -0.5 \\ 0 \\ 0.5 \end{bmatrix}$$

- Se o resíduo for o vetor nulo, temos uma solução exata do sistema $Ax = b$

MMQ — Exemplo

- Substituindo $\bar{x}$ no problema original, temos

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{7}{4} \\ \frac{3}{4} \end{bmatrix} = \begin{bmatrix} 2.5 \\ 1 \\ 2.5 \end{bmatrix} \neq \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

- Para avaliar a qualidade da nossa solução, calculamos o *resíduo* da solução de mínimos quadrados $\bar{x}$

$$r \equiv b - A\bar{x} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} 2.5 \\ 1 \\ 2.5 \end{bmatrix} = \begin{bmatrix} -0.5 \\ 0 \\ 0.5 \end{bmatrix}$$

- Se o resíduo for o vetor nulo, temos uma solução exata do sistema $Ax = b$
- Caso contrário, a norma do resíduo representa um **erro reverso**, uma medida do quão longe $\bar{x}$ está de ser uma solução.

MMQ — Erro

- Há três maneira usuais de se exprimir o “tamanho” do resíduo.

MMQ — Erro

- Há três maneira usuais de se exprimir o "tamanho" do resíduo.
  - O comprimento Euclideano (ou norma ℓ^2)

$$\|r\|_2 = \sqrt{r_1^2 + \ldots + r_m^2}$$

MMQ — Erro

- Há três maneira usuais de se exprimir o “tamanho” do resíduo.
  - O comprimento Euclideano (ou norma ℓ^2)

$$\|r\|_2 = \sqrt{r_1^2 + \ldots + r_m^2}$$

  - O Erro Quadrático

$$EQ = r_1^2 + \ldots + r_m^2$$

MMQ — Erro

- Há três maneira usuais de se exprimir o “tamanho” do resíduo.
  - O comprimento Euclideano (ou norma ℓ^2)

$$\|r\|_2 = \sqrt{r_1^2 + \ldots + r_m^2}$$

  - O Erro Quadrático

$$EQ = r_1^2 + \ldots + r_m^2$$

  - A Raiz do Erro Quadrático Médio (Root Mean Squared Error)

$$REQM = \sqrt{EQ/m} = \sqrt{(r_1^2 + \ldots + r_m^2)/m}$$

MMQ — Erro

- Há três maneira usuais de se exprimir o “tamanho” do resíduo.
  - O comprimento Euclideano (ou norma ℓ^2)

$$\|r\|_2 = \sqrt{r_1^2 + \ldots + r_m^2}$$

  - O Erro Quadrático

$$EQ = r_1^2 + \ldots + r_m^2$$

  - A Raiz do Erro Quadrático Médio (Root Mean Squared Error)

$$REQM = \sqrt{EQ/m} = \sqrt{(r_1^2 + \ldots + r_m^2)/m}$$

- Todas estas expressões estão relacionadas,

$$REQM = \frac{\sqrt{EQ}}{\sqrt{m}} = \frac{\|r\|_2}{\sqrt{m}}$$

MMQ — Erro

- Há três maneira usuais de se exprimir o “tamanho” do resíduo.
  - O comprimento Euclideano (ou norma ℓ^2)

$$\|r\|_2 = \sqrt{r_1^2 + \ldots + r_m^2}$$

  - O Erro Quadrático

$$EQ = r_1^2 + \ldots + r_m^2$$

  - A Raiz do Erro Quadrático Médio (Root Mean Squared Error)

$$REQM = \sqrt{EQ/m} = \sqrt{(r_1^2 + \ldots + r_m^2)/m}$$

- Todas estas expressões estão relacionadas,

$$REQM = \frac{\sqrt{EQ}}{\sqrt{m}} = \frac{\|r\|_2}{\sqrt{m}}$$

- de modo que o $\bar{x}$ que minimizar um deles, minimiza todos.

Ajustando modelos a dados

Ajuste de modelos

- Seja $(t_1, y_1), \ldots, (t_m, y_n)$ um conjunto de pontos no plano, os "dados".

Ajuste de modelos

- Seja $(t_1, y_1), \ldots, (t_m, y_n)$ um conjunto de pontos no plano, os "dados".
- Consideremos uma classe fixa de modelos matemáticos, por exemplo todas as retas $y = c_1 + c_2 t$

Ajuste de modelos

- Seja $(t_1, y_1), \ldots, (t_m, y_n)$ um conjunto de pontos no plano, os "dados".
- Consideremos uma classe fixa de modelos matemáticos, por exemplo todas as retas $y = c_1 + c_2 t$
- Qual destas retas é a que "melhor ajusta" os dados?

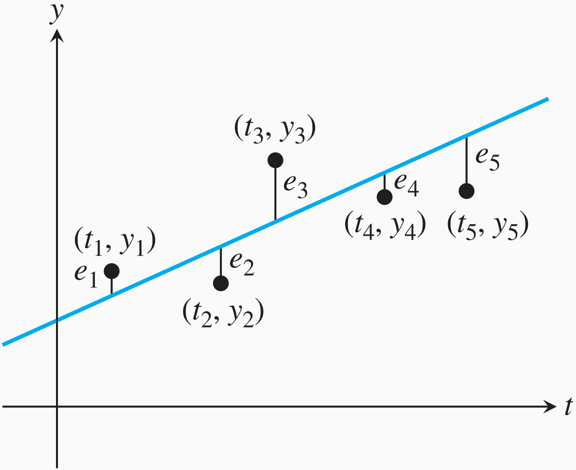
Ajuste de modelos

- Seja $(t_1, y_1), \ldots, (t_m, y_n)$ um conjunto de pontos no plano, os "dados".
- Consideremos uma classe fixa de modelos matemáticos, por exemplo todas as retas $y = c_1 + c_2 t$
- Qual destas retas é a que "melhor ajusta" os dados?
- Melhor em que sentido?

Ajuste de modelos

- Seja $(t_1, y_1), \ldots, (t_m, y_n)$ um conjunto de pontos no plano, os "dados".
- Consideremos uma classe fixa de modelos matemáticos, por exemplo todas as retas $y = c_1 + c_2 t$
- Qual destas retas é a que "melhor ajusta" os dados?
- Melhor em que sentido?
- MMQ: medir o resíduo usando o erro quadrado, e encontrar c_1, c_2 (*parâmetros* do modelo) que minimizam este erro.

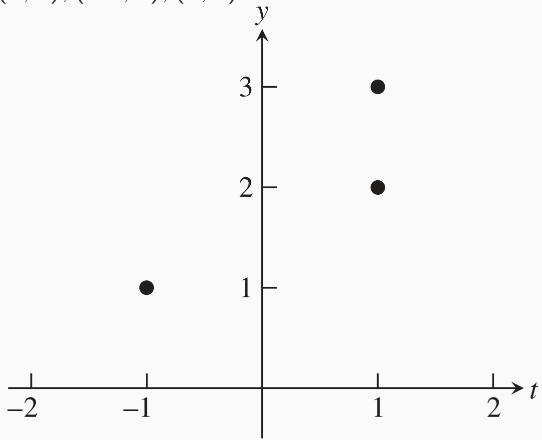
Ajuste de modelos



A "melhor" dentre todas as retas $y = c_1 + c_2 t$ é aquela para o qual o erro quadrático $e_1^2 + e_2^2 + \ldots + e_5^2$ é menor.

Ajuste de modelos — Exemplo

- Vamos encontrar a reta que melhor ajusta os dados $(t, y) = (1, 2), (-1, 1), (1, 3)$:



Ajuste de modelos — Exemplo

- Modelo linear geral: $y = c_1 + c_2 t$. Substituindo os pontos:

$$\begin{cases} c_1 + c_2(1) = 2 \\ c_1 + c_2(-1) = 1 \\ c_1 + c_2(1) = 3 \end{cases}$$

Ajuste de modelos — Exemplo

- Modelo linear geral: $y = c_1 + c_2 t$. Substituindo os pontos:

$$\begin{cases} c_1 + c_2(1) = 2 \\ c_1 + c_2(-1) = 1 \\ c_1 + c_2(1) = 3 \end{cases}$$

- Em forma matricial,

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

Ajuste de modelos — Exemplo

- Modelo linear geral: $y = c_1 + c_2 t$. Substituindo os pontos:

$$\begin{cases} c_1 + c_2(1) = 2 \\ c_1 + c_2(-1) = 1 \\ c_1 + c_2(1) = 3 \end{cases}$$

- Em forma matricial,

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

- Este sistema tem alguma solução (c_1, c_2)?

Ajuste de modelos — Exemplo

- Modelo linear geral: $y = c_1 + c_2 t$. Substituindo os pontos:

$$\begin{cases} c_1 + c_2(1) = 2 \\ c_1 + c_2(-1) = 1 \\ c_1 + c_2(1) = 3 \end{cases}$$

- Em forma matricial,

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

- Este sistema tem alguma solução (c_1, c_2)?
- Não. Se houvesse uma solução, a reta $y = c_1 + c_2 t$ passaria pelos três pontos — mas eles claramente não são colineares.

Ajuste de modelos — Exemplo

- Modelo linear geral: $y = c_1 + c_2 t$. Substituindo os pontos:

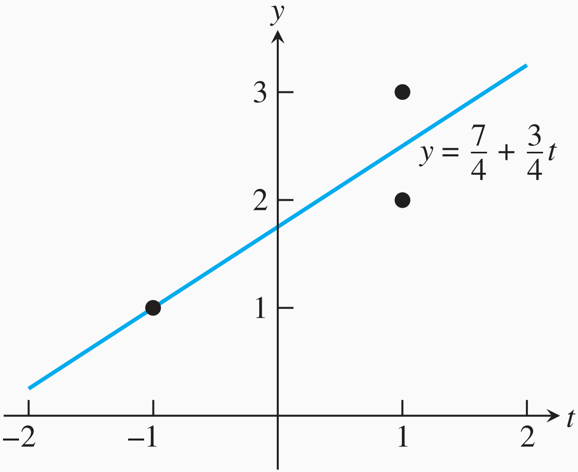
$$\begin{cases} c_1 + c_2(1) = 2 \\ c_1 + c_2(-1) = 1 \\ c_1 + c_2(1) = 3 \end{cases}$$

- Em forma matricial,

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

- Este sistema tem alguma solução (c_1, c_2)?
- Não. Se houvesse uma solução, a reta $y = c_1 + c_2 t$ passaria pelos três pontos — mas eles claramente não são colineares.
- Já vimos este sistema inconsistente anteriormente, e achamos uma solução por mínimos quadrados $(c_1, c_2) = (7/4, 3/4)$

Ajuste de modelos — Exemplo



Ajuste de modelos — Exemplo

- Como calcular o erro?

Ajuste de modelos — Exemplo

- Como calcular o erro?
- Basta ver a diferença entre o valor previsto pelo modelo e cada ponto dos dados

t	y	reta	erro
1	2	2.5	-0.5
-1	1	1.0	0
1	3	2.5	0.5
1	3	2.5	0.5

Ajuste de modelos — Exemplo

- Como calcular o erro?
- Basta ver a diferença entre o valor previsto pelo modelo e cada ponto dos dados

t	y	reta	erro
1	2	2.5	-0.5
-1	1	1.0	0
1	3	2.5	0.5
1	3	2.5	0.5

- O erro quadrático médio é $\frac{\sqrt{(-0.5)^2+0^2+(0.5)^2}}{\sqrt{3}} = \frac{1}{\sqrt{6}} \approx 0.408$

Ajuste de modelos por Mínimos Quadrados

Para um conjunto de m pontos ("dados") $(t_1, y_1), \ldots, (t_m, y_n)$

1. Escolha um modelo parametrizado, como $y = c_1 + c_2 t$, para ajustar aos dados.
2. Substitua os dados no modelo. Cada ponto corresponde a uma equação com os parâmetros como incógnitas (c_1, c_2 no modelo linear acima). Isto resulta num sistema $Ax = b$, com x representando os parâmetros desconhecidos.
3. A solução de mínimos quadrados é encontrada resolvendo as equações normais, ou seja o sistema $A^T Ax = A^T b$

Mínimos Quadrados — Compressão

- O MMQ é outro exemplo clássico de *compressão* de dados.

Mínimos Quadrados — Compressão

- O MMQ é outro exemplo clássico de *compressão* de dados.
- Entramos com dados, e a saída é um modelo que, com (relativamente) poucos parâmetros, ajusta os dados da melhor maneira possível.

Mínimos Quadrados — Compressão

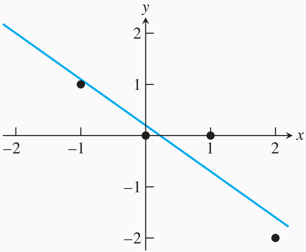
- O MMQ é outro exemplo clássico de *compressão* de dados.
- Entramos com dados, e a saída é um modelo que, com (relativamente) poucos parâmetros, ajusta os dados da melhor maneira possível.
- Geralmente, MMQ é utilizado para substituir dados ruidosos por uma representação matemática subjacente.

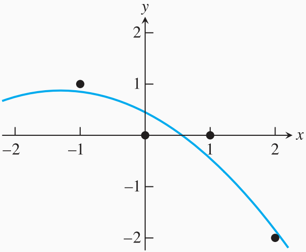
Mínimos Quadrados — Compressão

- O MMQ é outro exemplo clássico de *compressão* de dados.
- Entramos com dados, e a saída é um modelo que, com (relativamente) poucos parâmetros, ajusta os dados da melhor maneira possível.
- Geralmente, MMQ é utilizado para substituir dados ruidosos por uma representação matemática subjacente.
- O modelo obtido é então usado para fins de classificação ou predição

Ajuste de modelos pelo MMQ — Exemplo

- Encontre a melhor reta e a melhor parábola que ajustem $(-1,1),(0,0),(1,0),(2,-2)$.





Ajuste de modelo pelo MMQ — Caso linear

0) Dados $(-1, 1), (0, 0), (1, 0), (2, -2)$.

Ajuste de modelo pelo MMQ — Caso linear

0) Dados $(-1,1),(0,0),(1,0),(2,-2)$.
1) Modelo linear: $y = c_1 + c_2 x$ (dois parâmetros)

Ajuste de modelo pelo MMQ — Caso linear

0) Dados $(-1,1),(0,0),(1,0),(2,-2)$.
1) Modelo linear: $y = c_1 + c_2 x$ (dois parâmetros)
2) Substituindo dados no modelo:

$$\begin{cases} c_1 + c_2(-1) = 1 \\ c_1 + c_2(0) = 0 \\ c_1 + c_2(1) = 0 \\ c_1 + c_2(2) = -2 \end{cases} \qquad \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -2 \end{bmatrix}$$

Ajuste de modelo pelo MMQ — Caso linear

0) Dados $(-1,1),(0,0),(1,0),(2,-2)$.
1) Modelo linear: $y = c_1 + c_2 x$ (dois parâmetros)
2) Substituindo dados no modelo:

$$\begin{cases} c_1 + c_2(-1) = 1 \\ c_1 + c_2(0) = 0 \\ c_1 + c_2(1) = 0 \\ c_1 + c_2(2) = -2 \end{cases} \qquad \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -2 \end{bmatrix}$$

3) Equações normais

$$\begin{bmatrix} 4 & 2 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} -1 \\ -5 \end{bmatrix}$$

Ajuste de modelo pelo MMQ — Caso linear

- A solução das equações normais nos dá $(c_1, c_2) = (0.2, -0.9)$, ou seja o modelo é $y = 0.2 - 0.9x$

Ajuste de modelo pelo MMQ — Caso linear

- A solução das equações normais nos dá $(c_1, c_2) = (0.2, -0.9)$, ou seja o modelo é $y = 0.2 - 0.9x$
- Resíduos:

t	y	reta	erro
-1	1	1.1	-0.1
0	0	0.2	-0.2
1	0	-0.7	0.7
2	2	-1.6	-0.4

Ajuste de modelo pelo MMQ — Caso linear

- A solução das equações normais nos dá $(c_1, c_2) = (0.2, -0.9)$, ou seja o modelo é $y = 0.2 - 0.9x$
- Resíduos:

t	y	reta	erro
-1	1	1.1	-0.1
0	0	0.2	-0.2
1	0	-0.7	0.7
2	2	-1.6	-0.4

- Análise do erro: $EQ = (-.1)^2 + (-.2)^2 + (.7)^2 + (-.4)^2 = 0.7$ e $REQM = \frac{\sqrt{.7}}{\sqrt{4}} \approx 0.418$

Ajuste de modelo pelo MMQ — Caso quadrático

0) Dados $(-1, 1), (0, 0), (1, 0), (2, -2)$.

Ajuste de modelo pelo MMQ — Caso quadrático

0) Dados $(-1, 1), (0, 0), (1, 0), (2, -2)$.
1) Modelo quadrático: $y = c_1 + c_2 x + c_3 x^2$ (três parâmetros)

Ajuste de modelo pelo MMQ — Caso quadrático

0) Dados $(-1,1),(0,0),(1,0),(2,-2)$.
1) Modelo quadrático: $y = c_1 + c_2x + c_3x^2$ (três parâmetros)
2) Substituindo dados no modelo:

$$\begin{cases} c_1 + c_2(-1) + c_3(-1)^2 = 1 \\ c_1 + c_2(0) + c_3(0)^2 = 0 \\ c_1 + c_2(1) + c_3(1)^2 = 0 \\ c_1 + c_2(2) + c_3(2)^2 = -2 \end{cases} \qquad \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -2 \end{bmatrix}$$

Ajuste de modelo pelo MMQ — Caso quadrático

0) Dados $(-1, 1), (0, 0), (1, 0), (2, -2)$.
1) Modelo quadrático: $y = c_1 + c_2 x + c_3 x^2$ (três parâmetros)
2) Substituindo dados no modelo:

$$\begin{cases} c_1 + c_2(-1) + c_3(-1)^2 = 1 \\ c_1 + c_2(0) + c_3(0)^2 = 0 \\ c_1 + c_2(1) + c_3(1)^2 = 0 \\ c_1 + c_2(2) + c_3(2)^2 = -2 \end{cases} \qquad \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -2 \end{bmatrix}$$

3) Equações normais

$$\begin{bmatrix} 4 & 2 & 6 \\ 2 & 6 & 8 \\ 6 & 8 & 18 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -5 \\ -7 \end{bmatrix}$$

Ajuste de modelo pelo MMQ — Caso quadrático

- A solução das equações normais nos dá $(c_1, c_2, c_3) = (0.45, -0.65, -0.25)$, ou seja o modelo é $y = 0.45 - 0.65x - 0.25x^2$

Ajuste de modelo pelo MMQ — Caso quadrático

- A solução das equações normais nos dá $(c_1, c_2, c_3) = (0.45, -0.65, -0.25)$, ou seja o modelo é $y = 0.45 - 0.65x - 0.25x^2$
- Resíduos:

t	y	parábola	erro
-1	1	0.85	0.15
0	0	0.45	-0.45
1	0	-0.45	0.45
2	2	-1.85	-0.15

Ajuste de modelo pelo MMQ — Caso quadrático

- A solução das equações normais nos dá $(c_1, c_2, c_3) = (0.45, -0.65, -0.25)$, ou seja o modelo é $y = 0.45 - 0.65x - 0.25x^2$
- Resíduos:

t	y	parábola	erro
-1	1	0.85	0.15
0	0	0.45	-0.45
1	0	-0.45	0.45
2	2	-1.85	-0.15

- O Erro Quadrático é $EQ = (.15)^2 + (-.45)^2 + (.45)^2 + (-.15)^2 \approx 0.45$ e $REQM = \frac{\sqrt{.45}}{\sqrt{4}} \approx 0.335$

Ajuste de modelo pelo MMQ

- **Importante**: o MMQ pode ser utilizado (conforme acabamos de ver!) para ajustar tanto funções não-lineares como lineares.

Ajuste de modelo pelo MMQ

- **Importante**: o MMQ pode ser utilizado (conforme acabamos de ver!) para ajustar tanto funções não-lineares como lineares.
- O método funciona para qualquer tipo de modelo, desde que os *parâmetros* entrem de forma linear nas equações.

Ajuste de modelo pelo MMQ

- **Importante**: o MMQ pode ser utilizado (conforme acabamos de ver!) para ajustar tanto funções não-lineares como lineares.
- O método funciona para qualquer tipo de modelo, desde que os *parâmetros* entrem de forma linear nas equações.
- Veremos vários exemplos na próxima aula!

Mínimos Quadrados — Considerações Numéricas

MMQ — Condicionamento

- Dados estão sujeitos a erros (FeMec, incerteza experimental, etc)

MMQ — Condicionamento

- Dados estão sujeitos a erros (FeMec, incerteza experimental, etc)
- Ao ajustar modelos via MMQ, é importante então mitigar a amplificação de erros.

MMQ — Condicionamento

- Dados estão sujeitos a erros (FeMec, incerteza experimental, etc)
- Ao ajustar modelos via MMQ, é importante então mitigar a amplificação de erros.
- Que conceito representava a amplificação de erros na entrada?

MMQ — Condicionamento

- Dados estão sujeitos a erros (FeMec, incerteza experimental, etc)
- Ao ajustar modelos via MMQ, é importante então mitigar a amplificação de erros.
- Que conceito representava a amplificação de erros na entrada?
- Isto motiva o estudo do condicionamento das equações normais.

MMQ — Condicionamento

- Dados estão sujeitos a erros (FeMec, incerteza experimental, etc)
- Ao ajustar modelos via MMQ, é importante então mitigar a amplificação de erros.
- Que conceito representava a amplificação de erros na entrada?
- Isto motiva o estudo do condicionamento das equações normais.
- Vamos explorar este problema determinando o erro na saída (forward) da solução via equações normais (resolver o sistema $A^T Ax = A^T b$)

MMQ — Condicionamento

Exemplo: 11 pontos sobre polinômio de grau 7

- Dados $x_1 = 2.0, x_2 = 2.2, x_3 = 2.4, \ldots, x_{11} = 4.0$ equispaçados em $[2, 4]$

MMQ — Condicionamento

Exemplo: 11 pontos sobre polinômio de grau 7

- Dados $x_1 = 2.0, x_2 = 2.2, x_3 = 2.4, \ldots, x_{11} = 4.0$ equispaçados em $[2, 4]$
- Sejam $y_i = 1 + x_i + x_i^2 + x_i^3 + x_i^4 + x_i^5 + x_i^6 + x_i^7$, com $1 \leq i \leq 11$

MMQ — Condicionamento

Exemplo: 11 pontos sobre polinômio de grau 7

- Dados $x_1 = 2.0, x_2 = 2.2, x_3 = 2.4, \ldots, x_{11} = 4.0$ equispaçados em $[2, 4]$
- Sejam $y_i = 1 + x_i + x_i^2 + x_i^3 + x_i^4 + x_i^5 + x_i^6 + x_i^7$, com $1 \leq i \leq 11$
- Vamos usar as equações normais para achar o polinômio de mínimos quadrados $P(x) = c_1 + c_2 x + \ldots + c_8 x^7$ que ajusta os (x_i, y_i)

MMQ — Condicionamento

Exemplo: 11 pontos sobre polinômio de grau 7

- Dados $x_1 = 2.0, x_2 = 2.2, x_3 = 2.4, \ldots, x_{11} = 4.0$ equispaçados em $[2, 4]$
- Sejam $y_i = 1 + x_i + x_i^2 + x_i^3 + x_i^4 + x_i^5 + x_i^6 + x_i^7$, com $1 \leq i \leq 11$
- Vamos usar as equações normais para achar o polinômio de mínimos quadrados $P(x) = c_1 + c_2 x + \ldots + c_8 x^7$ que ajusta os (x_i, y_i)
- Qual é o resultado esperado?

MMQ — Condicionamento

Exemplo: 11 pontos sobre polinômio de grau 7

- Dados $x_1 = 2.0, x_2 = 2.2, x_3 = 2.4, \ldots, x_{11} = 4.0$ equispaçados em $[2, 4]$
- Sejam $y_i = 1 + x_i + x_i^2 + x_i^3 + x_i^4 + x_i^5 + x_i^6 + x_i^7$, com $1 \leq i \leq 11$
- Vamos usar as equações normais para achar o polinômio de mínimos quadrados $P(x) = c_1 + c_2 x + \ldots + c_8 x^7$ que ajusta os (x_i, y_i)
- Qual é o resultado esperado?
- Estamos ajustando um polinômio de grau 7 a 11 pontos, mas *todos* estão sobre o gráfico do polinômio (também de grau 7) $\bar{P}(x) = 1 + x + x^2 + x^3 + x^4 + x^5 + x^6 + x^7$

Exemplo: 11 pontos sobre polinômio de grau 7

- Dados $x_1 = 2.0, x_2 = 2.2, x_3 = 2.4, \ldots, x_{11} = 4.0$ equispaçados em $[2, 4]$
- Sejam $y_i = 1 + x_i + x_i^2 + x_i^3 + x_i^4 + x_i^5 + x_i^6 + x_i^7$, com $1 \leq i \leq 11$
- Vamos usar as equações normais para achar o polinômio de mínimos quadrados $P(x) = c_1 + c_2 x + \ldots + c_8 x^7$ que ajusta os (x_i, y_i)
- Qual é o resultado esperado?
- Estamos ajustando um polinômio de grau 7 a 11 pontos, mas *todos* estão sobre o gráfico do polinômio (também de grau 7) $\bar{P}(x) = 1 + x + x^2 + x^3 + x^4 + x^5 + x^6 + x^7$
- Ou seja, a solução é $c_1 = c_2 = \ldots = c_8 = 1$

MMQ — Condicionamento

Exemplo: 11 pontos sobre polinômio de grau 7

- Substituindo os dados no modelo $P(x)$, chegamos a um sistema $Ac = b$

$$\begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^7 \\ 1 & x_2 & x_2^2 & \dots & x_2^7 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{11} & x_{11}^2 & \dots & x_{11}^7 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_8 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{11} \end{bmatrix}$$

MMQ — Condicionamento

Exemplo: 11 pontos sobre polinômio de grau 7

- Substituindo os dados no modelo $P(x)$, chegamos a um sistema $Ac = b$

$$\begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^7 \\ 1 & x_2 & x_2^2 & \dots & x_2^7 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{11} & x_{11}^2 & \dots & x_{11}^7 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_8 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{11} \end{bmatrix}$$

- Ou seja, temos uma matriz de Vandermonde.

MMQ — Condicionamento

- Resolvendo o problema numericamente:

```
x = linspace(2,4,11)
y = 1 + x + x**2 + x**3 + x**4 + x**5 + x**6 + x**7
A = c_[x**0, x, x**2, x**3, x**4, x**5, x**6, x**7]
solve(A.T @ A, A.T @ y)
>>> [0.15796926,  3.08362063, -1.18744072,  2.2630896,  0.5666848,
        1.08833448,  0.99008967,  1.00047215]
cond(A.T @ A)
>>> 7.845335402995069e+18
```

MMQ — Condicionamento

- Resolvendo o problema numericamente:

```
x = linspace(2,4,11)
y = 1 + x + x**2 + x**3 + x**4 + x**5 + x**6 + x**7
A = c_[x**0, x, x**2, x**3, x**4, x**5, x**6, x**7]
solve(A.T @ A, A.T @ y)
>>> [0.15796926,  3.08362063, -1.18744072,  2.2630896,  0.5666848,
        1.08833448,  0.99008967,  1.00047215]
cond(A.T @ A)
>>> 7.845335402995069e+18
```

- Erramos feio! A soluçÃo exata é $c_1 = c_2 = \ldots = c_8 = 1$. Problema mal condicionado!

MMQ — Condicionamento

- Vimos que as equações normais levam a Matrizes de Vandermonde

MMQ — Condicionamento

- Vimos que as equações normais levam a Matrizes de Vandermonde
- Estas matrizes são notoriamente mal-condicionadas

MMQ — Condicionamento

- Vimos que as equações normais levam a Matrizes de Vandermonde
- Estas matrizes são notoriamente mal-condicionadas
- Em particular, o número de condicionamento de $A^T A$ é aproximadamente o *quadrado* daquele de A.

MMQ — Condicionamento

- Vimos que as equações normais levam a Matrizes de Vandermonde
- Estas matrizes são notoriamente mal-condicionadas
- Em particular, o número de condicionamento de $A^T A$ é aproximadamente o *quadrado* daquele de A.
- Ou seja, é provável que o problema seja mal-condicionado.

MMQ — Condicionamento

- Vimos que as equações normais levam a Matrizes de Vandermonde
- Estas matrizes são notoriamente mal-condicionadas
- Em particular, o número de condicionamento de $A^T A$ é aproximadamente o *quadrado* daquele de A.
- Ou seja, é provável que o problema seja mal-condicionado.
- MMQ via equações normais: só para problemas "pequenos"

MMQ — Condicionamento

- Vimos que as equações normais levam a Matrizes de Vandermonde
- Estas matrizes são notoriamente mal-condicionadas
- Em particular, o número de condicionamento de $A^T A$ é aproximadamente o *quadrado* daquele de A.
- Ou seja, é provável que o problema seja mal-condicionado.
- MMQ via equações normais: só para problemas "pequenos"
- Soluções: fatorações que eliminam a necessidade de calcular $A^T A$ (como fatoração QR), ou decomposição em valores singulares (SVD).

Fatoração QR

- Idéia geral: decompor uma matriz quadrada real em um produto $A = QR$

Fatoração QR

- Idéia geral: decompor uma matriz quadrada real em um produto $A = QR$
- Lembrando: em $A = LU$, capturamos a informação da eliminação Gaussiana

Fatoração QR

- Idéia geral: decompor uma matriz quadrada real em um produto $A = QR$
- Lembrando: em $A = LU$, capturamos a informação da eliminação Gaussiana
- Aqui, capturamos a *ortogonalização* da matriz A (construção de conjunto ortogonal de vetores que *gera* o espaço coluna de A)

Fatoração QR

- Idéia geral: decompor uma matriz quadrada real em um produto $A = QR$
- Lembrando: em $A = LU$, capturamos a informação da eliminação Gaussiana
- Aqui, capturamos a *ortogonalização* da matriz A (construção de conjunto ortogonal de vetores que *gera* o espaço coluna de A)
- Q é uma matriz *ortogonal* e R uma matriz triangular superior.

Fatoração QR

- Idéia geral: decompor uma matriz quadrada real em um produto $A = QR$
- Lembrando: em $A = LU$, capturamos a informação da eliminação Gaussiana
- Aqui, capturamos a *ortogonalização* da matriz A (construção de conjunto ortogonal de vetores que *gera* o espaço coluna de A)
- Q é uma matriz *ortogonal* e R uma matriz triangular superior.
- Matrizes ortogonais: $Q^T Q = QQ^T = 1$, ou $Q^T = Q^{-1} \rightarrow$ fáceis de inverter.

Fatoração QR

- Idéia geral: decompor uma matriz quadrada real em um produto $A = QR$
- Lembrando: em $A = LU$, capturamos a informação da eliminação Gaussiana
- Aqui, capturamos a *ortogonalização* da matriz A (construção de conjunto ortogonal de vetores que *gera* o espaço coluna de A)
- Q é uma matriz *ortogonal* e R uma matriz triangular superior.
- Matrizes ortogonais: $Q^T Q = QQ^T = 1$, ou $Q^T = Q^{-1} \rightarrow$ fáceis de inverter.
- Propriedade fundamental de matrizes ortogonais: preservam a norma euclideana de vetores $\rightarrow$ não amplificam erros.

Fatoração QR

- Por que preservam a norma Euclideana?

$$\|Qx\|_2^2 = (Qx)^T Qx = x^T Q^T Qx = x^T x = \|x\|_2^2$$

Fatoração QR

- Por que preservam a norma Euclideana?

$$\|Qx\|_2^2 = (Qx)^T Qx = x^T Q^T Qx = x^T x = \|x\|_2^2$$

- Infelizmente, a decomposição QR (há vários algoritmos, incluindo uso direto de Gram-Schmidt) é custosa ($\mathcal{O}(n^3)$).

Fatoração QR

- Por que preservam a norma Euclideana?

$$\|Qx\|_2^2 = (Qx)^T Qx = x^T Q^T Qx = x^T x = \|x\|_2^2$$

- Infelizmente, a decomposição QR (há vários algoritmos, incluindo uso direto de Gram-Schmidt) é custosa ($\mathcal{O}(n^3)$).
- Ainda assim, ela permite resolver mínimos quadrados sem construir $A^T A$.

Mínimos Quadrados via fatoração QR

- Dado um sistema inconsistente $m \times n$ $Ax = b$

Mínimos Quadrados via fatoração QR

- Dado um sistema inconsistente $m \times n$ $Ax = b$
- Encontramos a decomposição $A = QR$

Mínimos Quadrados via fatoração QR

- Dado um sistema inconsistente $m \times n$ $Ax = b$
- Encontramos a decomposição $A = QR$
- Fazemos $\hat{R} =$ submatriz superior $n \times n$ de R

Mínimos Quadrados via fatoração QR

- Dado um sistema inconsistente $m \times n$ $Ax = b$
- Encontramos a decomposição $A = QR$
- Fazemos $\hat{R} =$ submatriz superior $n \times n$ de R
- Fazemos $\hat{d} = n$ entradas superiores de $d = Q^T b$

Fatoração QR

Mínimos Quadrados via fatoração QR

- Dado um sistema inconsistente $m \times n$ $Ax = b$
- Encontramos a decomposição $A = QR$
- Fazemos $\hat{R} =$ submatriz superior $n \times n$ de R
- Fazemos $\hat{d} = n$ entradas superiores de $d = Q^T b$
- Resolvemos $\hat{R}\bar{x} = \hat{d}$ para obter a solução de mínimos quadrados $\bar{x}$.

Fatoração QR — Exemplo

- Para o exemplo mal-condicionado anterior:

```
x = linspace(2,4,11)
y = 1 + x + x**2 + x**3 + x**4 + x**5 + x**6 + x**7
A = c_[x**0, x, x**2, x**3, x**4, x**5, x**6, x**7]

Q,R = qr(A)
d = Q.T @ y
solve(R[0:8, 0:8], d[0:8])
>>> [0.99999985, 1.00000036, 0.99999962, 1.00000021, 0.99999993,
       1.00000001, 1.        , 1.        ]
```

Fatoração QR — Exemplo

- Para o exemplo mal-condicionado anterior:

```
x = linspace(2,4,11)
y = 1 + x + x**2 + x**3 + x**4 + x**5 + x**6 + x**7
A = c_[x**0, x, x**2, x**3, x**4, x**5, x**6, x**7]

Q,R = qr(A)
d = Q.T @ y
solve(R[0:8, 0:8], d[0:8])
>>> [0.99999985, 1.00000036, 0.99999962, 1.00000021, 0.99999993,
       1.00000001, 1.        , 1.        ]
```

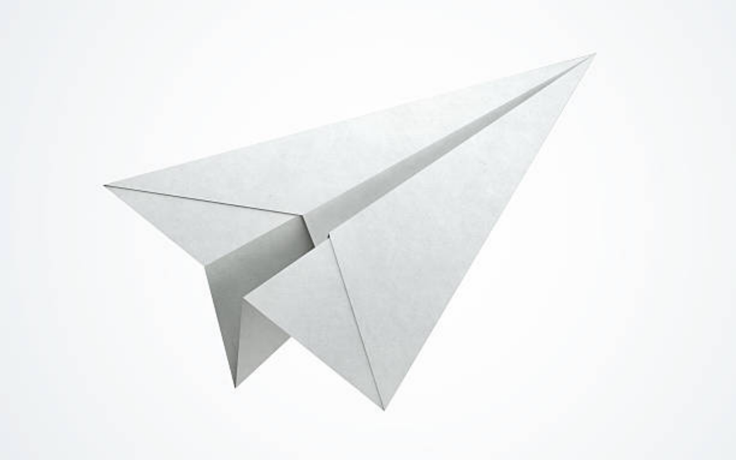
- Encontramos agora ao menos 6 casas decimais corretas da solução exata $c_1 = c_2 = \ldots = c_8 = 1$.

Ajuste de Modelos — Estudo de Casos

Modelos / analogias

Modelos / analogias

Modelos / analogias



Modelagem

- Assunto complexo! Questões abordadas pela epistemologia.

Modelagem

- Assunto complexo! Questões abordadas pela epistemologia.
- Modelo: aproximação da realidade, pode ser útil ou não.

Modelagem

- Assunto complexo! Questões abordadas pela epistemologia.
- Modelo: aproximação da realidade, pode ser útil ou não.
- Modelos quantitativos aqui abordados:

Modelagem

- Assunto complexo! Questões abordadas pela epistemologia.
- Modelo: aproximação da realidade, pode ser útil ou não.
- Modelos quantitativos aqui abordados:
  - derivados de princípios físicos subjacentes à fonte de dados

Modelagem

- Assunto complexo! Questões abordadas pela epistemologia.
- Modelo: aproximação da realidade, pode ser útil ou não.
- Modelos quantitativos aqui abordados:
    - derivados de princípios físicos subjacentes à fonte de dados
    - puramente baseados em fatores empíricos

Dados periódicos

- Há fenômenos que são intrinsecamente periódicos

Dados periódicos

- Há fenômenos que são intrinsecamente periódicos
- A temperatura do ar, por exemplo, varia periodicamente em várias escalas temporais

Dados periódicos

- Há fenômenos que são intrinsecamente periódicos
- A temperatura do ar, por exemplo, varia periodicamente em várias escalas temporais
  - rotação da terra

Dados periódicos

- Há fenômenos que são intrinsecamente periódicos
- A temperatura do ar, por exemplo, varia periodicamente em várias escalas temporais
  - rotação da terra
  - translação da terra

Dados periódicos

- Há fenômenos que são intrinsecamente periódicos
- A temperatura do ar, por exemplo, varia periodicamente em várias escalas temporais
  - rotação da terra
  - translação da terra
- Qual tipo de função podemos usar para ajustar dados periódicos? Ela é linear?

Dados periódicos: Temperatura

- A temperatura do ar em Londres (Reino Unido), no dia 1 de Janeiro de 2001, foi:

hora	$t(dia)$	temp (°C)
0 h	0	-2.2
3 h	$\frac{1}{8}$	-2.8
6 h	$\frac{1}{4}$	-6.1
9 h	$\frac{3}{8}$	-3.9
12 h	$\frac{1}{2}$	1.1
15 h	$\frac{5}{8}$	0.0
18 h	$\frac{3}{4}$	-0.6
21 h	$\frac{7}{8}$	-1.1

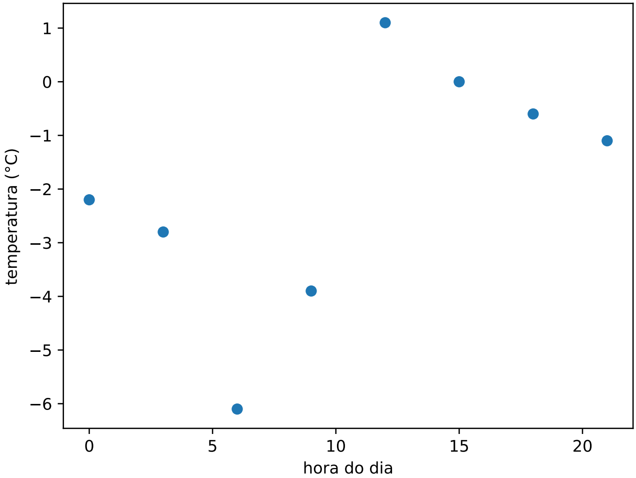
Dados periódicos: Temperatura

- A temperatura do ar em Londres (Reino Unido), no dia 1 de Janeiro de 2001, foi:

hora	$t(dia)$	temp (°C)
0 h	0	-2.2
3 h	$\frac{1}{8}$	-2.8
6 h	$\frac{1}{4}$	-6.1
9 h	$\frac{3}{8}$	-3.9
12 h	$\frac{1}{2}$	1.1
15 h	$\frac{5}{8}$	0.0
18 h	$\frac{3}{4}$	-0.6
21 h	$\frac{7}{8}$	-1.1

- Estes dados têm um "caráter" periódico?

Dados periódicos: Temperatura



Dados periódicos: Temperatura

- Como ajustar um modelo periódico aos dados?

Dados periódicos: Temperatura

- Como ajustar um modelo periódico aos dados?
- Vamos tentar $y = c_1 + c_2 \cos 2\pi t + c_3 \sin 2\pi t$

Dados periódicos: Temperatura

- Como ajustar um modelo periódico aos dados?
- Vamos tentar $y = c_1 + c_2 \cos 2\pi t + c_3 \sin 2\pi t$
- Qual é o período dos termos em seno, cosseno?

Dados periódicos: Temperatura

- Como ajustar um modelo periódico aos dados?
- Vamos tentar $y = c_1 + c_2 \cos 2\pi t + c_3 \sin 2\pi t$
- Qual é o período dos termos em seno, cosseno?
- Sabemos que a temperatura é aproximadamente periódica, com um período de 24 h

Dados periódicos: Temperatura

- Como ajustar um modelo periódico aos dados?
- Vamos tentar $y = c_1 + c_2 \cos 2\pi t + c_3 \sin 2\pi t$
- Qual é o período dos termos em seno, cosseno?
- Sabemos que a temperatura é aproximadamente periódica, com um período de 24 h
- Se medirmos t em dias, obtemos a coluna t na tabela anterior

Dados periódicos: Temperatura

- Ajuste pelo MMQ: substituimos os dados no modelo

$$
\begin{aligned}
c_1 + c_2 \cos 2\pi(0) + c_3 \sin 2\pi(0) &= -2.2 \\
c_1 + c_2 \cos 2\pi \left(\frac{1}{8}\right) + c_3 \sin 2\pi \left(\frac{1}{8}\right) &= -2.8 \\
c_1 + c_2 \cos 2\pi \left(\frac{1}{4}\right) + c_3 \sin 2\pi \left(\frac{1}{4}\right) &= -6.1 \\
c_1 + c_2 \cos 2\pi \left(\frac{3}{8}\right) + c_3 \sin 2\pi \left(\frac{3}{8}\right) &= -3.9 \\
c_1 + c_2 \cos 2\pi \left(\frac{1}{2}\right) + c_3 \sin 2\pi \left(\frac{1}{2}\right) &= 0.0 \\
c_1 + c_2 \cos 2\pi \left(\frac{5}{8}\right) + c_3 \sin 2\pi \left(\frac{5}{8}\right) &= 1.1 \\
c_1 + c_2 \cos 2\pi \left(\frac{3}{4}\right) + c_3 \sin 2\pi \left(\frac{3}{4}\right) &= -0.6 \\
c_1 + c_2 \cos 2\pi \left(\frac{7}{8}\right) + c_3 \sin 2\pi \left(\frac{7}{8}\right) &= -1.1
\end{aligned}
$$

Dados periódicos: Temperatura

- Chegamos ao sistema sobredeterminado

$$A = \begin{bmatrix} 1 & \cos 0 & \sin 0 \\ 1 & \cos\frac{\pi}{4} & \sin\frac{\pi}{4} \\ 1 & \cos\frac{\pi}{2} & \sin\frac{\pi}{2} \\ 1 & \cos\frac{3\pi}{4} & \sin\frac{3\pi}{4} \\ 1 & \cos\pi & \sin\pi \\ 1 & \cos\frac{5\pi}{4} & \sin\frac{5\pi}{4} \\ 1 & \cos\frac{3\pi}{2} & \sin\frac{3\pi}{2} \\ 1 & \cos\frac{7\pi}{4} & \sin\frac{7\pi}{4} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & \sqrt{2}/2 & \sqrt{2}/2 \\ 1 & 0 & 1 \\ 1 & -\sqrt{2}/2 & \sqrt{2}/2 \\ 1 & -1 & 0 \\ 1 & -\sqrt{2}/2 & -\sqrt{2}/2 \\ 1 & 0 & -1 \\ 1 & \sqrt{2}/2 & -\sqrt{2}/2 \end{bmatrix} \qquad b = \begin{bmatrix} -2.2 \\ -2.8 \\ -6.1 \\ -3.9 \\ 0.0 \\ 1.1 \\ -0.6 \\ -1.1 \end{bmatrix}.$$

Dados periódicos: Temperatura

- Determinamos e resolvemos as equações normais $A^T Ac = A^T b$:

$$\begin{bmatrix} 8 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} -15.6 \\ -2.9778 \\ -10.2376 \end{bmatrix}$$

Dados periódicos: Temperatura

- Determinamos e resolvemos as equações normais $A^T A c = A^T b$:

$$\begin{bmatrix} 8 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} -15.6 \\ -2.9778 \\ -10.2376 \end{bmatrix}$$

- este sistema tem solução (imediata) $c_1 = -1.95, c_2 = -0.7445, c_3 = -2.5594$

Dados periódicos: Temperatura

- Determinamos e resolvemos as equações normais $A^T A c = A^T b$:

$$\begin{bmatrix} 8 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} -15.6 \\ -2.9778 \\ -10.2376 \end{bmatrix}$$

- este sistema tem solução (imediata) $c_1 = -1.95, c_2 = -0.7445, c_3 = -2.5594$
- ou seja, o "melhor" (no sentido de MQ) modelo é:

$$y = -1.9500 - 0.7445 \cos 2\pi t - 2.5594 \sin 2\pi t$$

Dados periódicos: Temperatura

- Determinamos e resolvemos as equações normais $A^T A c = A^T b$:

$$\begin{bmatrix} 8 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} -15.6 \\ -2.9778 \\ -10.2376 \end{bmatrix}$$

- este sistema tem solução (imediata) $c_1 = -1.95, c_2 = -0.7445, c_3 = -2.5594$
- ou seja, o "melhor" (no sentido de MQ) modelo é:

$$y = -1.9500 - 0.7445 \cos 2\pi t - 2.5594 \sin 2\pi t$$

- A raiz do erro quadrático médio (mínima!) é $REQM \approx 1.063$

Demonstração

Dados periódicos: Temperatura

- Vamos agora tentar "aprimorar" nosso modelo (mesmos dados):

$$y = c_1 + c_2 \cos 2\pi t + c_3 \sin 2\pi t + c_4 \cos 4\pi t$$

Dados periódicos: Temperatura

- Vamos agora tentar "aprimorar" nosso modelo (mesmos dados):

$$y = c_1 + c_2 \cos 2\pi t + c_3 \sin 2\pi t + c_4 \cos 4\pi t$$

- Qual o período deste novo termo?

Dados periódicos: Temperatura

- Vamos agora tentar "aprimorar" nosso modelo (mesmos dados):

$$y = c_1 + c_2 \cos 2\pi t + c_3 \sin 2\pi t + c_4 \cos 4\pi t$$

- Qual o período deste novo termo?
- Será que o ajuste vai "melhorar"?

Dados periódicos: Temperatura

- Em forma matricial, temos o novo sistema

$$
\begin{aligned}
c_1 + c_2 \cos 2\pi(0) + c_3 \sin 2\pi(0) + c_4 \cos 4\pi(0) &= -2.2 \\
c_1 + c_2 \cos 2\pi\left(\frac{1}{8}\right) + c_3 \sin 2\pi\left(\frac{1}{8}\right) + c_4 \cos 4\pi\left(\frac{1}{8}\right) &= -2.8 \\
c_1 + c_2 \cos 2\pi\left(\frac{1}{4}\right) + c_3 \sin 2\pi\left(\frac{1}{4}\right) + c_4 \cos 4\pi\left(\frac{1}{4}\right) &= -6.1 \\
c_1 + c_2 \cos 2\pi\left(\frac{3}{8}\right) + c_3 \sin 2\pi\left(\frac{3}{8}\right) + c_4 \cos 4\pi\left(\frac{3}{8}\right) &= -3.9 \\
c_1 + c_2 \cos 2\pi\left(\frac{1}{2}\right) + c_3 \sin 2\pi\left(\frac{1}{2}\right) + c_4 \cos 4\pi\left(\frac{1}{2}\right) &= 0.0 \\
c_1 + c_2 \cos 2\pi\left(\frac{5}{8}\right) + c_3 \sin 2\pi\left(\frac{5}{8}\right) + c_4 \cos 4\pi\left(\frac{5}{8}\right) &= 1.1 \\
c_1 + c_2 \cos 2\pi\left(\frac{3}{4}\right) + c_3 \sin 2\pi\left(\frac{3}{4}\right) + c_4 \cos 4\pi\left(\frac{3}{4}\right) &= -0.6 \\
c_1 + c_2 \cos 2\pi\left(\frac{7}{8}\right) + c_3 \sin 2\pi\left(\frac{7}{8}\right) + c_4 \cos 4\pi\left(\frac{7}{8}\right) &= -1.1
\end{aligned}
$$

Dados periódicos: Temperatura

- Determinamos e resolvemos as equações normais $A^TAc = A^Tb$:

$$\begin{bmatrix} 8 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} -15.6 \\ -2.9778 \\ -10.2376 \\ 4.5 \end{bmatrix}$$

Dados periódicos: Temperatura

- Determinamos e resolvemos as equações normais $A^T A c = A^T b$:

$$\begin{bmatrix} 8 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} -15.6 \\ -2.9778 \\ -10.2376 \\ 4.5 \end{bmatrix}$$

- este sistema tem solução (imediata) $c_1 = -1.95, c_2 = -0.7445, c_3 = -2.5594, c_4 = 1.125$

Dados periódicos: Temperatura

- Determinamos e resolvemos as equações normais $A^T Ac = A^T b$:

$$\begin{bmatrix} 8 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} -15.6 \\ -2.9778 \\ -10.2376 \\ 4.5 \end{bmatrix}$$

- este sistema tem solução (imediata) $c_1 = -1.95, c_2 = -0.7445, c_3 = -2.5594, c_4 = 1.125$
- ou seja, o "melhor" (no sentido de MQ) modelo é:

$$y = -1.9500 - 0.7445 \cos 2\pi t - 2.5594 \sin 2\pi t + 1.125 \cos 4\pi t$$

Dados periódicos: Temperatura

- Determinamos e resolvemos as equações normais $A^T A c = A^T b$:

$$\begin{bmatrix} 8 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} -15.6 \\ -2.9778 \\ -10.2376 \\ 4.5 \end{bmatrix}$$

- este sistema tem solução (imediata) $c_1 = -1.95, c_2 = -0.7445, c_3 = -2.5594, c_4 = 1.125$
- ou seja, o "melhor" (no sentido de MQ) modelo é:

$$y = -1.9500 - 0.7445 \cos 2\pi t - 2.5594 \sin 2\pi t + 1.125 \cos 4\pi t$$

- A raiz do erro quadrático médio (mínima!) é $REQM \approx 0.705$

Dados periódicos: Temperatura

- Determinamos e resolvemos as equações normais $A^T Ac = A^T b$:

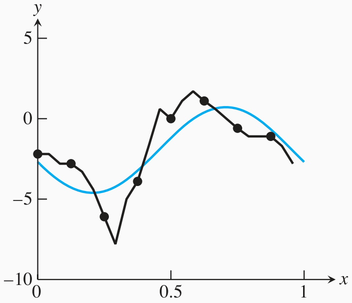
$$\begin{bmatrix} 8 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} -15.6 \\ -2.9778 \\ -10.2376 \\ 4.5 \end{bmatrix}$$

- este sistema tem solução (imediata) $c_1 = -1.95, c_2 = -0.7445, c_3 = -2.5594, c_4 = 1.125$
- ou seja, o "melhor" (no sentido de MQ) modelo é:

$$y = -1.9500 - 0.7445 \cos 2\pi t - 2.5594 \sin 2\pi t + 1.125 \cos 4\pi t$$

- A raiz do erro quadrático médio (mínima!) é $REQM \approx 0.705$
- "Melhorou" em relação ao anterior ($REQM \approx 1.063$)

Dados periódicos: Temperatura



Parênteses: Ortogonalidade

- Foi difícil resolver os sistemas anteriores?

Parênteses: Ortogonalidade

- Foi difícil resolver os sistemas anteriores?
- No exemplo anterior, as equações normais já saíram em forma diagonal.

Parênteses: Ortogonalidade

- Foi difícil resolver os sistemas anteriores?
- No exemplo anterior, as equações normais já saíram em forma diagonal.
- O problema dos mínimos quadrados pode ser simplificado consideravelmente com uma escolha cuidadosa das funções de base.

Parênteses: Ortogonalidade

- Foi difícil resolver os sistemas anteriores?
- No exemplo anterior, as equações normais já saíram em forma diagonal.
- O problema dos mínimos quadrados pode ser simplificado consideravelmente com uma escolha cuidadosa das funções de base.
- Esta é a essência das *expansões de Fourier*.

Linearização dos dados

- Quando a taxa de crescimento de uma população é proporcional ao seu tamanho, temos qual tipo de crescimento populacional?

Linearização dos dados

- Quando a taxa de crescimento de uma população é proporcional ao seu tamanho, temos qual tipo de crescimento populacional?
- Em condições controladas e ideais (ambiente controlado, "imutável"), população bem menor do que a *capacidade de carga* do ambiente, crescimento exponencial é um bom modelo:

$$y = c_1 e^{c_2 t}$$

Linearização dos dados

- Quando a taxa de crescimento de uma população é proporcional ao seu tamanho, temos qual tipo de crescimento populacional?
- Em condições controladas e ideais (ambiente controlado, "imutável"), população bem menor do que a *capacidade de carga* do ambiente, crescimento exponencial é um bom modelo:

$$y = c_1 e^{c_2 t}$$

- Sabemos ajustar uma função exponencial via MMQ?

Linearização dos dados

- Quando a taxa de crescimento de uma população é proporcional ao seu tamanho, temos qual tipo de crescimento populacional?
- Em condições controladas e ideais (ambiente controlado, "imutável"), população bem menor do que a *capacidade de carga* do ambiente, crescimento exponencial é um bom modelo:

$$y = c_1 e^{c_2 t}$$

- Sabemos ajustar uma função exponencial via MMQ?
- Não: temos uma equação não-linear **nos parâmetros**, de modo que nossas técnicas não funcionam.

Linearização dos dados

- Quando a taxa de crescimento de uma população é proporcional ao seu tamanho, temos qual tipo de crescimento populacional?
- Em condições controladas e ideais (ambiente controlado, "imutável"), população bem menor do que a *capacidade de carga* do ambiente, crescimento exponencial é um bom modelo:

$$y = c_1 e^{c_2 t}$$

- Sabemos ajustar uma função exponencial via MMQ?
- Não: temos uma equação não-linear **nos parâmetros**, de modo que nossas técnicas não funcionam.
- Duas soluções:

Linearização dos dados

- Quando a taxa de crescimento de uma população é proporcional ao seu tamanho, temos qual tipo de crescimento populacional?
- Em condições controladas e ideais (ambiente controlado, "imutável"), população bem menor do que a *capacidade de carga* do ambiente, crescimento exponencial é um bom modelo:

$$y = c_1 e^{c_2 t}$$

- Sabemos ajustar uma função exponencial via MMQ?
- Não: temos uma equação não-linear **nos parâmetros**, de modo que nossas técnicas não funcionam.
- Duas soluções:
    - "Linearizar" o modelo

Linearização dos dados

- Quando a taxa de crescimento de uma população é proporcional ao seu tamanho, temos qual tipo de crescimento populacional?
- Em condições controladas e ideais (ambiente controlado, "imutável"), população bem menor do que a *capacidade de carga* do ambiente, crescimento exponencial é um bom modelo:

$$y = c_1 e^{c_2 t}$$

- Sabemos ajustar uma função exponencial via MMQ?
- Não: temos uma equação não-linear **nos parâmetros**, de modo que nossas técnicas não funcionam.
- Duas soluções:
  - "Linearizar" o modelo
  - MMQ não linear (curso de Análise Numérica!)

Linearização dos dados

- Para linearizar o modelo exponencial, basta tomar o logaritmo em ambos os lados:

$$\ln y = \ln(c_1 e^{c_2 t}) = \ln c_1 + c_2 t$$

Linearização dos dados

- Para linearizar o modelo exponencial, basta tomar o logaritmo em ambos os lados:

$$\ln y = \ln(c_1 e^{c_2 t}) = \ln c_1 + c_2 t$$

- Ou seja, o gráfico de $\ln y$ é uma linha reta em t (para isso que servem gráficos em escala logarítmica!)

Linearização dos dados

- Para linearizar o modelo exponencial, basta tomar o logaritmo em ambos os lados:

$$\ln y = \ln(c_1 e^{c_2 t}) = \ln c_1 + c_2 t$$

- Ou seja, o gráfico de $\ln y$ é uma linha reta em t (para isso que servem gráficos em escala logarítmica!)
- Mas isso resolveu o problema? Apesar de c_2 ter passado a a ser linear, c_1 deixou de o ser...

Linearização dos dados

- Para linearizar o modelo exponencial, basta tomar o logaritmo em ambos os lados:

$$\ln y = \ln(c_1 e^{c_2 t}) = \ln c_1 + c_2 t$$

- Ou seja, o gráfico de $\ln y$ é uma linha reta em t (para isso que servem gráficos em escala logarítmica!)
- Mas isso resolveu o problema? Apesar de c_2 ter passado a a ser linear, c_1 deixou de o ser...
- Não importa! O logaritmo de uma constante é outra constante. Criamos $k = \ln c_1$:

$$\ln y = k + c_2 t$$

Linearização dos dados

- Para linearizar o modelo exponencial, basta tomar o logaritmo em ambos os lados:

$$\ln y = \ln(c_1 e^{c_2 t}) = \ln c_1 + c_2 t$$

- Ou seja, o gráfico de $\ln y$ é uma linha reta em t (para isso que servem gráficos em escala logarítmica!)
- Mas isso resolveu o problema? Apesar de c_2 ter passado a a ser linear, c_1 deixou de o ser...
- Não importa! O logaritmo de uma constante é outra constante. Criamos $k = \ln c_1$:

$$\ln y = k + c_2 t$$

- Temos então um modelo linear nos parâmetros, que sabemos ajustar via MMQ.

Linearização dos dados

- Note que tivemos que mudar o problema para eliminar os coeficientes não lineares.

Linearização dos dados

- Note que tivemos que mudar o problema para eliminar os coeficientes não lineares.
- O problema de MQ original era ajustar um modelo exponencial, ou seja encontrar c_1, c_2 que minimizassem

$$\left(c_1 e^{c_2 t_1} - y_1\right)^2 + \ldots + \left(c_1 e^{c_2 t_m} - y_m\right)^2$$

Linearização dos dados

- Note que tivemos que mudar o problema para eliminar os coeficientes não lineares.
- O problema de MQ original era ajustar um modelo exponencial, ou seja encontrar c_1, c_2 que minimizassem

$$\left(c_1 e^{c_2 t_1} - y_1\right)^2 + \ldots + \left(c_1 e^{c_2 t_m} - y_m\right)^2$$

- Agora, estamos resolvendo o modelo modificado minimizando o erro num "espaço logarítmico", ou seja c_1, c_2 que minimizem

$$(\ln c_1 + c_2 t_1 - \ln y_1)^2 + \ldots + (\ln c_1 + c_2 t_m - \ln y_m)^2$$

Linearização dos dados

- Note que tivemos que mudar o problema para eliminar os coeficientes não lineares.
- O problema de MQ original era ajustar um modelo exponencial, ou seja encontrar c_1, c_2 que minimizassem

$$\left(c_1 e^{c_2 t_1} - y_1\right)^2 + \ldots + \left(c_1 e^{c_2 t_m} - y_m\right)^2$$

- Agora, estamos resolvendo o modelo modificado minimizando o erro num "espaço logarítmico", ou seja c_1, c_2 que minimizem

$$(\ln c_1 + c_2 t_1 - \ln y_1)^2 + \ldots + (\ln c_1 + c_2 t_m - \ln y_m)^2$$

- Somente o primeiro caso é um problema de MQ segundo nossa definição. Em geral, os c_i vão ser distintos para os dois casos. Dependendo do contexto dos dados, minimizar erros ou log(erros) pode ser mais conveniente.

Linearização dos dados: Carros

- Vamos encontrar o "melhor" ajuste exponecial $y = c_1 e^{c_2 t}$ para o número de carros em circulação no mundo, como função do tempo:

ano	carros ($\times 10^6$)
1950	53.05
1955	73.04
1960	98.31
1965	139.78
1970	193.48
1975	260.20
1980	320.39

Linearização dos dados: Carros

- Vamos encontrar o "melhor" ajuste exponecial $y = c_1 e^{c_2 t}$ para o número de carros em circulação no mundo, como função do tempo:

ano	carros ($\times 10^6$)
1950	53.05
1955	73.04
1960	98.31
1965	139.78
1970	193.48
1975	260.20
1980	320.39

- definiremos a variável t como sendo o tempo desde 1950, para ajustar

$$\ln y = k + c_2 t$$

Linearização dos dados: Carros

- Resolvendo o problema de MQ, chegamos a $k \approx 3.9896, c_2 \approx 0.06152$. Como $c_1 = e^k \approx e^{3.9896} \approx 54.03$, o modelo é
$$y = 54.03e^{0.06152t}$$

Linearização dos dados: Carros

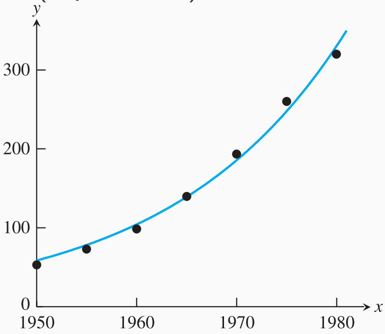
- Resolvendo o problema de MQ, chegamos a $k \approx 3.9896, c_2 \approx 0.06152$. Como $c_1 = e^k \approx e^{3.9896} \approx 54.03$, o modelo é
$$y = 54.03e^{0.06152t}$$
- Problema linearizado (erro no espaço log): $REQM \approx 0.0357$

Linearização dos dados: Carros

- Resolvendo o problema de MQ, chegamos a $k \approx 3.9896, c_2 \approx 0.06152$. Como $c_1 = e^k \approx e^{3.9896} \approx 54.03$, o modelo é

$$y = 54.03e^{0.06152t}$$

- Problema linearizado (erro no espaço log): $REQM \approx 0.0357$
- Problema original (exponencial): $REQM \approx 9.56$

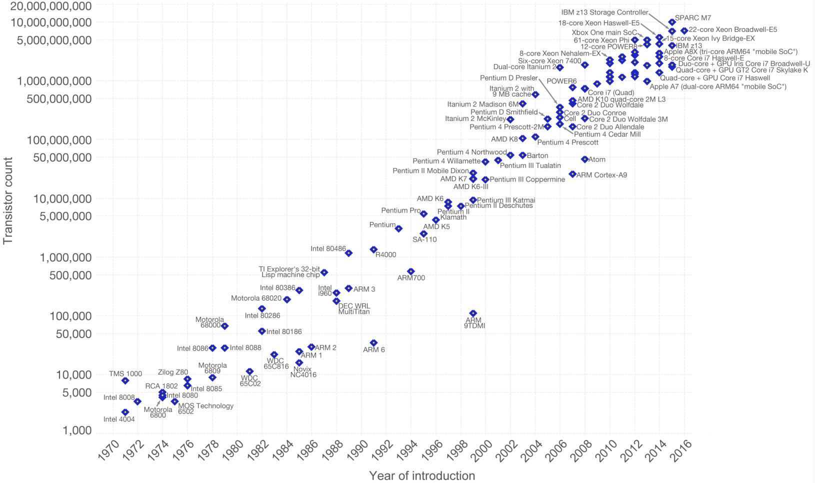


Lei de Moore

Moore's Law – The number of transistors on integrated circuit chips (1971-2016)

Our World in Data

Moore's law describes the empirical regularity that the number of transistors on integrated circuits doubles approximately every two years. This advancement is important as other aspects of technological progress – such as processing speed or the price of electronic products – are strongly linked to Moore's law.



Data source: Wikipedia (https://en.wikipedia.org/wiki/Transistor_count)

The data visualization is available at OurWorldinData.org. There you find more visualizations and research on this topic.

Linearização: CPUs

CPU	ano	transistores
4004	1971	2,250
8008	1972	2,500
8080	1974	5,000
8086	1978	29,000
286	1982	120,000
386	1985	275,000
486	1989	1,180,000
Pentium	1993	3,100,000
Pentium II	1997	7,500,000
Pentium III	1999	24,000,000
Pentium 4	2000	42,000,000
Itanium	2002	220,000,000
Itanium 2	2003	410,000,000

- Ajustaremos um modelo exponencial (linearizado) a estes dados.

Linearização: CPUs

- O modelo linearizado é (tomando $t = 0$ como 1970)

$$y = k + c_2 t$$

Linearização: CPUs

- O modelo linearizado é (tomando $t = 0$ como 1970)

$$y = k + c_2 t$$

- Substituindo os dados no modelo, temos

$$\begin{aligned} k + c_2(1) &= \ln 2250 \\ k + c_2(2) &= \ln 2500 \\ k + c_2(4) &= \ln 5000 \\ k + c_2(8) &= \ln 29000 \\ &\vdots \end{aligned}$$

Linearização: CPUs

- Na forma matricial,

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 4 \\ 1 & 8 \\ \vdots & \\ 1 & 33 \end{bmatrix} \qquad b = \begin{bmatrix} \ln 2250 \\ \ln 2500 \\ \ln 5000 \\ \ln 29000 \\ \vdots \\ \ln 410000000 \end{bmatrix}$$

Linearização: CPUs

- Na forma matricial,

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 4 \\ 1 & 8 \\ \vdots & \\ 1 & 33 \end{bmatrix} \qquad b = \begin{bmatrix} \ln 2250 \\ \ln 2500 \\ \ln 5000 \\ \ln 29000 \\ \vdots \\ \ln 410000000 \end{bmatrix}$$

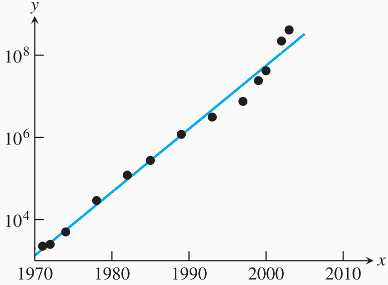
- Equações normais $A^T Ac = A^T b$ (com $c = (k, c_2)$)

$$A = \begin{bmatrix} 13 & 235 \\ 235 & 5927 \end{bmatrix} \begin{bmatrix} k \\ c_2 \end{bmatrix} = \begin{bmatrix} 176.90 \\ 3793.23 \end{bmatrix}$$

Linearização: CPU

- Encontramos a solução
  $k \approx 7.197, c_2 \approx 0.3546 \implies c_1 = e^k \approx 1335.3$:

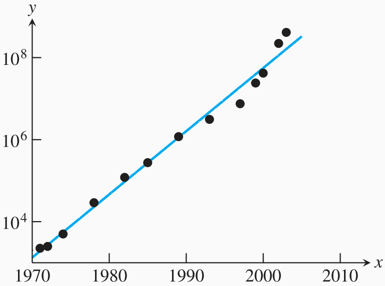
$$y = 1335.3e^{0.3546t}$$



Linearização: CPU

- Encontramos a solução
  $k \approx 7.197, c_2 \approx 0.3546 \implies c_1 = e^k \approx 1335.3$:

$$y = 1335.3e^{0.3546t}$$

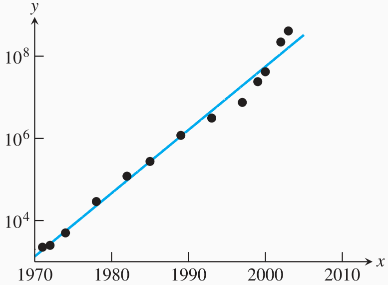


- Tempo para que o número tansistores *dobre*: $\ln \frac{2}{c_2} \approx 1.95$ anos.

Linearização: CPU

- Encontramos a solução
  $k \approx 7.197, c_2 \approx 0.3546 \implies c_1 = e^k \approx 1335.3$:

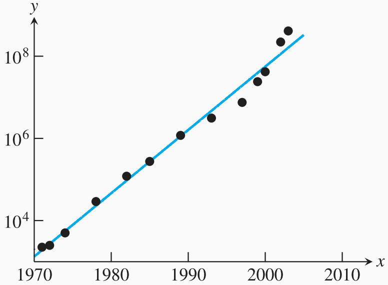
$$y = 1335.3e^{0.3546t}$$



- Tempo para que o número tansistores *dobre*: $\ln \frac{2}{c_2} \approx 1.95$ anos.
- Em 1965 Gordon Moore (Intel) previu que, durante a próxima década, o poder computacional iria dobrar a cada 2 anos.

Linearização: CPU

- Encontramos a solução
  $k \approx 7.197, c_2 \approx 0.3546 \implies c_1 = e^k \approx 1335.3$:
  $$y = 1335.3e^{0.3546t}$$



- Tempo para que o número tansistores *dobre*: $\ln \frac{2}{c_2} \approx 1.95$ anos.
- Em 1965 Gordon Moore (Intel) previu que, durante a próxima década, o poder computacional iria dobrar a cada 2 anos.
- Esta taxa exponencial continuou por mais de 40 anos!

Leis de Potência

- Outro tipo importante de modelo com coeficientes não lineares são as *leis de potência*:

$$y = c_1 t^{c_2}$$

Leis de Potência

- Outro tipo importante de modelo com coeficientes não lineares são as *leis de potência*:

$$y = c_1 t^{c_2}$$

- Também podemos linearizar este tipo de modelo tomando log:

$$\ln y = \ln(c_1 t^{c_2}) = \ln(c_1) + c_2 \ln(t) = k + c_2 \ln(t)$$

Leis de Potência

- Outro tipo importante de modelo com coeficientes não lineares são as *leis de potência*:

$$y = c_1 t^{c_2}$$

- Também podemos linearizar este tipo de modelo tomando log:

$$\ln y = \ln(c_1 t^{c_2}) = \ln(c_1) + c_2 \ln(t) = k + c_2 \ln(t)$$

- Substituindo os dados no modelo, temos

$$\begin{cases} k + c_2 \ln t_1 = \ln y_1 \\ \vdots \\ k + c_2 \ln t_n = \ln y_n \end{cases}$$

Leis de Potência

- Ou, em forma matricial,

$$A = \begin{bmatrix} 1 & \ln t_1 \\ \vdots & \vdots \\ 1 & \ln t_n \end{bmatrix} \qquad b = \begin{bmatrix} \ln y_1 \\ \vdots \\ \ln y_n \end{bmatrix}$$

Leis de Potência

- Ou, em forma matricial,

$$A = \begin{bmatrix} 1 & \ln t_1 \\ \vdots & \vdots \\ 1 & \ln t_n \end{bmatrix} \qquad b = \begin{bmatrix} \ln y_1 \\ \vdots \\ \ln y_n \end{bmatrix}$$

- Através das equações normais, encontramos k, c_2, e $c_1 = e^k$

Leis de Potência: Peso – Altura

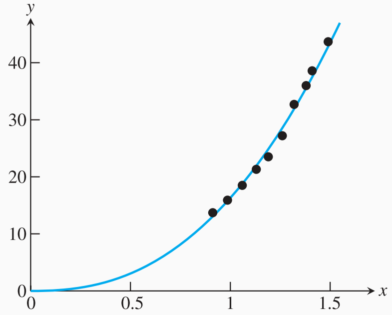
- Vamos ajustar uma lei de potência aos dados de altura-peso (CDC, 2002) abaixo:

idade (anos)	altura (m)	peso (kg)
2	0.9120	13.7
3	0.9860	15.9
4	1.0600	18.5
5	1.1300	21.3
6	1.1900	23.5
7	1.2600	27.2
8	1.3200	32.7
9	1.3800	36.0
10	1.4100	38.6
11	1.4900	43.7

Leis de Potência: Peso – Altura

- A lei de potência resultante é

$$p = 16.3h^{2.42}$$



Leis de Potência: Farmacologia

- A concentração de uma droga na corrente sanguínea é bem descrita por (t é o tempo após a administração da droga):

$$y = c_1 t e^{c_2 t}$$

Leis de Potência: Farmacologia

- A concentração de uma droga na corrente sanguínea é bem descrita por (t é o tempo após a administração da droga):

$$y = c_1 t e^{c_2 t}$$

- Este tipo de modelo reflete um aumento rápido, seguido por um decaimento exponencial mais lento da concentração.

Leis de Potência: Farmacologia

- A concentração de uma droga na corrente sanguínea é bem descrita por (t é o tempo após a administração da droga):

$$y = c_1 t e^{c_2 t}$$

- Este tipo de modelo reflete um aumento rápido, seguido por um decaimento exponencial mais lento da concentração.
- Podemos linearizar este modelo tomando log de cada lado:

$$\ln y = \ln c_1 + \ln t + c_2 t$$

$$k + c_2 t = \ln y - \ln t$$

Leis de Potência: Farmacologia

- A concentração de uma droga na corrente sanguínea é bem descrita por (t é o tempo após a administração da droga):

$$y = c_1 t e^{c_2 t}$$

- Este tipo de modelo reflete um aumento rápido, seguido por um decaimento exponencial mais lento da concentração.
- Podemos linearizar este modelo tomando log de cada lado:

$$\ln y = \ln c_1 + \ln t + c_2 t$$

$$k + c_2 t = \ln y - \ln t$$

- Isso nos leva $Ax = b$, com

$$A = \begin{bmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_n \end{bmatrix} \qquad b = \begin{bmatrix} \ln y_1 - \ln t_1 \\ \vdots \\ \ln y_n - \ln t_n \end{bmatrix}$$

Leis de Potência: Farmacologia

- Vamos ajustar este modelo aos dados de concentração de norfluoxetina (antidepressivo) no sangue de um paciente.

hora	concentração (ng/ml)
1	8.0
2	12.3
3	15.5
4	16.8
5	17.1
6	15.8
7	15.2
8	14.0

Leis de Potência: Farmacologia

- Vamos ajustar este modelo aos dados de concentração de norfluoxetina (antidepressivo) no sangue de um paciente.

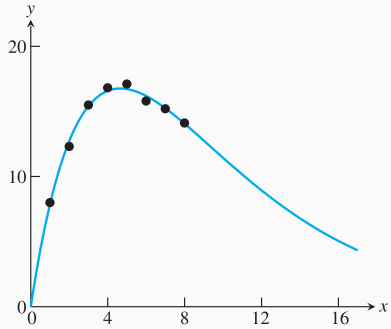
hora	concentração (ng/ml)
1	8.0
2	12.3
3	15.5
4	16.8
5	17.1
6	15.8
7	15.2
8	14.0

- as equações normais nos dão
  $k \approx 2.28, c_2 \approx -0.215 \implies c_1 \approx e^{2.28} \approx 9.77$

Leis de Potência: Farmacologia

- A lei de potência resultante é

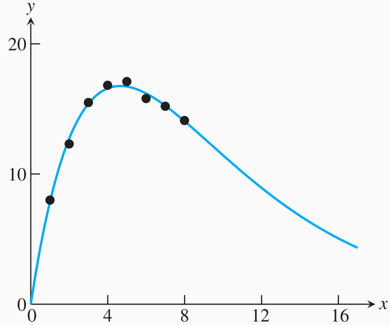
$$y = 9.77te^{-0.215t}$$



Leis de Potência: Farmacologia

- A lei de potência resultante é

$$y = 9.77te^{-0.215t}$$



- Deste gráfico, é possível estimar a *concentração de pico* (máximo da curva) e a meia-vida (tempo que a concentração leva para cair do valor de pico a metade deste valor), quantidades importantes em farmacologia.