Solução de um sistema de equações lineares por mínimos quadrados

Pablo Angulo, Fabricio Macià

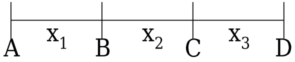
Universidade Pedagogica de Maputo

medições com erros

Na vida real, quando fazemos *medidas*, sempre o fazemos *com um certo erro*. Se mais tarde pretendermos resolver problemas usando essas medições, os erros de medição podem levar a um sistema de equações que não possuem soluções.

medições com erros

Na vida real, quando fazemos *medidas*, sempre o fazemos *com um certo erro*. Se mais tarde pretendermos resolver problemas usando essas medições, os erros de medição podem levar a um sistema de equações que não possuem soluções. Vejamos um exemplo:

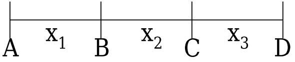


Suponha que tomamos as medidas $AD = 89m$, $AC = 67m$, $BD = 53m$, $AB = 35m$, $CD = 20m$.

Queremos determinar os comprimentos dos segmentos $x_1 = AB$, $x_2 = BC$, $x_3 = CD$.

medições com erros

Na vida real, quando fazemos *medidas*, sempre o fazemos *com um certo erro*. Se mais tarde pretendermos resolver problemas usando essas medições, os erros de medição podem levar a um sistema de equações que não possuem soluções. Vejamos um exemplo:



Suponha que tomamos as medidas
$AD = 89m$, $AC = 67m$, $BD = 53m$,
$AB = 35m$, $CD = 20m$.
Queremos determinar os comprimentos dos segmentos $x_1 = AB$, $x_2 = BC$, $x_3 = CD$.

Se queremos três valores para x_1, x_2, x_3 compatíveis com todas as nossas medições, eles precisam resolver o sistema de equações:

$$\begin{array}{cccl} x_1 & +x_2 & +x_3 & = 89 \\ x_1 & +x_2 & & = 67 \\ & x_2 & +x_3 & = 53 \\ x_1 & & & = 35 \\ & & x_3 & = 20 \end{array}, \quad \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 89 \\ 67 \\ 53 \\ 35 \\ 20 \end{bmatrix}$$

$$
\begin{array}{rrrl}
x_1 & +x_2 & +x_3 & = 89 \\
x_1 & +x_2 & & = 67 \\
 & x_2 & +x_3 & = 53 \\
x_1 & & & = 35 \\
 & & x_3 & = 20
\end{array}
, \quad
\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 89 \\ 67 \\ 53 \\ 35 \\ 20 \end{bmatrix}
$$

Este sistema não tem solução: a terceira e quinta equações implicam que $x_2 = 33$, mas a segunda e a quarta implicam que $x_2 = 32$.
Uma solução possível é ignorar algumas equações.

$$\begin{array}{llll} x_1 & +x_2 & +x_3 & = 89 \\ x_1 & +x_2 & & = 67 \\ & x_2 & +x_3 & = 53 \\ x_1 & & & = 35 \\ & & x_3 & = 20 \end{array}, \quad \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 89 \\ 67 \\ 53 \\ 35 \\ 20 \end{bmatrix}$$

Este sistema não tem solução: a terceira e quinta equações implicam que $x_2 = 33$, mas a segunda e a quarta implicam que $x_2 = 32$.
Uma solução possível é ignorar algumas equações.
Uma solução melhor é procurar os valores de x_1, x_2, x_3 que *minimizar o erro total*, para alguma noção sensata de erro total.

Definición

Erro quadrático total cometido por $\mathbf{x}$ no sistema $A \cdot \mathbf{x} = \mathbf{b}$ É o padrão quadrado da diferença entre $\mathbf{b}$ e $A \cdot \mathbf{x}$:

$$\|\mathbf{b} - A \cdot \mathbf{x}\|^2$$

um único desconhecido

Portanto, antes de um sistema de equações lineares $A \cdot \mathbf{x} = \mathbf{b}$ que não possui solução, queremos encontrar o vetor $\mathbf{x}$ cujo erro quadrático total $\|\mathbf{b} - A \cdot \mathbf{x}\|^2$ é mínimo.

um único desconhecido

Portanto, antes de um sistema de equações lineares $A \cdot \mathbf{x} = \mathbf{b}$ que não possui solução, queremos encontrar o vetor $\mathbf{x}$ cujo erro quadrático total $\|\mathbf{b} - A \cdot \mathbf{x}\|^2$ é mínimo.

Se tivermos apenas um desconhecido t, a matriz A é um vetor de coluna $\mathbf{a}$, e temos que minimizar $\|\mathbf{b} - t\mathbf{a}\|^2$.

um único desconhecido

Portanto, antes de um sistema de equações lineares $A \cdot \mathbf{x} = \mathbf{b}$ que não possui solução, queremos encontrar o vetor $\mathbf{x}$ cujo erro quadrático total $\|\mathbf{b} - A \cdot \mathbf{x}\|^2$ é mínimo.
Se tivermos apenas um desconhecido t, a matriz A é um vetor de coluna $\mathbf{a}$, e temos que minimizar $\|\mathbf{b} - t\mathbf{a}\|^2$.
Quando t viaja todos os valores reais, $t\mathbf{a}$ viaja todos os vetores da linha vetorial $\langle \mathbf{a} \rangle$.

um único desconhecido

Portanto, antes de um sistema de equações lineares $A \cdot \mathbf{x} = \mathbf{b}$ que não possui solução, queremos encontrar o vetor $\mathbf{x}$ cujo erro quadrático total $\|\mathbf{b} - A \cdot \mathbf{x}\|^2$ é mínimo.
Se tivermos apenas um desconhecido t, a matriz A é um vetor de coluna $\mathbf{a}$, e temos que minimizar $\|\mathbf{b} - t\mathbf{a}\|^2$.
Quando t viaja todos os valores reais, $t\mathbf{a}$ viaja todos os vetores da linha vetorial $\langle \mathbf{a} \rangle$.
Em outras palavras, *procuramos o vetor de* $E = \langle \mathbf{a} \rangle$ *mais próximo de* $\mathbf{b}$.

um único desconhecido

Portanto, antes de um sistema de equações lineares $A \cdot \mathbf{x} = \mathbf{b}$ que não possui solução, queremos encontrar o vetor $\mathbf{x}$ cujo erro quadrático total $\|\mathbf{b} - A \cdot \mathbf{x}\|^2$ é mínimo.
Se tivermos apenas um desconhecido t, a matriz A é um vetor de coluna $\mathbf{a}$, e temos que minimizar $\|\mathbf{b} - t\mathbf{a}\|^2$.
Quando t viaja todos os valores reais, $t\mathbf{a}$ viaja todos os vetores da linha vetorial $\langle \mathbf{a} \rangle$.
Em outras palavras, *procuramos o vetor de $E = \langle \mathbf{a} \rangle$ mais próximo de* $\mathbf{b}$.
Já sabemos a solução: $t\mathbf{a}$ é a projeção ortogonal de $\mathbf{b}$ em $E = \langle \mathbf{a} \rangle$:

$$t\mathbf{a} = P_E(\mathbf{b}) = \left(\frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\|^2} \right) \mathbf{a}, \quad \text{então} \quad t = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\|^2}$$

ou, na notação matricial

$$t = \frac{\mathbf{a}^T \mathbf{b}}{\mathbf{a}^T \mathbf{a}}$$

Várias incógnitas

Lembramos da propriedade da distância mínima da projeção ortogonal: E É um subespaço vetorial de $\mathbb{R}^d$ de qualquer dimensão e $\mathbf{w}^* = P_E\mathbf{b}$ é a projeção ortogonal de $\mathbf{b}$ em E.

Para qualquer vetor $\mathbf{w} \in E$, você tem que

$$\|\mathbf{w} - \mathbf{b}\| \geq \|\mathbf{w}^* - \mathbf{b}\|$$

Então, se tivermos várias incógnitas:

- O sistema $A \cdot \mathbf{x} = \mathbf{b}$ não possui solução exata.

Várias incógnitas

Lembramos da propriedade da distância mínima da projeção ortogonal: E É um subespaço vetorial de $\mathbb{R}^d$ de qualquer dimensão e $\mathbf{w}^* = P_E\mathbf{b}$ é a projeção ortogonal de $\mathbf{b}$ em E.

Para qualquer vetor $\mathbf{w} \in E$, você tem que

$$\|\mathbf{w} - \mathbf{b}\| \geq \|\mathbf{w}^* - \mathbf{b}\|$$

Então, se tivermos várias incógnitas:

- O sistema $A \cdot \mathbf{x} = \mathbf{b}$ não possui solução exata.
- As colunas $\mathbf{a}_1, \ldots, \mathbf{a}_n$ de A geraram o subespaço vetorial E

Várias incógnitas

Lembramos da propriedade da distância mínima da projeção ortogonal: E É um subespaço vetorial de $\mathbb{R}^d$ de qualquer dimensão e $\mathbf{w}^* = P_E\mathbf{b}$ é a projeção ortogonal de $\mathbf{b}$ em E.

Para qualquer vetor $\mathbf{w} \in E$, você tem que

$$\|\mathbf{w} - \mathbf{b}\| \geq \|\mathbf{w}^* - \mathbf{b}\|$$

Então, se tivermos várias incógnitas:

- O sistema $A \cdot \mathbf{x} = \mathbf{b}$ não possui solução exata.
- As colunas $\mathbf{a}_1, \ldots, \mathbf{a}_n$ de A geraram o subespaço vetorial E
- Queremos $A \cdot \mathbf{x} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$ o mais próximo possível de $\mathbf{b}$.

Várias incógnitas

Lembramos da propriedade da distância mínima da projeção ortogonal:
E É um subespaço vetorial de $\mathbb{R}^d$ de qualquer dimensão e $\mathbf{w}^* = P_E\mathbf{b}$ é a projeção ortogonal de $\mathbf{b}$ em E.

Para qualquer vetor $\mathbf{w} \in E$, você tem que

$$\|\mathbf{w} - \mathbf{b}\| \geq \|\mathbf{w}^* - \mathbf{b}\|$$

Então, se tivermos várias incógnitas:

- O sistema $A \cdot \mathbf{x} = \mathbf{b}$ não possui solução exata.
- As colunas $\mathbf{a}_1, \dots, \mathbf{a}_n$ de A geraram o subespaço vetorial E
- Queremos $A \cdot \mathbf{x} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$ o mais próximo possível de $\mathbf{b}$.
- Queremos encontrar $x_1 \dots x_n$ de modo que $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$ seja a projeção ortogonal de $\mathbf{b}$ em E.

Objetivo original: Encontre o vetor $\mathbf{x}$ de modo que $\|A \cdot \mathbf{x} - \mathbf{b}\|$ seja mínimo.

Objetivo original: Encontre o vetor $\mathbf{x}$ de modo que $\|A \cdot \mathbf{x} - \mathbf{b}\|$ seja mínimo.

Em outras palavras: Encontre o vetor $\mathbf{w}$ de $E = \langle \mathbf{a}_1, \dots, \mathbf{a}_n \rangle$ mais próximo de $\mathbf{b}$ ($\mathbf{a}_1, \dots, \mathbf{a}_n$ estão *os vetores da coluna de* A).

Mas cuidado: Também queremos saber $x_1, \dots, x_n$ (ou seja, o vetor $\mathbf{x}$), que resolve

$$A\mathbf{x} = \mathbf{w}.$$

Objetivo original: Encontre o vetor $\mathbf{x}$ de modo que $\|A \cdot \mathbf{x} - \mathbf{b}\|$ seja mínimo.

Em outras palavras: Encontre o vetor $\mathbf{w}$ de $E = \langle \mathbf{a}_1, \ldots, \mathbf{a}_n \rangle$ mais próximo de $\mathbf{b}$ ($\mathbf{a}_1, \ldots, \mathbf{a}_n$ estão *os vetores da coluna de* A).

Mas cuidado: Também queremos saber $x_1, \ldots, x_n$ (ou seja, o vetor $\mathbf{x}$), que resolve

$$A\mathbf{x} = \mathbf{w}.$$

Objetivo equivalente: Encontre a projeção ortogonal $P_E\mathbf{b}$ de $\mathbf{b}$ em E e as coordenadas de $P_E\mathbf{b}$ na base $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$.

$$A\mathbf{x} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = P_E\mathbf{b}.$$

Este sistema possui uma solução (diferentemente $A \cdot \mathbf{x} = \mathbf{b}$), porque $P_E\mathbf{b} \in E$, e o subespaço E é gerado pelas colunas de A. Além disso, *A solução é única se os vetores da coluna de* A *forem linearmente independentes*.

Objetivo original: Encontre o vetor $\mathbf{x}$ de modo que $\|A \cdot \mathbf{x} - \mathbf{b}\|$ seja mínimo.
Em outras palavras: Encontre o vetor $\mathbf{w}$ de $E = \langle \mathbf{a}_1, \dots, \mathbf{a}_n \rangle$ mais próximo de $\mathbf{b}$ ($\mathbf{a}_1, \dots, \mathbf{a}_n$ estão *os vetores da coluna de* A).
Mas cuidado: Também queremos saber $x_1, \dots, x_n$ (ou seja, o vetor $\mathbf{x}$), que resolve

$$A\mathbf{x} = \mathbf{w}.$$

Objetivo equivalente: Encontre a projeção ortogonal $P_E\mathbf{b}$ de $\mathbf{b}$ em E e as coordenadas de $P_E\mathbf{b}$ na base $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$.

$$A\mathbf{x} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = P_E\mathbf{b}.$$

Este sistema possui uma solução (diferentemente $A \cdot \mathbf{x} = \mathbf{b}$), porque $P_E\mathbf{b} \in E$, e o subespaço E é gerado pelas colunas de A. Além disso, *A solução é única se os vetores da coluna de A forem linearmente independentes*.
Podemos resolver esse problema com as ferramentas que conhecemos, mas procuraremos *outro método mais conveniente*.

Como $A\mathbf{x} = P_E\mathbf{b}$ é a projeção ortogonal de $\mathbf{b}$ em E, o termo

$$A\mathbf{x} - \mathbf{b} = P_E\mathbf{b} - \mathbf{b}$$

Tem que ser ortogonal para E.

Como $A\mathbf{x} = P_E\mathbf{b}$ é a projeção ortogonal de $\mathbf{b}$ em E, o termo

$$A\mathbf{x} - \mathbf{b} = P_E\mathbf{b} - \mathbf{b}$$

Tem que ser ortogonal para E. Como Éntão,

$$\mathbf{a}_1 \cdot (A\mathbf{x} - \mathbf{b}) = 0, \quad \mathbf{a}_2 \cdot (A\mathbf{x} - \mathbf{b}) = 0, ..., \quad \mathbf{a}_n \cdot (A\mathbf{x} - \mathbf{b}) = 0.$$

Como $A\mathbf{x} = P_E\mathbf{b}$ é a projeção ortogonal de $\mathbf{b}$ em E, o termo

$$A\mathbf{x} - \mathbf{b} = P_E\mathbf{b} - \mathbf{b}$$

Tem que ser ortogonal para E. Como Éntão,

$$\mathbf{a}_1 \cdot (A\mathbf{x} - \mathbf{b}) = 0, \quad \mathbf{a}_2 \cdot (A\mathbf{x} - \mathbf{b}) = 0, ..., \quad \mathbf{a}_n \cdot (A\mathbf{x} - \mathbf{b}) = 0.$$

Escrito como um produto Matrices, ou seja:

$$(\mathbf{a}_1)^T (A\mathbf{x} - \mathbf{b}) = 0, \quad (\mathbf{a}_2)^T (A\mathbf{x} - \mathbf{b}) = 0, ..., \quad (\mathbf{a}_n)^T (A\mathbf{x} - \mathbf{b}) = 0,$$

Como $A\mathbf{x} = P_E\mathbf{b}$ é a projeção ortogonal de $\mathbf{b}$ em E, o termo

$$A\mathbf{x} - \mathbf{b} = P_E\mathbf{b} - \mathbf{b}$$

Tem que ser ortogonal para E. Como Éntão,

$$\mathbf{a}_1 \cdot (A\mathbf{x} - \mathbf{b}) = 0, \quad \mathbf{a}_2 \cdot (A\mathbf{x} - \mathbf{b}) = 0, ..., \quad \mathbf{a}_n \cdot (A\mathbf{x} - \mathbf{b}) = 0.$$

Escrito como um produto Matrices, ou seja:

$$(\mathbf{a}_1)^T (A\mathbf{x} - \mathbf{b}) = 0, \quad (\mathbf{a}_2)^T (A\mathbf{x} - \mathbf{b}) = 0, ..., \quad (\mathbf{a}_n)^T (A\mathbf{x} - \mathbf{b}) = 0,$$

Ou o que é o mesmo:

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

isto é:

$$A^T A\mathbf{x} = A^T\mathbf{b}.$$

A matriz $B = A^T A$ possui linhas n e n colunas.

Portanto, o vetor **x** pesquisado é uma solução de:

$$A^T A\mathbf{x} = A^T \mathbf{b}.$$

Este sistema tem alguma solução?É único?

Portanto, o vetor $\mathbf{x}$ pesquisado é uma solução de:

$$A^T A\mathbf{x} = A^T \mathbf{b}.$$

Este sistema tem alguma solução?É único?
Sabemos que o sistema tem uma solução única se e somente se o sistema

$$A^T A\mathbf{x} = \mathbf{0}.$$

Tem uma solução única.Seja $\mathbf{x}$ de modo que $A^T A\mathbf{x} = \mathbf{0}$.
Isso é equivalente a dizer que $A\mathbf{x}$ é ortogonal para $\mathbf{a}_1$, ..., $\mathbf{a}_n$, porque

$$A^T \mathbf{v} = \begin{bmatrix} \mathbf{a}_1^T \mathbf{v} \\ \dots \\ \mathbf{a}_n^T \mathbf{v} \end{bmatrix}$$

Mas, como $A\mathbf{x} \in E$, deve ser $A\mathbf{x} = \mathbf{0}$.

Portanto, o vetor $\mathbf{x}$ pesquisado é uma solução de:

$$A^T A\mathbf{x} = A^T\mathbf{b}.$$

Este sistema tem alguma solução?É único?
Sabemos que o sistema tem uma solução única se e somente se o sistema

$$A^T A\mathbf{x} = \mathbf{0}.$$

Tem uma solução única.Seja $\mathbf{x}$ de modo que $A^T A\mathbf{x} = \mathbf{0}$.
Isso é equivalente a dizer que $A\mathbf{x}$ é ortogonal para $\mathbf{a}_1$, ..., $\mathbf{a}_n$, porque

$$A^T\mathbf{v} = \begin{bmatrix} \mathbf{a}_1^T\mathbf{v} \\ \dots \\ \mathbf{a}_n^T\mathbf{v} \end{bmatrix}$$

Mas, como $A\mathbf{x} \in E$, deve ser $A\mathbf{x} = \mathbf{0}$.
Se as colunas de A são linearmente independentes, então $A\mathbf{x} = \mathbf{0}$ implica $\mathbf{x} = \mathbf{0}$.

Solução por quadrados mínimos

Resumo:

- Eles elevam um sistema de equações $A \cdot \mathbf{x} = \mathbf{b}$ que não possui solução.
- As colunas $\mathbf{a}_1, \ldots, \mathbf{a}_n$ de A geraram o subespaço vetorial E
- Queremos encontrar $x_1 \ldots x_n$ de modo que $a_1\mathbf{x}_1 + \cdots + a_n\mathbf{x}_n$ seja a projeção ortogonal de $\mathbf{b}$ em E.
- *Se as colunas de A forem linearmente independentes*, então podemos encontrá-lo resolvendo:

$$A^T A\mathbf{x} = A^T\mathbf{b}.$$

Solução por quadrados mínimos

Resumo:

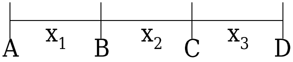
- Eles elevam um sistema de equações $A \cdot \mathbf{x} = \mathbf{b}$ que não possui solução.
- As colunas $\mathbf{a}_1, \ldots, \mathbf{a}_n$ de A geraram o subespaço vetorial E
- Queremos encontrar $x_1 \ldots x_n$ de modo que $a_1\mathbf{x}_1 + \cdots + a_n\mathbf{x}_n$ seja a projeção ortogonal de $\mathbf{b}$ em E.
- *Se as colunas de A forem linearmente independentes*, então podemos encontrá-lo resolvendo:

$$A^T A\mathbf{x} = A^T\mathbf{b}.$$

Seja A uma matriz $d \times n$ cujos vetores de coluna são linearmente independentes e é $\mathbf{b}$ um vetor de $\mathbb{R}^d$.
O **solução por quadrados mínimos** de $A\mathbf{x} = \mathbf{b}$ é:

$$\mathbf{x} = (A^T A)^{-1} A^T \mathbf{b}$$

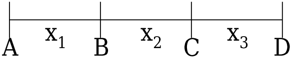
Podemos resolver o problema com o qual começamos?



$$A \cdot \mathbf{x} = \mathbf{b}$$

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 89 \\ 67 \\ 53 \\ 35 \\ 20 \end{bmatrix}$$

Podemos resolver o problema com o qual começamos?



$$A \cdot \mathbf{x} = \mathbf{b}$$

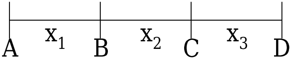
$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 89 \\ 67 \\ 53 \\ 35 \\ 20 \end{bmatrix}$$

Nesse caso, queremos encontrar x_1, x_2, x_3 para que

$$A \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

Esteja o mais próximo possível do vetor **b**.

Podemos resolver o problema com o qual começamos?



$$A \cdot \mathbf{x} = \mathbf{b}$$

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 89 \\ 67 \\ 53 \\ 35 \\ 20 \end{bmatrix}$$

Nesse caso, queremos encontrar x_1, x_2, x_3 para que

$$A \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

Esteja o mais próximo possível do vetor $\mathbf{b}$. Quando x_1, x_2, x_3 viaja todas as listas de valores reais, $A \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ viaja todos os vetores do subespaço $\langle \mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3 \rangle$, onde $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ são as colunas de A.

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 89 \\ 67 \\ 53 \\ 35 \\ 20 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 89 \\ 67 \\ 53 \\ 35 \\ 20 \end{bmatrix}$$

A matriz $A^T A$ é:

$$A^T A = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 3 \end{bmatrix}, \; A^T b = \begin{bmatrix} 191 \\ 209 \\ 162 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 89 \\ 67 \\ 53 \\ 35 \\ 20 \end{bmatrix}$$

A matriz $A^T A$ é:

$$A^T A = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 3 \end{bmatrix}, \; A^T b = \begin{bmatrix} 191 \\ 209 \\ 162 \end{bmatrix}$$

A solução é:

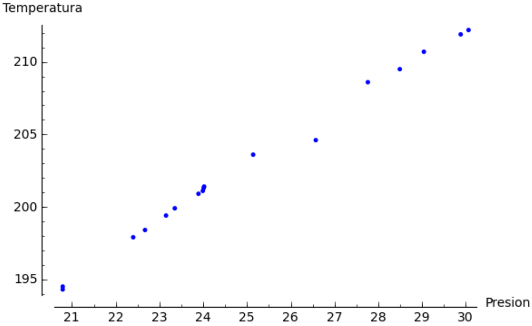
$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = (A^T A)^{-1} A^T \mathbf{b} = \begin{bmatrix} \frac{5}{8} & -\frac{1}{2} & \frac{1}{8} \\ -\frac{1}{2} & 1 & -\frac{1}{2} \\ \frac{1}{8} & -\frac{1}{2} & \frac{5}{8} \end{bmatrix} \cdot \begin{bmatrix} 191 \\ 209 \\ 162 \end{bmatrix} = \begin{bmatrix} 35{,}125 \\ 32{,}5 \\ 20{,}625 \end{bmatrix}$$

Dados reais

Vejamos outro exemplo com um conjunto de dados “ clássico ”. Uma pessoa medida em diferentes lugares e momentos, a pressão atmosférica e a temperatura na qual a água entrou em fervura:

Presion :	20,79	20,79	22,4	22,67	23,15	23,35	23,89	23,99	24,02
Temperatura :	194,5	194,3	197,9	198,4	199,4	199,9	20,9	201,1	201,4

Presion :	24,01	25,14	26,57	28,49	27,76	29,04	29,88	30,06
Temperatura :	201,3	203,6	204,6	209,5	208,6	210,7	211,9	212,20



O relacionamento é “ mais ou menos linear

É tentador concluir que a relação entre ambas as variáveis é aproximadamente linear. Com isso, queremos dizer que as medidas foram feitas com um certo erro, e também que a temperatura de ebulição depende de outras variáveis, mas que seu efeito é muito menor que o da pressão atmosférica.

O relacionamento é " mais ou menos linear

É tentador concluir que a relação entre ambas as variáveis é aproximadamente linear. Com isso, queremos dizer que as medidas foram feitas com um certo erro, e também que a temperatura de ebulição depende de outras variáveis, mas que seu efeito é muito menor que o da pressão atmosférica.

" Todos os modelos são falsos, alguns são úteis. "

O relacionamento é “ mais ou menos linear

É tentador concluir que a relação entre ambas as variáveis é aproximadamente linear. Com isso, queremos dizer que as medidas foram feitas com um certo erro, e também que a temperatura de ebulição depende de outras variáveis, mas que seu efeito é muito menor que o da pressão atmosférica.

“ Todos os modelos são falsos, alguns são úteis. ”

Por isso, tentamos procurar a relação linear entre as duas variáveis que cometem o menor erro possível. É claro que precisamos de um termo independente:

$$T = c + mP$$

se o modelo fosse exato para dois números reais a e b, teríamos que

$$T_i = c + m P_i$$

para cada medição (P_i, T_i) de um valor de pressão e outra temperatura de ebulição. Escrevemos nosso sistema de equações para incógnitas a e b:

se o modelo fosse exato para dois números reais a e b, teríamos que

$$T_i = c \,+\, m\,P_i$$

para cada medição (P_i, T_i) de um valor de pressão e outra temperatura de ebulição. Escrevemos nosso sistema de equações para incógnitas a e b:

$$\begin{bmatrix} 1{,}0000000000000 & 20{,}7900000000000 \\ 1{,}0000000000000 & 20{,}7900000000000 \\ 1{,}0000000000000 & 22{,}4000000000000 \\ 1{,}0000000000000 & 22{,}6700000000000 \\ 1{,}0000000000000 & 23{,}1500000000000 \\ 1{,}0000000000000 & 23{,}3500000000000 \\ 1{,}0000000000000 & 23{,}8900000000000 \\ 1{,}0000000000000 & 23{,}9900000000000 \\ 1{,}0000000000000 & 24{,}0200000000000 \\ 1{,}0000000000000 & 24{,}0100000000000 \\ 1{,}0000000000000 & 25{,}1400000000000 \\ 1{,}0000000000000 & 26{,}5700000000000 \\ 1{,}0000000000000 & 28{,}4900000000000 \\ 1{,}0000000000000 & 27{,}7600000000000 \\ 1{,}0000000000000 & 29{,}0400000000000 \\ 1{,}0000000000000 & 29{,}8800000000000 \\ 1{,}0000000000000 & 30{,}0600000000000 \end{bmatrix} \cdot \begin{bmatrix} c \\ m \end{bmatrix} = \begin{bmatrix} 194{,}500000000000 \\ 194{,}300000000000 \\ 197{,}900000000000 \\ 198{,}400000000000 \\ 199{,}400000000000 \\ 199{,}900000000000 \\ 200{,}900000000000 \\ 201{,}100000000000 \\ 201{,}400000000000 \\ 201{,}300000000000 \\ 203{,}600000000000 \\ 204{,}600000000000 \\ 209{,}500000000000 \\ 208{,}600000000000 \\ 210{,}700000000000 \\ 211{,}900000000000 \\ 212{,}200000000000 \end{bmatrix}$$

se o modelo fosse exato para dois números reais a e b, teríamos que

$$T_i = c \;+\; m\, P_i$$

para cada medição (P_i, T_i) de um valor de pressão e outra temperatura de ebulição. Escrevemos nosso sistema de equações para incógnitas a e b:

$$\begin{bmatrix} 1{,}0000000000000 & 20{,}7900000000000 \\ 1{,}0000000000000 & 20{,}7900000000000 \\ 1{,}0000000000000 & 22{,}4000000000000 \\ 1{,}0000000000000 & 22{,}6700000000000 \\ 1{,}0000000000000 & 23{,}1500000000000 \\ 1{,}0000000000000 & 23{,}3500000000000 \\ 1{,}0000000000000 & 23{,}8900000000000 \\ 1{,}0000000000000 & 23{,}9900000000000 \\ 1{,}0000000000000 & 24{,}0200000000000 \\ 1{,}0000000000000 & 24{,}0100000000000 \\ 1{,}0000000000000 & 25{,}1400000000000 \\ 1{,}0000000000000 & 26{,}5700000000000 \\ 1{,}0000000000000 & 28{,}4900000000000 \\ 1{,}0000000000000 & 27{,}7600000000000 \\ 1{,}0000000000000 & 29{,}0400000000000 \\ 1{,}0000000000000 & 29{,}8800000000000 \\ 1{,}0000000000000 & 30{,}0600000000000 \end{bmatrix} \cdot \begin{bmatrix} c \\ m \end{bmatrix} = \begin{bmatrix} 194{,}500000000000 \\ 194{,}300000000000 \\ 197{,}900000000000 \\ 198{,}400000000000 \\ 199{,}400000000000 \\ 199{,}900000000000 \\ 200{,}900000000000 \\ 201{,}100000000000 \\ 201{,}400000000000 \\ 201{,}300000000000 \\ 203{,}600000000000 \\ 204{,}600000000000 \\ 209{,}500000000000 \\ 208{,}600000000000 \\ 210{,}700000000000 \\ 211{,}900000000000 \\ 212{,}200000000000 \end{bmatrix}$$

Isso é impossível, e procuramos a abordagem de quadrados mínimos.

Fazemos os cálculos necessários:

$$A^T A = \begin{bmatrix} 17 & 426 \\ 426 & 10820{,}9966 \end{bmatrix}$$

$$(A^T A)^{-1} = \begin{bmatrix} 4{,}36164800614875 & -0{,}171708957991852 \\ -0{,}171708957991852 & 0{,}00685223541282037 \end{bmatrix}$$

$$A^T \mathbf{b} = \begin{bmatrix} 3450{,}2 \\ 86735{,}495 \end{bmatrix}$$

$$\mathbf{x} = \begin{bmatrix} c \\ m \end{bmatrix} = (A^T A)^{-1}(A^T \mathbf{b}) = \begin{bmatrix} 155{,}296483456967 \\ 1{,}90178352401767 \end{bmatrix}$$

Fazemos os cálculos necessários:

$$A^T A = \begin{bmatrix} 17 & 426 \\ 426 & 10820{,}9966 \end{bmatrix}$$

$$(A^T A)^{-1} = \begin{bmatrix} 4{,}36164800614875 & -0{,}171708957991852 \\ -0{,}171708957991852 & 0{,}00685223541282037 \end{bmatrix}$$

$$A^T\mathbf{b} = \begin{bmatrix} 3450{,}2 \\ 86735{,}495 \end{bmatrix}$$

$$\mathbf{x} = \begin{bmatrix} c \\ m \end{bmatrix} = (A^T A)^{-1}(A^T\mathbf{b}) = \begin{bmatrix} 155{,}296483456967 \\ 1{,}90178352401767 \end{bmatrix}$$

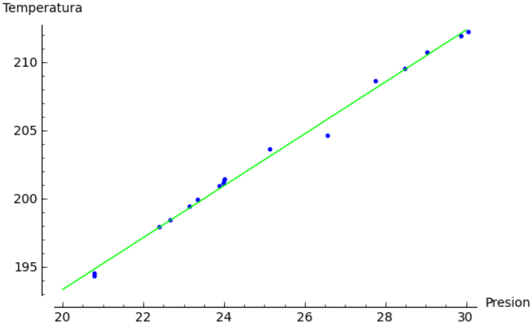
Isto é, isso

$$T = 155{,}296483456967 + 1{,}90178352401767P$$

É a melhor abordagem linear para o sistema.Em outras palavras, é a relação linear entre as duas variáveis que cometem menos erro.

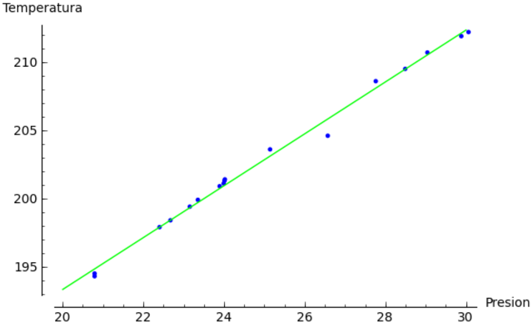
Encontramos a relação linear entre as duas variáveis que cometem um erro menor:

$$T = 155{,}296483456967 + 1{,}90178352401767P$$



Encontramos a relação linear entre as duas variáveis que cometem um erro menor:

$$T = 155{,}296483456967 + 1{,}90178352401767P$$



Esta não é uma lei universal e claramente falha para pressões muito baixas. É apenas uma lei *aproximadamente certa em uma certa gama de pressão atmosférica.*