# Método dos Mínimos Quadrados INF1608 – Análise Numérica

Departamento de Informática, PUC-Rio





#### Sistemas inconsistentes

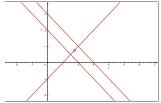
Considere o sistema de equações

$$\begin{cases} x_0 + x_1 = 2 \\ x_0 - x_1 = 1 \\ x_0 + x_1 = 3 \end{cases}$$

Primeira e terceira equações são contraditórias!

#### Objetivo

Achar solução que melhor aproxime todas as equações







- Sistemas lineares inconsistentes
  - Número de equações é maior que o número de incógnitas
  - ▶ Pode-se achar a "melhor" solução aproximada

$$A_{m \times n} \mathbf{x}_n = \mathbf{b}_m, \quad n < m$$

► Se "melhor" for menor distância Euclidiana ⇒ MMQ





Reescrevendo o exemplo em forma matricial:

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

Alternativamente:

$$x_0 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + x_1 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

▶ Qualquer sistema  $m \times n$  na forma  $A\mathbf{x} = \mathbf{b}$  pode ser expresso por uma equação vetorial na forma:

$$x_0 \mathbf{v}_0 + x_1 \mathbf{v}_1 + ... + x_{n-1} \mathbf{v}_{n-1} = \mathbf{b}$$

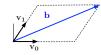
isto é, **b** é expresso como uma combinação linear de *n* vetores  $\mathbf{v}_i$ 's com coeficientes  $x_i$ 's, no espaço  $\mathbb{R}^m$ 





No nosso exemplo: Combinação linear de 2 vetores no espaço 3D

- ► Combinação linear de dois vetores formam um plano
- Qualquer vetor neste plano pode ser representado



- ► Plano formado pelos vetores: Ax
- Solução exata só existe se b estiver neste plano



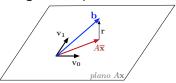


No nosso exemplo: Combinação linear de 2 vetores no espaço 3D

- Combinação linear de dois vetores formam um plano
- Qualquer vetor neste plano pode ser representado



- ► Plano formado pelos vetores: Ax
- ► Solução exata só existe se **b** estiver neste plano
- ▶ Melhor solução aproxima b por sua projeção no plano: Ax̄
- ▶ Vetor residual ortogonal ao plano:  $\mathbf{r} = \mathbf{b} A\overline{\mathbf{x}}$







Perpendicularidade do vetor residual:

$$(\mathbf{b} - A\overline{\mathbf{x}}) \perp A\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n$$





Perpendicularidade do vetor residual:

$$(\mathbf{b} - A\overline{\mathbf{x}}) \perp A\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n$$

Logo:

$$(A\mathbf{x})^{\mathsf{T}}(b - A\overline{\mathbf{x}}) = 0 \ \forall \ \mathbf{x} \in \mathbb{R}^{n}$$
  
 $\mathbf{x}^{\mathsf{T}}A^{\mathsf{T}}(b - A\overline{\mathbf{x}}) = 0 \ \forall \ \mathbf{x} \in \mathbb{R}^{n}$ 

Então:

$$\mathbf{x}^T \perp A^T (b - A\overline{\mathbf{x}})$$

▶ para qualquer  $\mathbf{x} \in \mathbb{R}^n$ , inclusive o próprio  $A^T(b - A\overline{\mathbf{x}})$ 





Perpendicularidade do vetor residual:

$$(\mathbf{b} - A\overline{\mathbf{x}}) \perp A\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n$$

Logo:

$$(A\mathbf{x})^T(b - A\overline{\mathbf{x}}) = 0 \ \forall \ \mathbf{x} \in \mathbb{R}^n$$
  
 $\mathbf{x}^T A^T(b - A\overline{\mathbf{x}}) = 0 \ \forall \ \mathbf{x} \in \mathbb{R}^n$ 

Então:

$$\mathbf{x}^T \perp A^T (b - A\overline{\mathbf{x}})$$

▶ para qualquer  $\mathbf{x} \in \mathbb{R}^n$ , inclusive o próprio  $A^T(b - A\overline{\mathbf{x}})$ 

Isso é verdade apenas se o vetor  $A^{T}(b - A\overline{\mathbf{x}})$  for nulo:

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

$$\mathbf{A}^{\mathsf{T}}\mathbf{A}\overline{\mathbf{x}}=\mathbf{A}^{\mathsf{T}}\mathbf{b}$$





Dado um sistema incosistente:

$$Ax = b$$

MMQ pode ser resolvido pelo sistema de equações normais:

$$(A^TA) \bar{\mathbf{x}} = A^T\mathbf{b}$$

- que minimiza o vetor residual  $\mathbf{r} = \mathbf{b} A\overline{\mathbf{x}}$
- ► Note que A<sup>T</sup>A é sempre simétrica

$$(A^TA)^T = A^T(A^T)^T = A^TA$$





Retomando nosso exemplo:

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

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

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





Ficamos com o sistema linear:

$$\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} \overline{x_0} \\ \overline{x_1} \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

► Solução:

$$\left[\begin{array}{c} \overline{x_0} \\ \overline{x_1} \end{array}\right] = \left[\begin{array}{c} \frac{7}{4} \\ \frac{3}{4} \end{array}\right]$$

Substituindo no sistema original:

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{7}{4} \\ \frac{3}{4} \\ \frac{10}{4} \end{bmatrix} = \begin{bmatrix} \frac{10}{4} \\ \frac{4}{4} \\ \frac{10}{4} \\ 2.5 \end{bmatrix} \neq \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$





#### Resíduo

$$\mathbf{r} = \mathbf{b} - A\overline{\mathbf{x}} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} 2.5 \\ 1 \\ 2.5 \end{bmatrix} = \begin{bmatrix} -0.5 \\ 0.0 \\ 0.5 \end{bmatrix}$$

- Erro regressivo: MMQ minimiza este erro
  - ► Pela norma-2 do vetor r

$$\|\mathbf{r}\|_2 = \sqrt{r_0^2 + r_1^2 + ... r_{n-1}^2}$$

- Pela raiz da média do erro ao guadrado
  - ► RMSE (root mean squared error)

$$RMSE = \frac{\|\mathbf{r}\|_2}{\sqrt{n}}$$

- Se o erro for nulo, encontra-se a solução do problema original
  - Uma solução existia
  - O vetor **b** estava no plano A**x**





# Ajuste de modelos

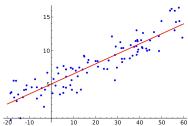
#### Regressão Linear

Dado um conjunto de pontos  $(x_i, y_i)$ , achar a reta que minimiza o erro

Achar a e b em:

$$y_i = a + bx_i + e_i$$

▶ tal que  $\sum e_i^2$  seja mínimo







### Regressão Linear

Achar mínimo da função:

$$S(a,b) = \sum (y_i - a - bx_i)^2$$

Derivando e igualando a zero (mínimo de função):

$$\frac{\partial S}{\partial a} = -2\sum (y_i - a - bx_i) = 0$$
$$\frac{\partial S}{\partial b} = -2\sum x_i (y_i - a - bx_i) = 0$$





### Regressão Linear

▶ Dividindo a primeira equação por 2n:

$$-\frac{1}{n}\sum_{i}(y_{i}-a-bx_{i})=-\frac{1}{n}\left(\sum_{i}y_{i}-\sum_{i}a-\sum_{i}bx_{i}\right)=$$

$$=-\frac{1}{n}(n\overline{y}-na-nb\overline{x})=-\overline{y}+a+b\overline{x}=0$$

$$\therefore a=\overline{y}-b\overline{x}$$





## Regressão Linear

▶ Dividindo a primeira equação por 2n:

$$-\frac{1}{n}\sum_{i}(y_{i}-a-bx_{i})=-\frac{1}{n}\left(\sum_{i}y_{i}-\sum_{i}a-\sum_{i}bx_{i}\right)=$$

$$=-\frac{1}{n}\left(n\overline{y}-na-nb\overline{x}\right)=-\overline{y}+a+b\overline{x}=0$$

$$\therefore \boxed{a=\overline{y}-b\overline{x}}$$

Substituindo na segunda equação:

 $b = \frac{\sum x_i y_i - n \overline{x} \overline{y}}{\sum x_i^2 - n \overline{x}^2}$ 

$$-2\sum_{i}x_{i}(y_{i}-\overline{y}+b\overline{x}-bx_{i}) = \sum_{i}(x_{i}(y_{i}-\overline{y})+x_{i}b(\overline{x}-x_{i})) =$$

$$=\sum_{i}x_{i}y_{i}-\overline{y}\sum_{i}x_{i}+b(\overline{x}\sum_{i}x_{i}-\sum_{i}x_{i}^{2})=0$$

$$\sum_{i}x_{i}y_{i}-n\overline{x}\overline{y}=b(\sum_{i}x_{i}^{2}-n\overline{x}^{2}) // b \text{ troca de sinal}$$





#### Ajuste linear por MMQ

$$A\mathbf{x} = \mathbf{b} \Rightarrow A^T A \overline{\mathbf{x}} = A^T \mathbf{b}$$

Sistema inconsistente:  $y_i = a + bx_i$ 

$$\begin{bmatrix} 1 & x_0 \\ 1 & x_1 \\ \dots & \dots \\ 1 & x_{n-1} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} y_0 \\ y_1 \\ \dots \\ y_{n-1} \end{bmatrix}$$

Solução

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ x_0 & x_1 & \dots & x_{n-1} \end{bmatrix} \begin{bmatrix} 1 & x_0 \\ 1 & x_1 \\ \dots & \dots \\ 1 & x_{n-1} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_0 & x_1 & \dots & x_{n-1} \end{bmatrix} \begin{bmatrix} y_0 \\ y_1 \\ \dots \\ y_{n-1} \end{bmatrix}$$





#### Ajuste linear por MMQ

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ x_0 & x_1 & \dots & x_{n-1} \end{bmatrix} \begin{bmatrix} 1 & x_0 \\ 1 & x_1 \\ \dots & \dots \\ 1 & x_{n-1} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_0 & x_1 & \dots & x_{n-1} \end{bmatrix} \begin{bmatrix} y_0 \\ y_1 \\ \dots \\ y_{n-1} \end{bmatrix}$$

$$\begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix}$$





#### Ajuste linear por MMQ

$$\begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix}$$





#### Ajuste linear por MMQ

$$\begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix} \therefore \begin{bmatrix} n & n\overline{x} \\ n\overline{x} & \sum x_i^2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} n\overline{y} \\ \sum x_i y_i \end{bmatrix}$$





#### Ajuste linear por MMQ

$$\begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix} : \begin{bmatrix} n & n\overline{x} \\ n\overline{x} & \sum x_i^2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} n\overline{y} \\ \sum x_i y_i \end{bmatrix}$$

Da primeira equação:

$$na + n\overline{x}b = n\overline{y}$$
 $a = \overline{y} - \overline{x}b$ 

Da segunda equação:

$$n\overline{x}a + b\sum_{i} x_{i}^{2} = \sum_{i} x_{i}y_{i}$$

$$n\overline{x}(\overline{y} - \overline{x}b) + b\sum_{i} x_{i}^{2} = \sum_{i} x_{i}y_{i}$$

$$n\overline{x}\overline{y} - n\overline{x}^{2}b + b\sum_{i} x_{i}^{2} = \sum_{i} x_{i}y_{i}$$

$$b = \frac{\sum_{i} x_{i}y_{i} - n\overline{x}\overline{y}}{\sum_{i} x_{i}^{2} - n\overline{x}^{2}}$$





#### Ajuste linear por MMQ

$$\begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix} : \begin{bmatrix} n & n\overline{x} \\ n\overline{x} & \sum x_i^2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} n\overline{y} \\ \sum x_i y_i \end{bmatrix}$$

▶ Da primeira equação:

$$na + n\overline{x}b = n\overline{y}$$

$$a = \overline{y} - \overline{x}b$$

Da segunda equação:

$$n\overline{x}a + b\sum_{i} x_{i}^{2} = \sum_{i} x_{i}y_{i}$$

$$n\overline{x}(\overline{y} - \overline{x}b) + b\sum_{i} x_{i}^{2} = \sum_{i} x_{i}y_{i}$$

$$n\overline{x}\overline{y} - n\overline{x}^{2}b + b\sum_{i} x_{i}^{2} = \sum_{i} x_{i}y_{i}$$

$$b = \frac{\sum_{i} x_{i}y_{i} - n\overline{x}\overline{y}}{\sum_{i} x_{i}^{2} - n\overline{x}^{2}}$$

Que são as mesmas fórmulas da Regressão Linear





#### Aplicações

- ► Ajustar um modelo simples a um conjunto de pontos
  - ► Ao invés de ajustar um polinômio que honre todos os pontos
- Ajustar a um conjunto de pontos com erros de aquisição

#### Uso

- 1. Definir um modelo, como y = a + bx
- 2. Forçar o ajuste do modelo a partir de um sistema Ax = b
  - Nesse caso, vetor x = [a, b]
- 3. Resolver o sistema para encontrar o vetor solução aproximado
  - Que pode ser uma solução exata





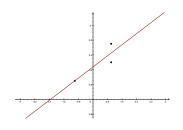
# Ajuste linear

#### Exemplo:

- Achar melhor reta que aproxima um conjunto de pontos  $(x_i, y_i)$ .
- ▶ Modelo (linha reta): y = a + bx
- ► Exemplo: (1,2), (-1,1) e (1,3)
  - Sistema inconsistente

$$a+1b=2$$
$$a+(-1)b=1$$
$$a+1b=3$$

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$



► Como vimos:

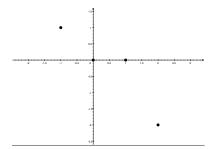
$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \frac{7}{4} \\ \frac{3}{4} \end{bmatrix}$$
$$\Rightarrow y = \frac{7}{4} + \frac{3}{4}x$$





Considere os pontos (-1,1), (0,0), (1,0), (2,-2)

▶ Ajuste de uma reta: y = a + bx



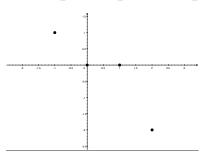


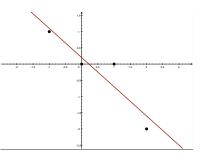


Considere os pontos (-1,1),(0,0),(1,0),(2,-2)

▶ Ajuste de uma reta: y = a + bx

$$\begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -2 \end{bmatrix} \Rightarrow \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0.2 \\ -0.9 \end{bmatrix}$$



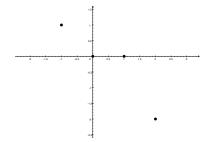






Considere os pontos (-1,1),(0,0),(1,0),(2,-2)

▶ Ajuste de uma parábola:  $y = a + bx + cx^2$ 



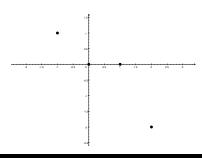


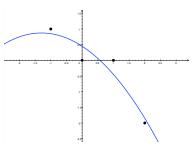


Considere os pontos (-1,1), (0,0), (1,0), (2,-2)

▶ Ajuste de uma parábola:  $y = a + bx + cx^2$ 

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -2 \end{bmatrix} \Rightarrow \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0.45 \\ -0.65 \\ -0.25 \end{bmatrix}$$









#### Estudo de modelos

- Dados periódicos
  - Exemplo: variação de temperatura durante o dia
    - ► Washington, DC (01/01/2001)

h	t	Т
0	0	-2.2
3	1/8	-2.8
6	1/4	-6.1
9	3/8	-3.9
12	1/2	-0.0
15	5/8	1.1
18	3/4	-0.6
21	7/8	-1.1

#### Estudo de modelos

- Dados periódicos
  - Exemplo: variação de temperatura durante o dia
    - ► Washington, DC (01/01/2001)

h	t	Т
0	0	-2.2
3	1/8	-2.8
6	1/4	-6.1
9	3/8	-3.9
12	1/2	-0.0
15	5/8	1.1
18	3/4	-0.6
21	7/8	-1.1

#### Modelo:

$$y = c_0 + c_1 \cos 2\pi t + c_2 \sin 2\pi t$$



#### Estudo de modelos

- Dados periódicos
  - Exemplo: variação de temperatura durante o dia
    - ► Washington, DC (01/01/2001)

h	t	Т
0	0	-2.2
3	1/8	-2.8
6	1/4	-6.1
9	3/8	-3.9
12	1/2	-0.0
15	5/8	1.1
18	3/4	-0.6
21	7/8	$\mid -1.1 \mid$

#### Modelo:

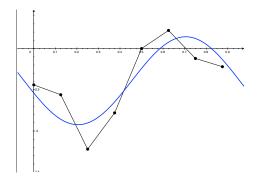
$$y = c_0 + c_1 \cos 2\pi t + c_2 \sin 2\pi t$$

#### Sistema inconsistente

$$\begin{bmatrix} 1 & \cos 2\pi t_0 & \sin 2\pi t_0 \\ 1 & \cos 2\pi t_1 & \sin 2\pi t_1 \\ \vdots & \vdots & & \vdots \\ 1 & \cos 2\pi t_{n-1} & \sin 2\pi t_{n-1} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} T_0 \\ T_1 \\ \vdots \\ T_{n-1} \end{bmatrix}$$

#### Dados periódicos

- Exemplo: variação de temperatura durante o dia
  - ► Modelo 1:  $y = c_0 + c_1 \cos 2\pi t + c_2 \sin 2\pi t$ 
    - Anti-simetria a.m. e p.m.

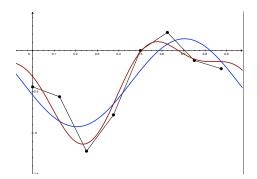






#### Dados periódicos

- Exemplo: variação de temperatura durante o dia
  - ► Modelo 1:  $y = c_0 + c_1 \cos 2\pi t + c_2 \sin 2\pi t$ 
    - Anti-simetria a.m. e p.m.
  - ► Modelo 2:  $y = c_0 + c_1 \cos 2\pi t + c_2 \sin 2\pi t + c_3 \cos 4\pi t$ 
    - Modelo melhorado







Como usar MMQ para problemas não lineares?

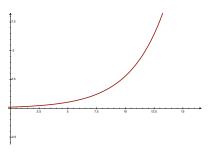
#### Exemplo

► Crescimento populacional exponencial

#### Modelo exponencial:

$$y = a e^{bx}$$

 Não pode ser resolvido diretamente pelo MMQ







#### Modelos não lineares





#### Modelo exponencial

$$y = a e^{bx}$$





#### Modelo exponencial

$$y = a e^{bx}$$

Linearizando:

$$\ln y = \ln \left( a \ e^{b x} \right)$$

$$\ln y = \ln a + b x$$

$$\ln y = k + b x, \ a = e^k$$





#### Modelo exponencial

$$y = a e^{bx}$$

Linearizando:

$$\ln y = \ln \left( a \ e^{b x} \right)$$

$$\ln y = \ln a + b x$$

$$\ln y = k + bx, \ a = e^k$$

- ► Note que estamos resolvendo um problema diferente
  - Ao invés de minimizar o erro:

$$\sum \left(ae^{bx_i}-y_i\right)^2$$

Estamos minimizando o erro no espaço logarítmico:

$$\sum \left(\ln a + bx_i - \ln y_i\right)^2$$





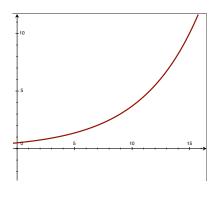
Modelo de potência simples

$$y = a x^b$$

Linearizando:

$$\log y = \log \left( a \, x^b \right)$$
$$\log y = \log a + b \, \log x$$

$$\log y = k + b \log x, \ a = 10^k$$







#### Modelo de crescimento com saturação

$$y = a \frac{x}{x+b}$$

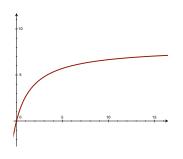
#### Linearizando:

$$\frac{1}{y} = \frac{x+b}{ax}$$

$$\frac{1}{v} = \frac{1}{a} + \frac{b}{a} \frac{1}{x}$$

$$\frac{1}{y} = k_1 + k_2 \frac{1}{x}$$

$$\overline{\mathsf{com}\left[\mathsf{a} = \frac{1}{\mathsf{k}_1}\right]} \,\mathsf{e}\left[\mathsf{b} = \mathsf{k}_2\,\mathsf{a}\right]$$







# Exercício proposto

1. Usando Mínimos Quadrados, como fazer o ajuste de um conjunto de pontos  $(x_i, y_i)$  num modelo de potência simples:  $y = ax^b$ . Qual o erro que será minimizado?



