Mini curso

Método dos Mínimos Quadrados

com ênfase em covariâncias

Otaviano Helene

Instituto de Física da USP

Baseado no livro *Método dos Mínimos Quadrados com Formalismo Matricial* e no blog https://livrommq.blogspot.com/

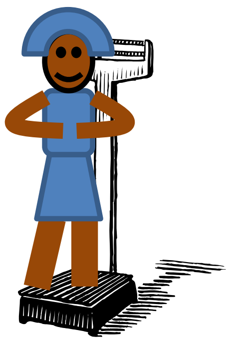
Semana Integrada da Física de São Carlos

IFSC – USP Agosto 2019

Conteúdo

- **I – Solução do MMQ para um problema sem solução**
- **II – O desvio padrão e alguns exemplos de ajustes de parâmetros**
- **III – Variâncias e covariâncias; propagação**
- **IV a – Equações do MMQ com matrizes**
- **IV b – Variâncias e covariâncias dos parâmetros ajustados**
- **V – Algumas propriedades do MMQ e o TCL**
- **VI – Outros desenvolvimentos**

Problema sem solução



$m_0 \rightarrow$ 60 kg $\quad m_0+f_0 \rightarrow$ 65 kg $\quad f_0 \rightarrow$ 4 kg

m=60

m+f=65 SISTEMA INCONSISTENTE

f=4

Quanto pesam criança e mãe?

I – Solução do MMQ

- Procurar valores de *m* e *f* que minimizem

$$Q(m,f) = (m-60)^2 + (m+f-65)^2 + (f-4)^2$$

Valores de *Q* para diferentes valores de *m* e *f*

$$Q(m,f)=(m-60)^2+(m+f-65)^2+(f-4)^2$$

	Mãe, *m* (kg)					
Filho, *f* (kg)		58	59	60	61	62
	3,5	16,5	7,5	2,5	1,5	4,5
	4,0	13,0	5,0	1,0	1,0	5,0
	4,5	10,5	3,5	**0,5**	1,5	6,5
	5,0	9,0	3,0	1,0	3,0	9,0
	5,5	8,5	3,5	2,5	5,5	12,5

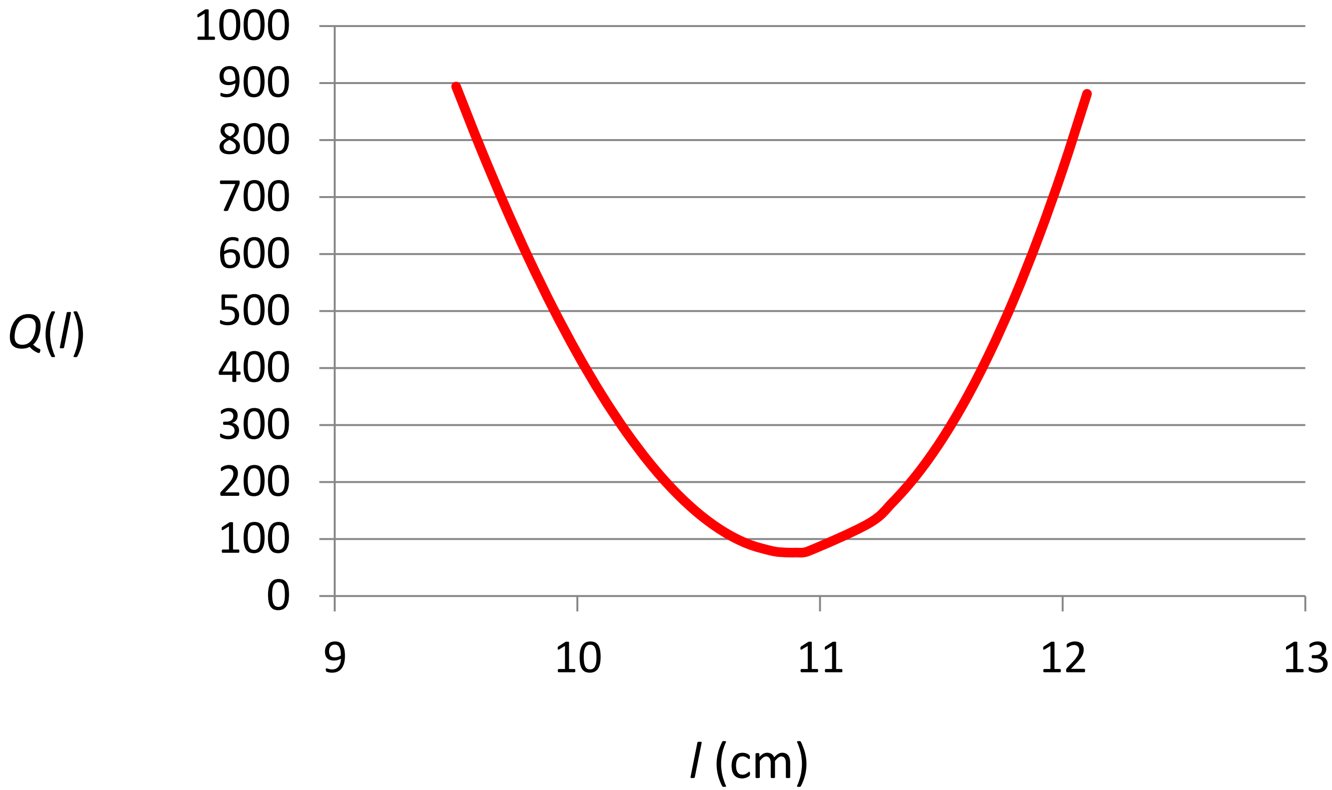
Medidas de um quadrado

- Lado = 10 cm
- Área = 120 cm^2
- Perímetro = 35 cm

Valores evidentemente inconsistentes

Que valor adotar para o lado desse quadrado?

$$Q(l) = (l-10)^2 + (l^2-120)^2 + (4l-35)^2$$



O “melhor” valor para *l* é 10,9 cm

Solução analítica

- Quando os **dados dependem linearmente dos parâmetros** a serem ajustados, há uma solução analítica para os valores que minimizam *Q*. Exemplo
  - $m \rightarrow 60$
  - $m+f \rightarrow 65$
  - $f \rightarrow 4$

$$Q(m, f) = (m - 60)^2 + (m + f - 65)^2 + (f - 4)^2$$

$$Q(m, f) = (m - 60)^2 + (m + f - 65)^2 + (f - 4)^2$$

$$\left.\frac{\partial Q}{\partial m}\right|_{\tilde{m},\, \tilde{f}} = 0$$

$$\left.\frac{\partial Q}{\partial f}\right|_{\tilde{m},\, \tilde{f}} = 0$$

Os valores ajustados têm um til

$$2(\tilde{m} - 60) + 2(\tilde{m} + \tilde{f} - 65) = 0$$

$$2(\tilde{m} + f - 65) + 2(\tilde{f} - 4) = 0$$

$$2\tilde{m} + \tilde{f} = 125$$

$$\tilde{m} + 2f = 69$$

Escrevendo na forma de matrizes

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} \tilde{m} \\ \tilde{f} \end{pmatrix} = \begin{pmatrix} 125 \\ 69 \end{pmatrix}$$

$$\begin{pmatrix} \tilde{m} \\ \tilde{f} \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 125 \\ 69 \end{pmatrix} = \begin{pmatrix} 60{,}3 \\ 4{,}3 \end{pmatrix}$$

A menos de arredondamento, são os mesmo valores obtidos numericamente (60kg e 4,5kg)

- Vamos nos restringir ao caso linear.

- O caso não linear é uma aproximação, e algumas propriedades do MMQ não são válidas.

Exercício para resolver

- Determinar o rendimento (em litros por quilômetro) de um veículo

Distâncias (km)		Consumo (l)
A 80 km/h	A 120 km/h	
200	100	29
100	200	31
500	100	49

R: 0,075 l/km a 80 km/h e 0,12 l/km a 120 km/h

Sobre o MMQ

1) O MMQ é mais geral do que se pensa.

Não serve apenas para "ajustar funções" – ajusta parâmetros

Os dados não precisam obedecer a distribuições normais (Verbete *Método dos mínimos quadrados* da Wikipédia em português está errado – jul/2019)

2) O MMQ tem limitações que nem sempre são consideradas (p. ex., nos casos em que há erro na "variável dependente" ou a relação entre dados e parâmetros não é linear).

II – O desvio padrão e alguns exemplos de ajustes de parâmetros

Alguns dados são mais precisos do que outros. Exemplos:

- combinar medidas com instrumentos diferentes
- medidas de uma mesma grandeza com técnicas diferentes
- medidas feitas por pessoas com habilidades diferentes

Aprendendo o que fazer a partir de medidas equivalentes

- x_1, x_2 e x_3 três medidas equivalentes (mesmos procedimentos, equipamentos, experimentadores, condições externas ...).

- O melhor resultado (veremos em breve o que "melhor" quer dizer) é uma média simples:

$$\bar{x} = \frac{x_1 + x_2 + x_3}{3}$$

Suponha, agora, que x_2 e x_3 tenham sido combinados,

$$\bar{x}_{2,3} = \frac{x_2 + x_3}{2}$$

e não saibamos mais seus valores, apenas o valor médio $x_{2,3}$. Como combiná-lo com x_1?

Como $x_{2,3}$ foi calculado usando dois dados, é razoável supor que ele tenha peso 2. Assim, a média deve ser

$$\bar{x}_{1;2,3} = \frac{x_1 + 2\bar{x}_{2,3}}{3} = \frac{x_1 + x_2 + x_3}{3} = \bar{x}$$

O peso deve refletir a precisão de um dado

- O peso é proporcional ao inverso da variância: $1/\sigma^2$.
- σ^2 é a variância e σ é o desvio padrão.
- σ é uma medida de quanto um valor flutua em torno do valor verdadeiro (e desconhecido) da grandeza medida.

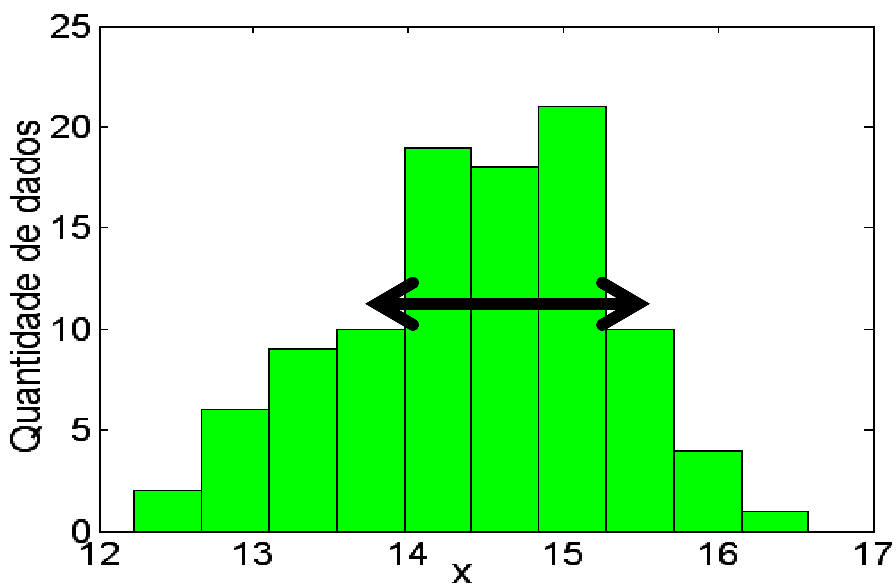
- Exemplo: $\sigma_{régua} \approx 0{,}3$ mm; $\sigma_{paquímetro} \approx 0{,}05$ mm

Como saber o valor do desvio padrão σ? (Vamos usar a palavra *erro* apenas para a diferença entre o valor experimental e o valor verdadeiro – desconhecido, claro.)

- O fabricante de um equipamento pode informar
- Fazendo várias medidas de uma mesma grandeza
- Estudando as propriedades físicas do processo
- Conhecimento anterior

Estimando o d.p.

I – Uma forma usual de estimar o d.p. é a partir de medidas independentes de uma mesma grandeza, p. ex., x_1, x_2, … x_n



$$
{}^{2} \approx \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2
$$

II – O **desvio padrão da média** de *n* dados independentes é

$$\sigma_{média} = \frac{\sigma}{\sqrt{n}}$$

III – Eventos contáveis, como o número de decaimentos de uma fonte radioativa (distribuição de Poisson). Se *N* eventos forem observados, $\sigma^2 \approx N$.

No que segue:

- Vamos nos restringir ao caso em que os parâmetros a serem ajustados dependem linearmente dos dados.

- Vamos supor conhecidos os desvios padrões.

- Abrir mão dessas limitações têm consequências que veremos mais adiante.

Dois exemplos

- 1 – Consumo de combustível: g_c e g_e (litros/km)

distância percorrida (km)		consumo (l)
cidade	estrada	
100	200	28
300	100	43
400	0	49

a equação do MMQ

$$Q(g_c, g_e) = (100g_c + 200g_e - 28)^2$$
$$+ (300g_c + 100g_e - 43)^2 + (400g_c - 49)^2$$

Derivando em relação a g_c e g_e, igualando a zero

$$2600\tilde{g}_c + 500\tilde{g}_e = 353$$
$$500\tilde{g}_c + 500\tilde{g}_e = 99$$

Solução:
$$\tilde{g}_c = 0{,}12\, l/km$$
$$\tilde{g}_e = 0{,}077\, l/km$$

- 2 – Ajuste dos parâmetros da função $y=a\cdot x^2+b\cdot x^3$ aos dados abaixo

x	y
-1	3,9
0	-1,6
1	1,4
3	-15

$$Q(a,b) = \sum \left(y_i - a x_i^2 - b x_i^3 \right)^2$$

$$Q(a,b) = \sum \left(y_i - a x_i^2 - b x_i^3\right)^2$$

- Derivando em relação à *a* e igualando a zero

$$\frac{\partial Q}{\partial a} = -2 \sum \left(y_i - \tilde{a} {x_i}^2 - \tilde{b} {x_i}^3\right) {x_i}^2 = 0$$

- Rearranjando os termos

$$\tilde{a} \sum {x_i}^2 {x_i}^2 + \tilde{b} \sum {x_i}^2 {x_i}^3 = \sum y_i {x_i}^2$$

- Derivando em relação a *a* e *b* ...

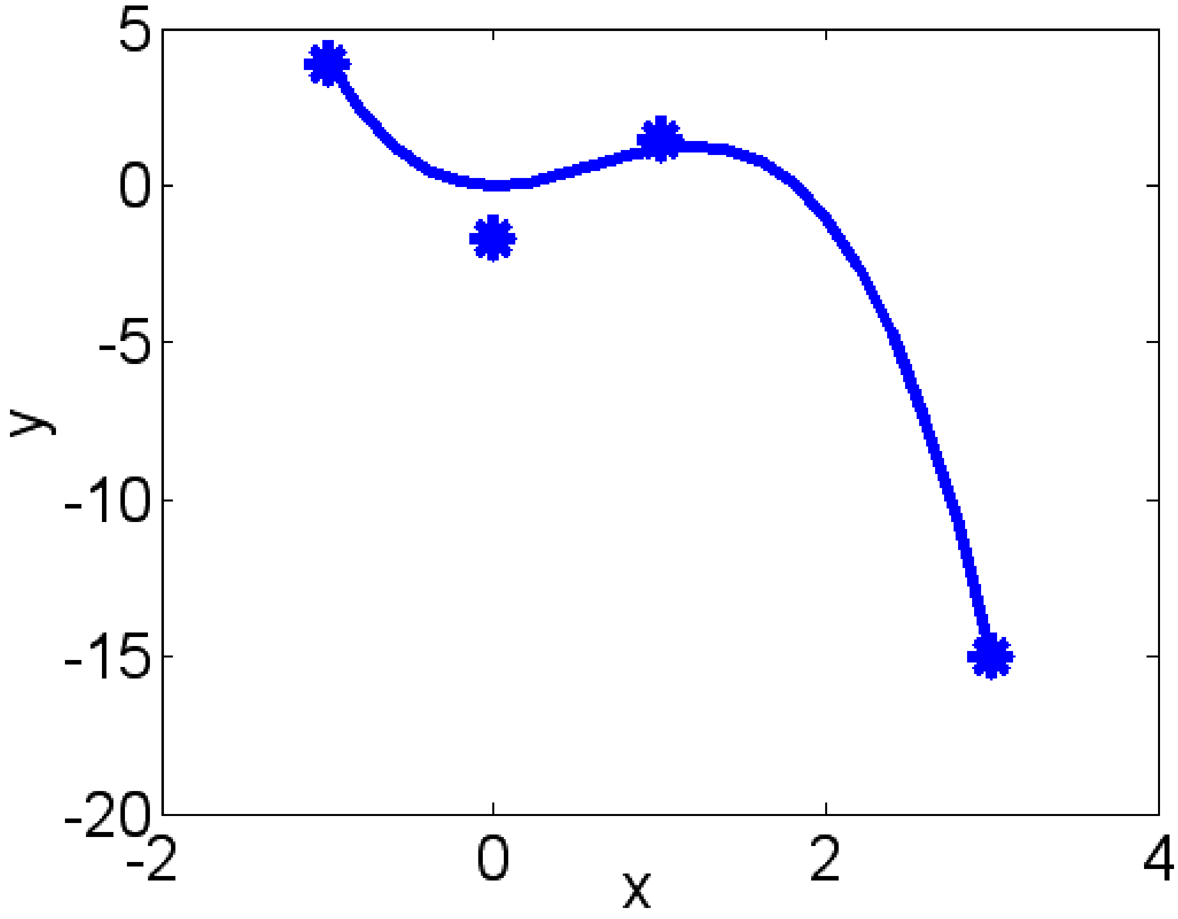
$$\tilde{a}\sum x_i^2 x_i^2 + \tilde{b}\sum x_i^2 x_i^3 = \sum x_i^2 y_i$$

$$\tilde{a}\sum x_i^3 x_i^2 + \tilde{b}\sum x_i^3 x_i^3 = \sum x_i^3 y_i$$

Solução

$$\tilde{a} = 2{,}6$$

$$\tilde{b} = -1{,}4$$


5
0
-5
-10
-15
-20
y
-2
0
2
4
x

$$\tilde{a}\sum x_i^2 x_i^2 + \tilde{b}\sum x_i^2 x_i^3 = \sum x_i^2 y_i$$
$$\tilde{a}\sum x_i^3 x_i^2 + \tilde{b}\sum x_i^3 x_i^3 = \sum x_i^3 y_i$$

- Essa forma estranha de escrever é pelo seguinte

A função é $y=a \cdot x^2+b \cdot x^3$

$$\tilde{a}\sum x_i^2 x_i^2 + \tilde{b}\sum x_i^2 x_i^3 = \sum x_i^2 y_i$$
$$\tilde{a}\sum x_i^3 x_i^2 + \tilde{b}\sum x_i^3 x_i^3 = \sum x_i^3 y_i$$

Isso facilita as coisas

$$y = a_0 \cdot f + b_0 \cdot g + c_0 \cdot h + \ldots,$$

f, g, h...são valores conhecidos e a_0, b_0, c_0 ... parâmetros a serem ajustados.

$$Q(a,b,c,\ldots) = \sum_{i=1}^{n} \frac{\left(y_i - af_i - bg_i - ch_i \ldots\right)^2}{\sigma_i^2}$$

$$\begin{pmatrix} \sum \frac{f_i y_i}{\sigma_i^2} \\ \sum \frac{g_i y_i}{\sigma_i^2} \\ \sum \frac{h_i y_i}{\sigma_i^2} \\ \vdots \end{pmatrix} = \begin{pmatrix} \sum \frac{f_i^2}{\sigma_i^2} & \sum \frac{f_i g_i}{\sigma_i^2} & \sum \frac{f_i h_i}{\sigma_i^2} & \cdots \\ \sum \frac{f_i g_i}{\sigma_i^2} & \sum \frac{g_i^2}{\sigma_i^2} & \sum \frac{g_i h_i}{\sigma_i^2} & \\ \sum \frac{f_i h_i}{\sigma_i^2} & \sum \frac{g_i h_i}{\sigma_i^2} & \sum \frac{h_i^2}{\sigma_i^2} & \\ \vdots & & & \ddots \end{pmatrix} \cdot \begin{pmatrix} \tilde{a} \\ \tilde{b} \\ \tilde{c} \\ \vdots \end{pmatrix}$$

Um cuidado importante
Tendenciosidade e não tendenciosidade
Exemplo

- Suponha que o tempo verdadeiro a ser medido seja de t
- Suponha que haja um medidor de tempo não tendencioso que meça os valores abaixo com as probabilidades indicadas:

resultado	probabilidade
t-5s	25%
T	50%
t+5s	25%

Não sabemos o valor de t, mas sabemos que, em média, o medidor fornece o valor verdadeiro t, pois o valor esperado da medida é

$$(t\text{-}5s)\times 0{,}25+t\times 0{,}50+(t+5s)\times 0{,}25=t$$

que é o valor verdadeiro (e desconhecido) do tempo.

Note que se o tempo medido não for muito maior do que 5 s, é possível que alguns valores sejam negativos. Preserve-os

No entanto, esse equipamento é tendencioso para o inverso do tempo, pois

$$(1/(t\text{-}1s))\times 0{,}25+(1/t)\times 0{,}50+(1/(t+1s))\times 0{,}25\neq 1/t$$

- **Outro exemplo de tendenciosidade – medida da massa de um liquido em um recipiente**

Dez medidas não tendenciosas das massas do líquido no recipiente e do recipiente vazio.

As diferenças são os dados correspondentes ao líquido apenas

L+R (g)	R (g)	Diferenças (g)
2,56	1,41	1,15
1,22	2,96	-1,74
3,51	1,91	1,60
3,9	1,74	2,16
4,09	1,84	2,25
2,38	1,84	0,54
2,27	0,67	1,60
3,78	2,25	1,53
2,58	3,05	-0,47
3,99	0,76	3,23

- Se as medidas de massa são não tendenciosas, então as diferenças também não o são.
- A média das diferenças também não é tendenciosa. Essa média é 1,19 g.

- **Entretanto, se os dados "não físicos" (massas negativas) fossem descartados, a média seria tendenciosa**
- **Os valores negativos são *não físicos*, mas estatisticamente necessários!**

III – Variâncias e covariâncias

Já vimos a origem das variâncias e como estimá-las.

Covariância é uma espécie de variância em comum entre dois dados

Exemplo:

Massa 1= Massa líq. 1 + Massa recipiente

Massa 2= Massa líq. 2 + Massa recipiente

Representando variâncias e covariâncias

$$\mathbf{V_Y} = \begin{pmatrix} \sigma_1^2 & \mathrm{cov}(y_1 y_2) & \cdots & \mathrm{cov}(y_1 y_n) \\ \mathrm{cov}(y_1 y_2) & \sigma_2^2 & \cdots & \mathrm{cov}(y_2 y_n) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{cov}(y_1 y_n) & \mathrm{cov}(y_2 y_n) & \cdots & \sigma_n^2 \end{pmatrix}$$

Propagação de variâncias –uma variável

Esboço de uma dedução. Se *z* é uma função que depende de *y*, *z(y)*, expandindo *z* até primeira ordem em torno de y_0

$$z(y) \cong z(y_0) + \left.\frac{dz}{dy}\right|_{y_0} (y - y_0) \qquad \text{então}$$

$$z_0 = \langle z(y) \rangle \cong z(y_0) + \left.\frac{dz}{dy}\right|_{y_0} \langle (y - y_0) \rangle = z(y_0)$$

pois $\langle y \rangle = y_0$

- A variância de *z* pode ser calculada assim:

$$\sigma_z^2 = \left\langle (z - z_0)^2 \right\rangle \cong \left\langle \left(\left(\frac{dz}{dy_0} \right) (y - y_0) \right)^2 \right\rangle$$

$$\sigma_z^2 \cong \left(\frac{dz}{dy_0} \right)^2 \left\langle (y - y_0)^2 \right\rangle = \left(\frac{dz}{dy_0} \right)^2 \sigma_y^2$$

Propagação de matrizes de covariância

- Repetindo o procedimento com mais funções e mais variáveis

$$z_1(y_1,y_2,\ldots y_n),\ z_2(y_1,y_2,\ldots y_n),\ \ldots\ z_m(y_1,y_2,\ldots y_n),$$

são m funções de n variáveis:

$$\mathbf{V_Z} \cong \mathbf{D} \cdot \mathbf{V_Y} \cdot \mathbf{D}^t$$

$$\mathbf{D} = \begin{pmatrix} \dfrac{\partial z_1}{\partial y_{01}} & \dfrac{\partial z_1}{\partial y_{02}} & \cdots & \dfrac{\partial z_1}{\partial y_{0n}} \\ \dfrac{\partial z_2}{\partial y_{01}} & \dfrac{\partial z_2}{\partial y_{02}} & \cdots & \dfrac{\partial z_2}{\partial y_{0n}} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial z_m}{\partial y_{01}} & \dfrac{\partial z_m}{\partial y_{02}} & \cdots & \dfrac{\partial z_m}{\partial y_{0n}} \end{pmatrix}$$

Exemplo 1

- Lados de um retângulo: y_1=105 (4), y_2=48 (3)

$$\mathbf{V} = \begin{bmatrix} 16 & 0 \\ 0 & 9 \end{bmatrix}$$

  - Variância do perímetro, $z=2(y_1+y_2)$

$$\mathbf{D} = \left(\partial z / \partial y_1 \quad \partial z / \partial y_2 \right) = \left(2 \quad 2 \right)$$

fazendo as contas ..

$$\mathbf{V}_z = \begin{pmatrix} 2 & 2 \end{pmatrix} \cdot \begin{pmatrix} 16 & 0 \\ 0 & 9 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 2 \end{pmatrix} = (100)$$

o desvio padrão do perímetro é

$$\sigma_z = \sqrt{100} = 10$$

Perímetro é 306 (10) ou 306±10

Exemplo 2

- Variância da área $z=y_1 \times y_2$. Neste caso não linear, a variância é aproximada

$$\mathbf{D} = \begin{pmatrix} \partial z/\partial y_1 & \partial z/\partial y_2 \end{pmatrix} = \begin{pmatrix} y_2 & y_1 \end{pmatrix} = \begin{pmatrix} 48 & 105 \end{pmatrix}$$

$$\mathbf{V}_z \approx \begin{pmatrix} 48 & 105 \end{pmatrix} \cdot \begin{pmatrix} 16 & 0 \\ 0 & 9 \end{pmatrix} \cdot \begin{pmatrix} 48 \\ 105 \end{pmatrix} = (136089)$$

área = 5040 (368)

Exemplo 3

- Perímetro e área, $z_1=2(y_1+y_2)$, $z_2=y_1 \cdot y_2$. Variância e covariância

$$\mathbf{D} = \begin{pmatrix} \partial z_1/\partial y_1 & \partial z_1/\partial y_2 \\ \partial z_2/\partial y_1 & \partial z_2/\partial y_2 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 48 & 105 \end{pmatrix}$$

$$\mathbf{V} = \begin{pmatrix} 2 & 2 \\ 48 & 105 \end{pmatrix} \begin{pmatrix} 16 & 0 \\ 0 & 9 \end{pmatrix} \begin{pmatrix} 2 & 48 \\ 2 & 105 \end{pmatrix} = \begin{pmatrix} 100 & 3426 \\ 3426 & 136089 \end{pmatrix}$$

$$\mathbf{V} = \begin{pmatrix} 100 & 3426 \\ 3426 & 136089 \end{pmatrix}$$

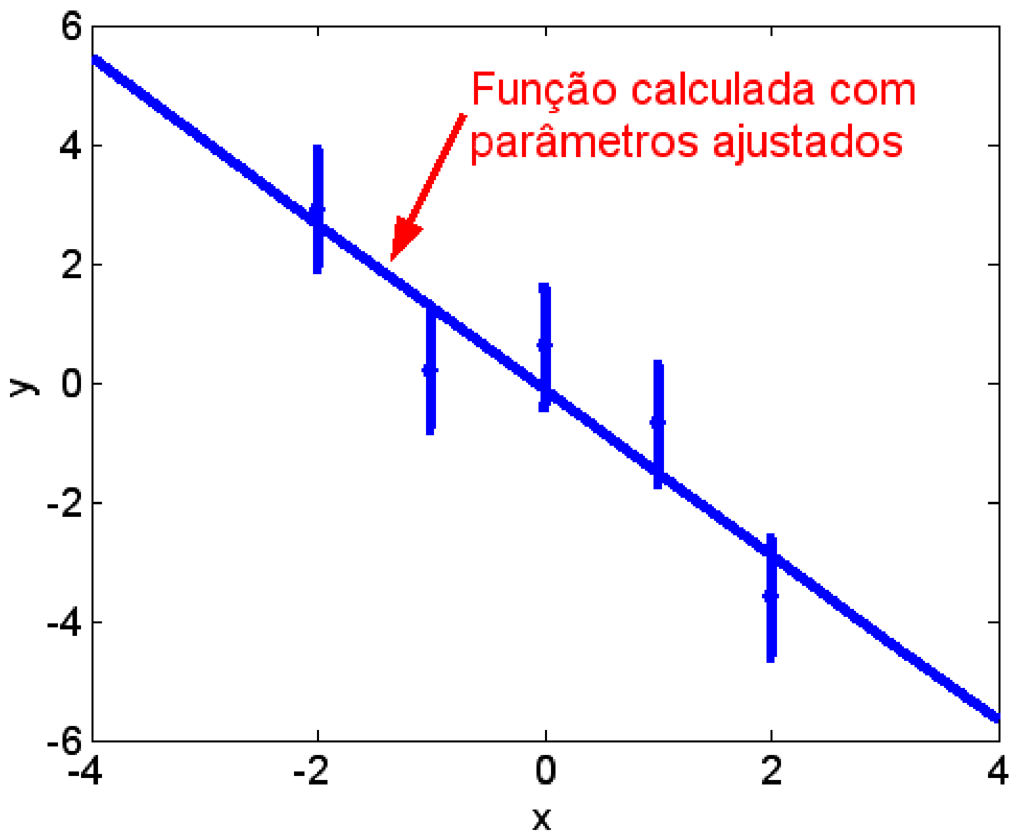
- As variâncias do perímetro e da área aparecem na diagonal e a covariância entre ambas, fora da diagonal

Perímetro = 306±10
Área = 5040±368
Covariância = 3426

Exemplo importante

Ajuste dos parâmetros de uma reta *y=a+bx* (mais adiante, veremos como estimar a matriz de covariância dos parâmetros ajustados).

x	y
-2	2,9
-1	0,2
0	0,6
1	-0,7
2	-3,6



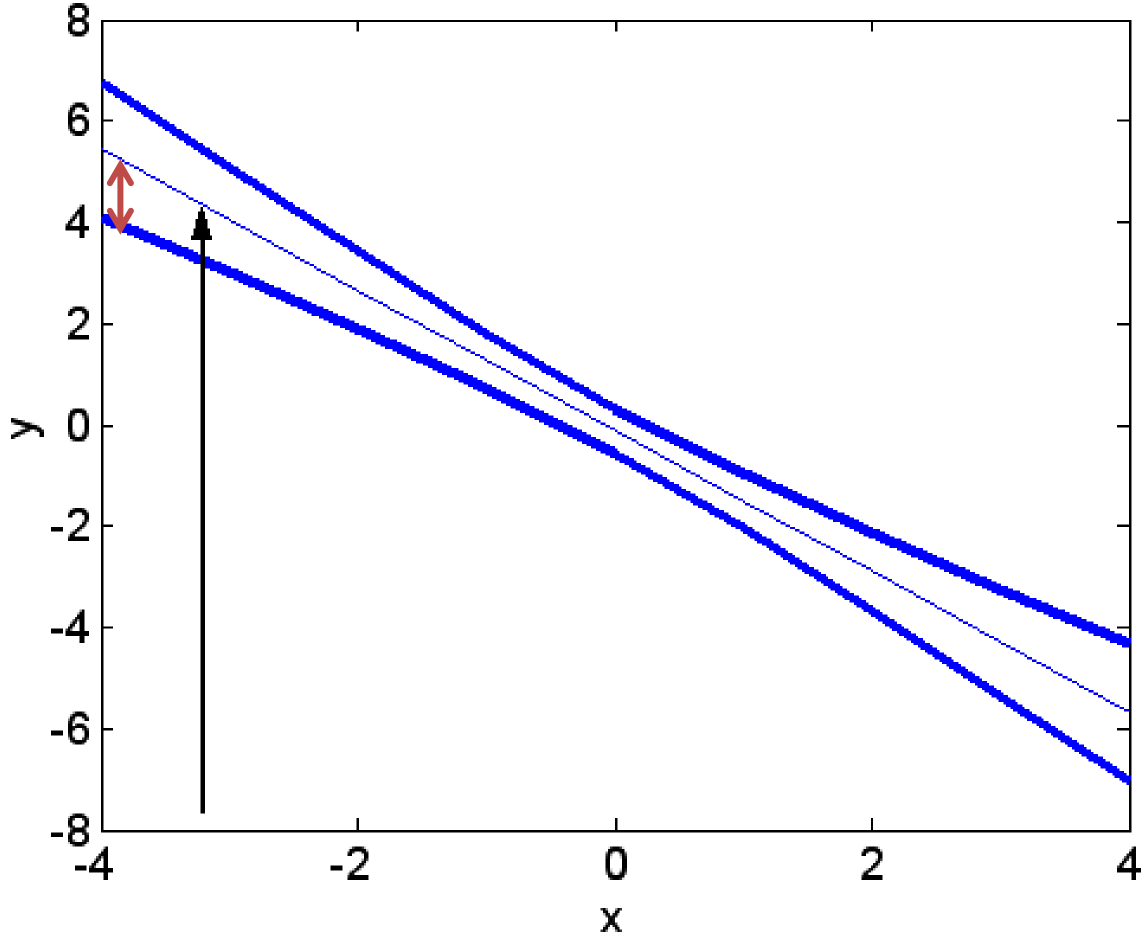
Matriz de covariância dos parâmetros *a*, *b* ajustados

0,20	0
0	0,10

Variâncias em interpolações $y_{int}=a_{aj}+b_{aj}x$

$$(\sigma_y^2) = \begin{pmatrix} 1 & x \end{pmatrix} \cdot \begin{pmatrix} 0,20 & 0 \\ 0 & 0,10 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ x \end{pmatrix} = (0,20 + 0,10 \cdot x^2)$$

Incertezas nas interpolações e extrapolações



- Covariâncias entre valores interpolados ou extrapolados:

$$\mathbf{V}_{y_a y_b} = \begin{pmatrix} 1 & x_a \\ 1 & x_b \end{pmatrix} \cdot \begin{pmatrix} 0{,}20 & 0 \\ 0 & 0{,}10 \end{pmatrix} \cdot \begin{pmatrix} 1 & x_a \\ 1 & x_b \end{pmatrix}^t$$

$$\mathbf{V}_{y_a y_b} = \begin{pmatrix} 0{,}20 + 0{,}10 x_a^2 & 0{,}20 + 0{,}10 x_a x_b \\ 0{,}20 + 0{,}10 x_a x_b & 0{,}20 + 0{,}10 x_b^2 \end{pmatrix}$$

Valores interpolados ou extrapolados são covariantes; são correlacionados uns com os outros. Isso é importante.

Relevância da correlação entre valores interpolados: cálculo da incerteza de $\Delta y = y_a - y_b$

$$(\sigma^2_{\Delta y}) = \begin{pmatrix} 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 0{,}20 + 0{,}10 x_a^2 & 0{,}20 + 0{,}10 x_a x_b \\ 0{,}20 + 0{,}10 x_a x_b & 0{,}20 + 0{,}10 x_b^2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$= [0{,}10 \cdot (x_a^2 + x_b^2) - 2 \cdot 0{,}10 \cdot x_a \cdot x_b]$$

Se $x_a = x_b$, evidentemente a incerteza em Δy será nula, como esperado.

Fontes de covariância

- Cálculos a partir de valores ajustados
- Medidas com um mesmo equipamento (a incerteza do equipamento afetará todos os dados igualmente). Ex: régua, aceleradores

Não confundir covariância com *erro sistemático*

IV a – Equações do MMQ com matrizes

- As equações do MMQ escritas de forma matricial são mais simples do que na forma tradicional. Como isso só ficará totalmente claro daqui a pouco, peço alguma paciência.

Estebelecimento do problema
a) Peso mãe e criança

$m_0 \rightarrow$ 60 kg $\qquad$ $m_0 + f_0 \rightarrow$ 65 kg $\qquad$ $f_0 \rightarrow$ 4 kg

$m = 60$

$m + f = 65$ $\qquad$ SISTEMA INCONSISTENTE, POIS HÁ ERROS

$f = 4$

$$60kg = m_0 + e_1$$
$$65kg = m_0 + f_0 + e_2$$
$$4kg = f_0 + e_3$$

m_0, f_0 ➔ valores verdadeiros (e desconhecidos) das massas da mãe e do filho

e_1 etc. ➔ **erros** de medida.

Essas equações podem ser escritas como

$$\begin{pmatrix} 60 \\ 65 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} m_0 \\ f_0 \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix}$$

- **Atenção: erro= diferença entre valor verdadeiro e valor experimental.**

- **Não confundir *erro* com *desvio padrão***

- **Erro é desconhecido (desvio padrão é conhecido)**

b) Parâmetros de uma reta

$$y = a_0 + b_0 x$$

x_1	y_1
x_2	y_2
x_3	y_3

$$y_1 = a_0 + b_0 \cdot x_1 + e_1$$
$$y_2 = a_0 + b_0 \cdot x_2 + e_2$$
$$y_3 = a_0 + b_0 \cdot x_3 + e_3$$

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix}$$

Notem: as equações são lineares nos parâmetros (*m_0, f_0, a_0, b_0*).

c) Caso geral

$$y_1 = a_{01}x_{11} + a_{02}x_{12} + \ldots + a_{0m}x_{1m} + e_1$$
$$y_2 = a_{01}x_{21} + a_{02}x_{22} + \ldots + a_{0m}x_{2m} + e_2$$
$$\vdots$$
$$y_n = a_{01}x_{n1} + a_{02}x_{n2} + \ldots + a_{0m}x_{nm} + e_n$$

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{pmatrix} \cdot \begin{pmatrix} a_{01} \\ a_{02} \\ \vdots \\ a_{0m} \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}$$

Escrevendo o problema

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{pmatrix} \cdot \begin{pmatrix} a_{01} \\ a_{02} \\ \vdots \\ a_{0m} \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}$$

$$\mathbf{Y} = \mathbf{X} \cdot \mathbf{A_0} + \mathbf{e}$$

$$\mathbf{Y} = \mathbf{X} \cdot \mathbf{A_0} + \mathbf{e}$$

O que precisamos minimizar:

$$Q(\mathbf{A}) = (\mathbf{Y} - \mathbf{X} \cdot \mathbf{A})^{\mathbf{t}} \cdot \mathbf{V}^{-1} \cdot (\mathbf{Y} - \mathbf{X} \cdot \mathbf{A})$$

V é a matriz de covariância dos dados
Quando não há covariância entre os dados, essa expressão se reduz à

$$Q(a,b,c,\ldots) = \sum_{i=1}^{n} \frac{\left(y_i - af_i - bg_i - ch_i \ldots\right)^2}{\sigma_i^2}$$

Estimativa dos parâmetros pelo MMQ

Derivando $Q(\mathbf{A})$ em relação a cada um dos parâmetros e igualando a zero obtemos o valor ajustado,

$$\frac{\partial Q}{\partial a_j} = 0\,,\; j = 1,\, 2,\, \ldots m$$

Vamos chamar o valor ajustado de **Ã**. Ele é dado por

$$\tilde{\mathbf{A}} = (\mathbf{X}^{\mathrm{t}} \cdot \mathbf{V}^{-1} \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^{\mathrm{t}} \cdot \mathbf{V}^{-1} \cdot \mathbf{Y}$$

Exemplo 1: pesos, mãe e filho

$$\begin{pmatrix} 60 \\ 65 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} m_0 \\ f_0 \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix}$$

- Para continuar o problema, vamos supor que todos os desvios padrões sejam iguais a 1 (**V**=matriz identidade de ordem *n*=3) e não haja covariância entre os dados

$$\begin{pmatrix} 60 \\ 65 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} m_0 \\ f_0 \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix}$$

$$\mathbf{Y} = \begin{bmatrix} 60 \\ 65 \\ 4 \end{bmatrix} \qquad \mathbf{X} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \qquad \mathbf{V} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\tilde{\mathbf{A}} = (\mathbf{X}^{\mathbf{t}} \cdot \mathbf{V}^{-1} \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^{\mathbf{t}} \cdot \mathbf{V}^{-1} \cdot \mathbf{Y}$$

$$\tilde{\mathbf{A}} = \begin{bmatrix} 60{,}3 \\ 4{,}3 \end{bmatrix}$$

Evidentemente, o mesmo resultado que havíamos obtido usando o procedimento tradicional

Exemplo 2: parâmetros de uma reta

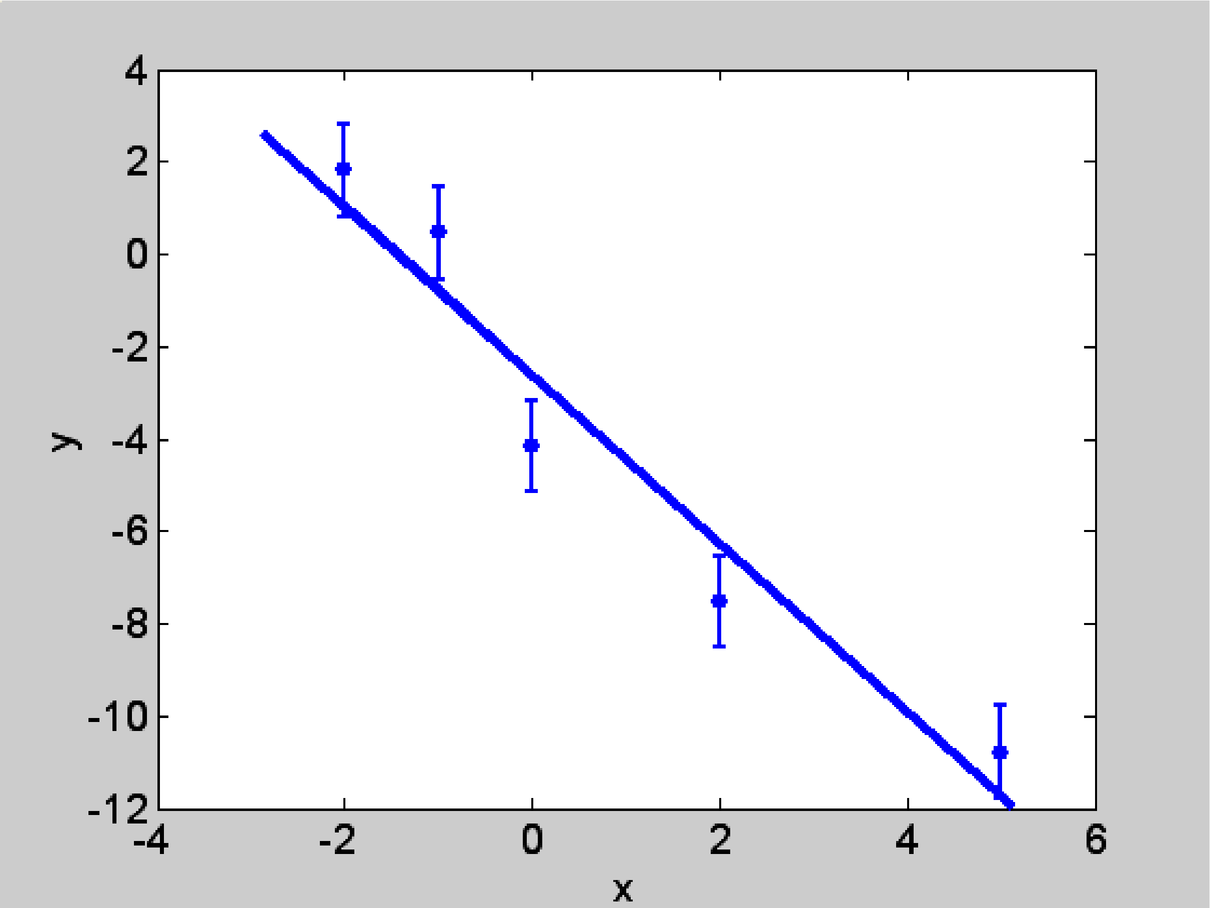
$$y = a_0 + b_o x$$

x	y
-2	1,8
-1	0,4
0	-4,2
2	-7,5
5	-10,8

$$y_1 = a_0 + b_o x_1 + e_1$$
$$y_2 = a_0 + b_o x_2 + e_2$$
$$y_3 = a_0 + b_o x_3 + e_3$$
$$y_4 = a_0 + b_o x_4 + e_4$$
$$y_5 = a_0 + b_o x_5 + e_5$$

$$\mathbf{X} = \begin{bmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 2 \\ 1 & 5 \end{bmatrix} \qquad \mathbf{Y} = \begin{bmatrix} 1{,}8 \\ 0{,}4 \\ -4{,}2 \\ -7{,}5 \\ -10{,}8 \end{bmatrix} \qquad \mathbf{V} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\tilde{\mathbf{A}} = (\mathbf{X}^t \cdot \mathbf{V}^{-1} \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^t \cdot \mathbf{V}^{-1} \cdot \mathbf{Y} \qquad \tilde{\mathbf{A}} = \begin{pmatrix} -2{,}57 \\ -1{,}85 \end{pmatrix}$$



Exemplo 3: média de dois dados

a) y_1 e y_2 com mesmos desvios padrões σ:

- $y_1 = y_o + e_1$
- $y_2 = y_o + e_2$

$$\mathbf{Y} = \mathbf{X} \cdot \mathbf{A_0} + \mathbf{e}$$

$$\mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \qquad \mathbf{X} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \mathbf{A}_0 = \left(y_o \right)$$

$$\tilde{\mathbf{A}} = (\mathbf{X}^{\mathbf{t}} \cdot \mathbf{V}^{-1} \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^{\mathbf{t}} \cdot \mathbf{V}^{-1} \cdot \mathbf{Y} \qquad \mathbf{V} = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}$$

$$\tilde{y} = \frac{y_1 + y_2}{2}$$

b) y_1 e y_2 com desvios padrões diferentes:

- $y_1 = y_o + e_1$
- $y_2 = y_o + e_2$

$$\mathbf{Y} = \mathbf{X} \cdot \mathbf{A_0} + \mathbf{e}$$

$$\mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \qquad \mathbf{X} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \mathbf{A}_0 = \left(y_o \right)$$

$$\tilde{\mathbf{A}} = (\mathbf{X}^{\mathbf{t}} \cdot \mathbf{V}^{-1} \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^{\mathbf{t}} \cdot \mathbf{V}^{-1} \cdot \mathbf{Y}$$

$$\tilde{y} = \frac{y_1 / \sigma_1^2 + y_2 / \sigma_2^2}{1/\sigma_1^2 + 1/\sigma_2^2}$$

IV b – Variâncias e covariâncias dos parâmetros ajustados

$$\tilde{\mathbf{A}} = (\mathbf{X}^{\mathbf{t}} \cdot \mathbf{V}^{-1} \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^{\mathbf{t}} \cdot \mathbf{V}^{-1} \cdot \mathbf{Y}$$

Como Ã depende de Y, sua matriz de covariância depende da matriz de covariância de Y. Vamos calcular isso por propagação.

$$\tilde{\mathbf{A}} = (\mathbf{X}^{\mathbf{t}} \cdot \mathbf{V}^{-1} \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^{\mathbf{t}} \cdot \mathbf{V}^{-1} \cdot \mathbf{Y}$$

$$\mathbf{D}_{\tilde{\mathbf{A}}} = \begin{pmatrix} \frac{\partial A_1}{\partial y_1} & \frac{\partial A_1}{\partial y_2} & \cdots & \frac{\partial A_1}{\partial y_n} \\ \frac{\partial A_2}{\partial y_1} & \frac{\partial A_2}{\partial y_2} & \cdots & \frac{\partial A_2}{\partial y_n} \\ \vdots & & \ddots & \\ \frac{\partial A_m}{\partial y_1} & \frac{\partial A_m}{\partial y_2} & \cdots & \frac{\partial A_m}{\partial y_n} \end{pmatrix} = (\mathbf{X}^{\mathbf{t}}\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^{\mathbf{t}}\mathbf{V}^{-1}$$

- Portanto,

$$\mathbf{V}_{\tilde{\mathbf{A}}} = \mathbf{D}_{\tilde{\mathbf{A}}} \cdot \mathbf{V} \cdot \mathbf{D}_{\tilde{\mathbf{A}}}^{t} = (\mathbf{X}^{\mathbf{t}}\mathbf{V}^{-1}\mathbf{X})^{-1}$$

$$\tilde{\mathbf{A}} = (\mathbf{X}^t \cdot \mathbf{V}^{-1} \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^t \cdot \mathbf{V}^{-1} \cdot \mathbf{Y}$$

$$\boxed{\mathbf{V}_{\tilde{\mathbf{A}}} = (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1}}$$

Quatro equações básicas

$$\mathbf{Y} = \mathbf{X} \cdot \mathbf{A}_0 + \mathbf{e}$$

$$\tilde{\mathbf{A}} = (\mathbf{X}^t \cdot \mathbf{V}^{-1} \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^t \cdot \mathbf{V}^{-1} \cdot \mathbf{Y}$$

$$\mathbf{V}_{\tilde{\mathbf{A}}} = (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1}$$

$$\mathbf{V}_\mathbf{Z} \cong \mathbf{D} \cdot \mathbf{V}_\mathbf{Y} \cdot \mathbf{D}^t$$

Exemplo 1: pesos, mãe e filho

$$\begin{pmatrix} 60 \\ 65 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} m_0 \\ f_0 \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} \qquad \mathbf{X} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$$

Vamos supor que todos os desvios padrões sejam iguais a 1 e não haja covariância entre os dados

$$\mathbf{V} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1}$$

$$\mathbf{V}_{\tilde{\mathbf{A}}} = (\mathbf{X}^{\mathbf{t}}\mathbf{V}^{-1}\mathbf{X})^{-1}$$

$$\mathbf{V}_{\tilde{\mathbf{A}}} = \begin{pmatrix} 0{,}67 & -0{,}33 \\ -0{,}33 & 0{,}67 \end{pmatrix} \qquad \tilde{\mathbf{A}} = \begin{bmatrix} 60{,}3 \\ 4{,}3 \end{bmatrix}$$

m ➔ 60,3±0,8 kg

f ➔ 4,3±0,8 kg

Covariância (*m*,*f*)= - 0,33

Exemplo 2: Consumo de combustível: g_c e g_e (litros/km)

distância percorrida (km)		consumo (l)
cidade	estrada	
100	200	28
300	100	43
400	0	49

$$\begin{pmatrix} 28 \\ 43 \\ 49 \end{pmatrix} = \begin{pmatrix} 100 & 200 \\ 300 & 100 \\ 400 & 0 \end{pmatrix} \cdot \begin{pmatrix} g_c \\ g_e \end{pmatrix} + \mathbf{e}$$

Vamos supor que **V** seja uma matriz identidade,

$$\mathbf{V} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1}$$

$$\tilde{\mathbf{A}} = (\mathbf{X}^{\mathbf{t}} \cdot \mathbf{V}^{-1} \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^{\mathbf{t}} \cdot \mathbf{V}^{-1} \cdot \mathbf{Y}$$

$$\tilde{\mathbf{A}} = \begin{pmatrix} 0{,}12 \\ 0{,}077 \end{pmatrix}$$

$$\mathbf{V}_{\tilde{\mathbf{A}}} = (\mathbf{X}^{\mathbf{t}}\mathbf{V}^{-1}\mathbf{X})^{-1}$$

$$\mathbf{V}_{\tilde{\mathbf{A}}} = \begin{pmatrix} 0{,}048 & -0{,}048 \\ -0{,}048 & 0{,}248 \end{pmatrix} \cdot 10^{-4}$$

Consumo em quilômetros/litro?

$y_c=1/g_c=8{,}27\text{km/l}$, $y_e=1/g_e=13{,}0\text{km/l}$

$$\mathbf{V}_y \approx \mathbf{D}\cdot\mathbf{V}\cdot\mathbf{D}^t \qquad \mathbf{D} = \begin{pmatrix} \dfrac{\partial y_c}{\partial g_c} & \dfrac{\partial y_c}{\partial g_e} \\ \dfrac{\partial y_e}{\partial g_c} & \dfrac{\partial y_e}{\partial g_e} \end{pmatrix} = \begin{pmatrix} \dfrac{-1}{g_c^2} & 0 \\ 0 & \dfrac{-1}{g_e^2} \end{pmatrix}$$

$$\mathbf{V}_y = \begin{pmatrix} 0.022 & -0.055 \\ -0.055 & 0.703 \end{pmatrix}$$

$y_c=(8{,}27\pm0{,}15)$ km/l

$y_e=(13{,}0\pm0{,}15)$ km/l

Vale a pena insistir: resumo

Relação linear entre dados e parâmetros

$$\mathbf{Y} = \mathbf{X} \cdot \mathbf{A_0} + \mathbf{e}$$

A matriz de covariância dos dados deve ser conhecida (os erros são desconhecidos)

$$\tilde{\mathbf{A}} = (\mathbf{X}^t \cdot \mathbf{V}^{-1} \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^t \cdot \mathbf{V}^{-1} \cdot \mathbf{Y} \qquad \mathbf{V}_{\tilde{\mathbf{A}}} = (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1}$$

Em nenhum momento foi feita qualquer hipótese quanto a forma da função densidade de probabilidade dos dados.

V – Algumas propriedades do MMQ e o TCL
(a) Não tendenciosidade

resultado	probabilidade
1,8 s	25%
2,0 s	50%
2,2 s	25%

Espera-se que se *y* é uma medida de uma grandeza (um dado, uma média, o resultado de um ajuste...), seu valor esperado seja igual ao valor verdadeiro da grandeza,

$$\langle y \rangle = y_0$$

- **O MMQ é não tendencioso quando os dados também são não tendenciosos**

Se y_i representa os dados e $<y_i>=y_{0i}$, como $y_i=y_{0i}+e_i$, onde e_i é o erro, então $<e_i>=0$.

Vamos ver como fica o valor esperado de parâmetros ajustados pelo MMQ:

$$\tilde{\mathbf{A}} = (\mathbf{X}^{\mathbf{t}} \cdot \mathbf{V}^{-1} \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^{\mathbf{t}} \cdot \mathbf{V}^{-1} \cdot \mathbf{Y}$$

O valor esperado de Ã é

$$\left\langle \tilde{\mathbf{A}} \right\rangle = (\mathbf{X}^{\mathbf{t}} \cdot \mathbf{V}^{-1} \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^{\mathbf{t}} \cdot \mathbf{V}^{-1} \cdot \left\langle \mathbf{Y} \right\rangle$$

Como os dados se relacionam com os parâmetros na forma

$$\mathbf{Y} = \mathbf{X} \cdot \mathbf{A_0} + \mathbf{e}$$

então

$$\left\langle \mathbf{Y} \right\rangle = \mathbf{X} \cdot \mathbf{A_0} + \left\langle \mathbf{e} \right\rangle = \mathbf{X} \cdot \mathbf{A_0}$$

portanto,

$$\left\langle \widetilde{\mathbf{A}} \right\rangle = (\mathbf{X}^{\mathbf{t}} \cdot \mathbf{V}^{-1} \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^{\mathbf{t}} \cdot \mathbf{V}^{-1} \cdot \left\langle \mathbf{Y} \right\rangle$$
$$= (\mathbf{X}^{\mathbf{t}} \cdot \mathbf{V}^{-1} \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^{\mathbf{t}} \cdot \mathbf{V}^{-1} \cdot \mathbf{X} \cdot \mathbf{A_0}$$

$$\left\langle \widetilde{\mathbf{A}} \right\rangle = \mathbf{A_0}$$

Conclusão: se os dados não são tendenciosos, os parâmetros ajustados pelo MMQ também não serão tendenciosos

(b) Mínima variância

- **Entre todas as estimativas não tendenciosas de parâmetros que dependem linearmente dos dados, o MMQ fornece aquelas com menor variância**

- Exemplo: dois dados com mesmo desvio padrão σ e não covariantes. O resultado do MMQ é

$$\bar{x} = \frac{x_1 + x_2}{2} \qquad\qquad \sigma_{\bar{x}}^2 = \sigma^2 / 2$$

Outra estimativa geral, linear nos dados e não tendenciosa:

$$\overline{x}' = a x_1 + (1-a) x_2$$

Qual valor de *a* leva à menor variância dessa estimativa? Por propagação de incerteza,

$$\sigma^2_{\overline{x}'} = (a^2 + (1-a)^2)\sigma^2$$

- Derivando em relação à *a* e igualando a zero, obtemos *a*=1/2 e, portanto,

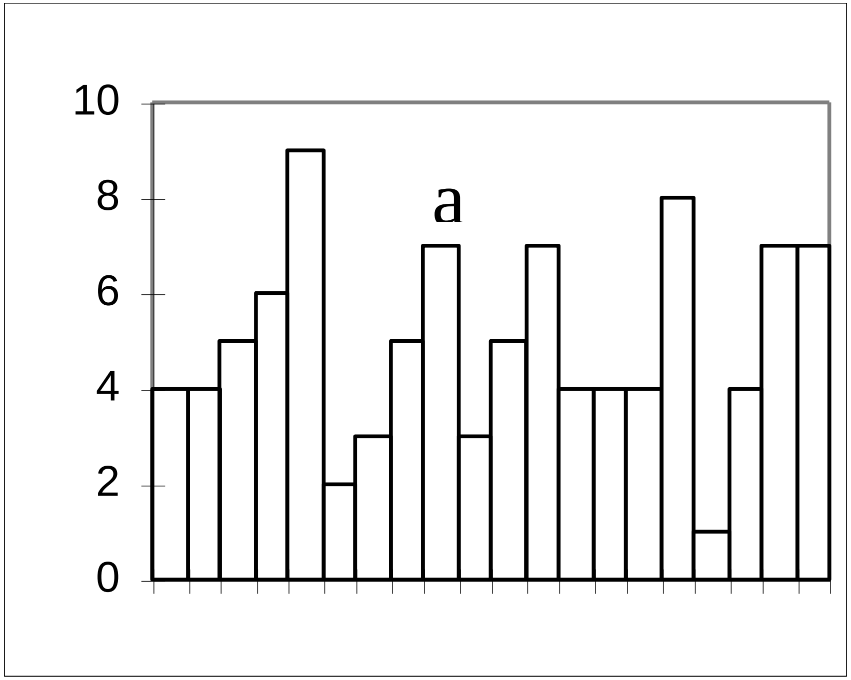
$$\bar{x} = \frac{x_1 + x_2}{2}$$

que é o resultado do MMQ:

(c) O Teorema Central do Limite

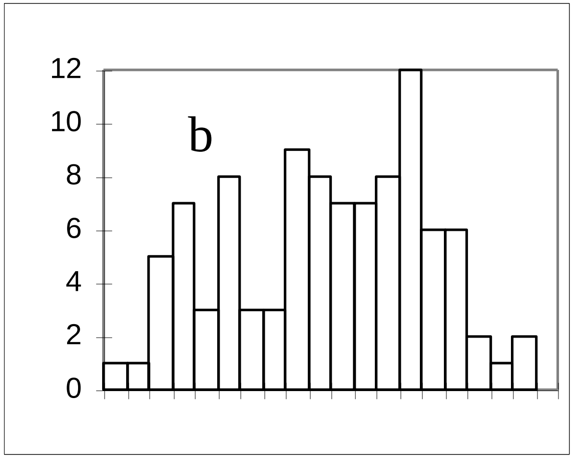
O valor ajustado obedece a uma distribuição normal quando ele é resultado de muitos dados

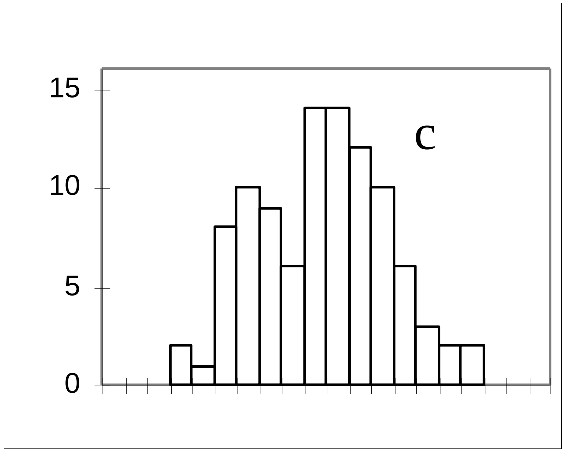
Exemplo: médias de dados uniformemente distribuídos entre zero e um.

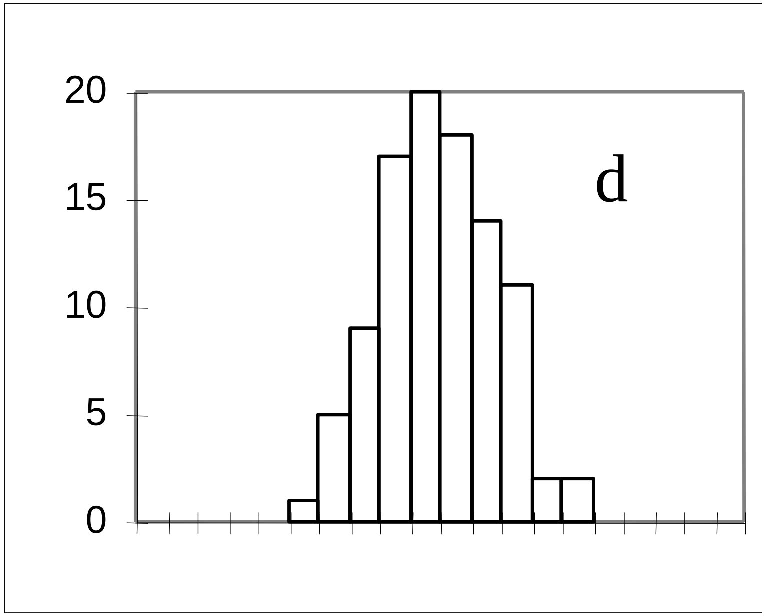


Um só dado

- Média de dois, quatro e oito dados







- **O TCL garante que**

$$\bar{x} = \frac{x_1/\sigma_1^2 + x_2/\sigma_2^2 + \cdots + x_n/\sigma_n^2}{1/\sigma_1^2 + 1/\sigma_2^2 + \cdots + 1/\sigma_n^2}$$

tende uma gaussiana com desvio padrão igual a 1, quaisquer que sejam as f.d.p.s dos dados

O TCL no contexto do MMQ

$$\tilde{\mathbf{A}} = (\mathbf{X}^{\mathbf{t}} \cdot \mathbf{V}^{-1} \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^{\mathbf{t}} \cdot \mathbf{V}^{-1} \cdot \mathbf{Y}$$

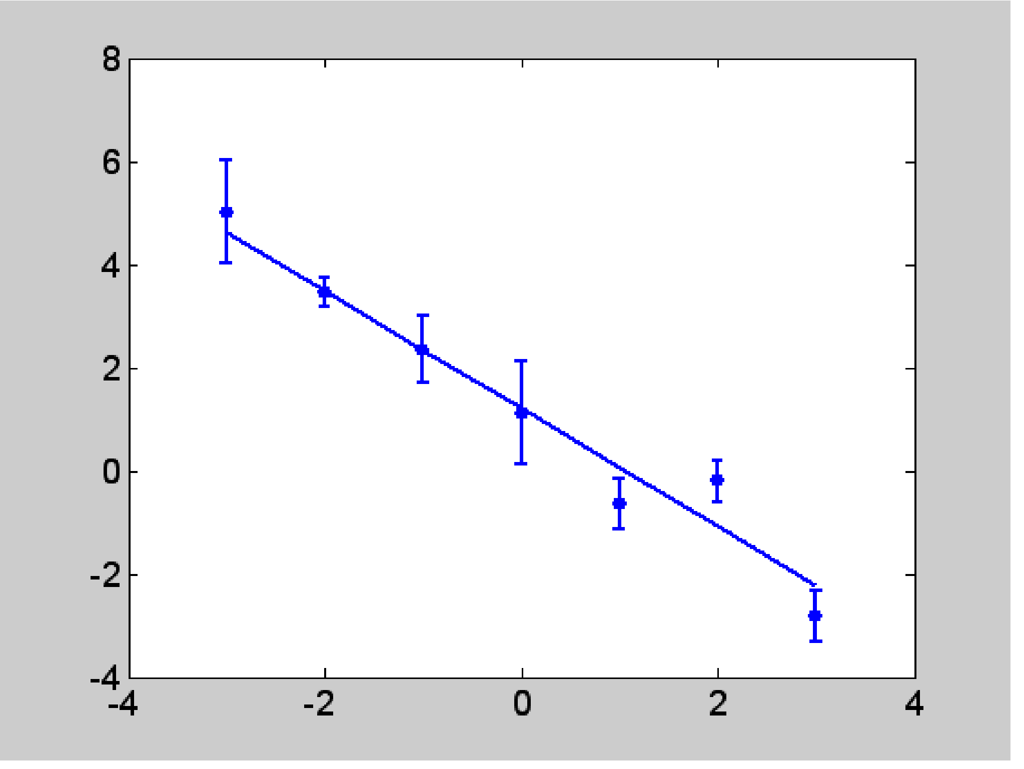
Ã é uma média dos vários dados experimentais y_i, com pesos inversamente às respectivas variâncias

Desde que a quantidade de dados seja suficientemente grande, cada parâmetro deverá obedecer a uma distribuição normal, independente da forma da f. d. p. de cada dado

Exemplo

x	y	d.p.	Tipo
-3	5,0	1,0	gaussiana
-2	3,5	0,3	uniforme
-1	2,3	0,7	tipo qui2
0	1,1	1,0	gaussiana
1	-0,6	0,5	binomial
2	-0,2	0,4	duas uniformes
3	-2,8	0,5	binomial

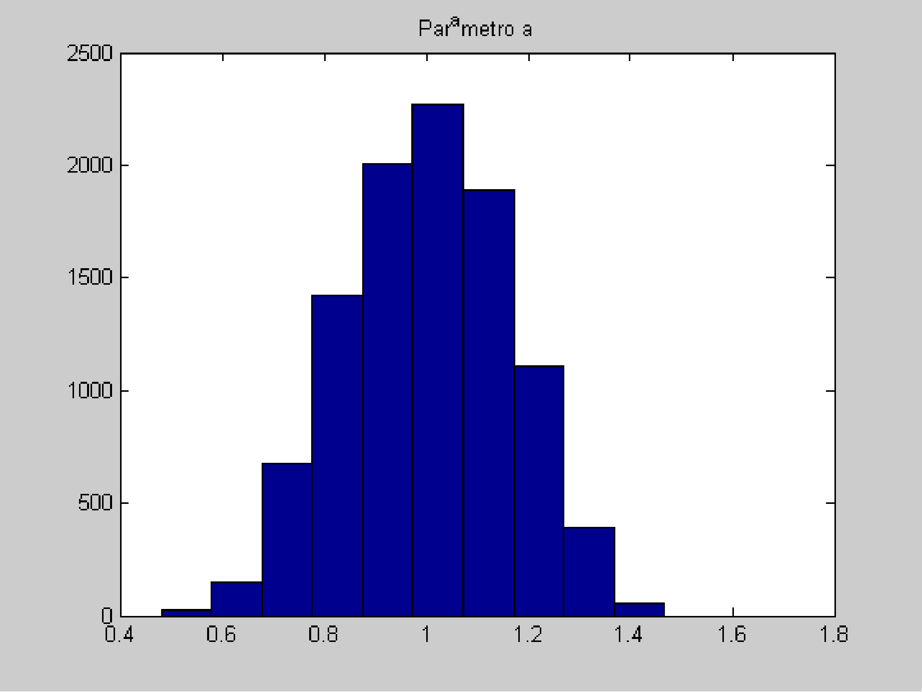
$y=a+bx$, onde os diferentes valores y_i obedecem a diferentes f.d.p.s

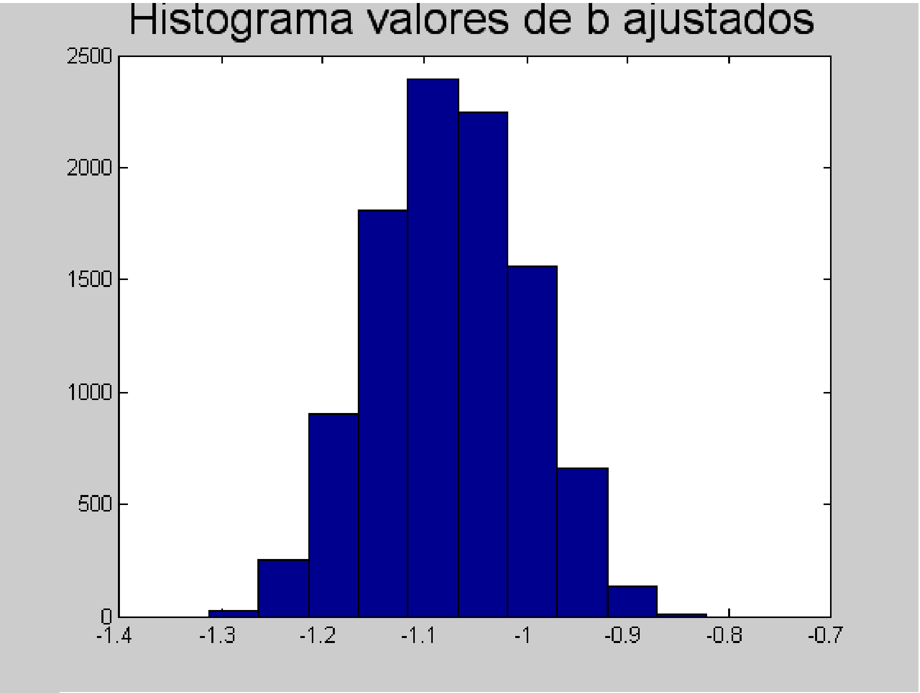


Parâmetros *a* e *b* ajustados: 1,20(18); -1,14(9) devem obedecer a uma f,d,p, gaussiana ou bem perto disso,

Seria exatamente gaussiana se a quantidade de dados fosse infinita

Ilustração com muitas simulações
Valores ajustados de *a* e *b*





VI – Outros desenvolvimentos

- a) Medida de uma grandeza altera valores de outras
- b) Como incorporar um dado perdido
- c) Vínculos entre parâmetros
- d) Média de dados correlacionados

- As próximas transparências discutem esses problemas

a) Medida de uma grandeza altera outras

- Exemplo: *a* e *b* foram medidos, obtendo-se os valores 10,0 e 20,0, com matriz de covariância

$$\mathbf{V} = \begin{pmatrix} 10 & 8 \\ 8 & 10 \end{pmatrix}$$

Uma medida independente de *b* forneceu o valor 25,0 com variância igual a 10. Vamos ao MMQ.

$$\mathbf{Y} = \begin{pmatrix} 10 \\ 20 \\ 25 \end{pmatrix} \qquad \mathbf{V} = \begin{pmatrix} 10 & 8 & 0 \\ 8 & 10 & 0 \\ 0 & 0 & 10 \end{pmatrix} \qquad \begin{pmatrix} 10 \\ 20 \\ 25 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix}$$

$$\tilde{\mathbf{A}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{Y} = \begin{pmatrix} 12{,}0 \\ 22{,}5 \end{pmatrix} \qquad \mathbf{V} = \begin{pmatrix} 6{,}8 & 4{,}0 \\ 4{,}0 & 5{,}0 \end{pmatrix}$$

- Evidentemente, valor adotado para *b* foi alterado; seu d. p., idem.
- Menos evidente: o valor adotado para *a* é alterado e seu d.p., idem.

b) Como incorporar um novo dado a um ajuste

Parâmetros a e b foram ajustados $\qquad y = a + bx$

$$\tilde{\boldsymbol{A}} = \begin{bmatrix} \tilde{a} \\ \tilde{b} \end{bmatrix} \qquad \boldsymbol{V}_{\tilde{\boldsymbol{A}}} = \begin{bmatrix} \sigma_a^2 & \rho.\sigma_a.\sigma_b \\ \rho.\sigma_a.\sigma_b & \sigma_b^2 \end{bmatrix}$$

Novo dado é obtido $\qquad (x, y, \sigma)$

$$\begin{bmatrix} \tilde{a} \\ \tilde{b} \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & x \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ b_0 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix} \qquad \boldsymbol{V}_{\tilde{\boldsymbol{a}},\tilde{\boldsymbol{b}},\boldsymbol{y}} = \begin{bmatrix} \sigma_a^2 & \rho.\sigma_a.\sigma_b & 0 \\ \rho.\sigma_a.\sigma_b & \sigma_b^2 & 0 \\ 0 & 0 & \sigma^2 \end{bmatrix}$$

c) Vínculos entre parâmetros

- Exemplo: mede-se os três ângulos internos de um triângulo, digamos, θ_1, θ_2, e θ_3, sendo os resultados y_1,σ_1; y_2, σ_2; y_3, σ_3. Mas queremos impor a condição $\tilde{\theta}_1 + \tilde{\theta}_2 + \tilde{\theta}_3 = 180^o$.

- Um procedimento geral é ajustar valores para os parâmetros, θ_1, θ_2 e θ_3, impondo o vínculo.

- Uma forma geral de escrever vínculo lineares:

$$\sum_{j=1}^{m} g_{ij} a_j = r_i$$

onde a_j são os parâmetros a serem ajustados e g_{ij} e r_i são valores conhecidos.

No caso do triangulo só há um vínculo (r=180º) e todos os gs são iguais a 1.

A equações de vínculo podem ser escritas como

$$\mathbf{GA} = \mathbf{R}$$

Solução do MMQ:

$$\tilde{\mathbf{A}} = \mathbf{V}_{\tilde{\mathbf{A}}} \mathbf{X}^t \mathbf{V}^{-1} \mathbf{Y} - \mathbf{BC}^{-1} \mathbf{R}$$

$$\mathbf{V}_{\tilde{\mathbf{A}}} = \mathbf{H}^{-1} + \mathbf{BC}^{-1} \mathbf{B}^t$$

$$\mathbf{H} = \mathbf{X}^t \mathbf{V}^{-1} \mathbf{X} \qquad \mathbf{B} = \mathbf{H}^{-1} \mathbf{G}^t \qquad \mathbf{C} = -\mathbf{GH}^{-1} \mathbf{G}^t$$

Exemplo

Ângulos de um triângulo: 48º, 41º, 94º (soma=183º) todos com σ=1 e não covariantes. G=[1 1 1] e R=[180].

$$\mathbf{Y} = \begin{pmatrix} 48 \\ 41 \\ 94 \end{pmatrix} \qquad \mathbf{V} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Resultado: Valores ajustados: 47º, 40º, 93º . Soma é 180º: obedece vínculo. As variâncias diminuem

$$\mathbf{V} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \Rightarrow \mathbf{V} = \begin{pmatrix} 0{,}67 & -0{,}33 & -0{,}33 \\ -0{,}33 & 0{,}67 & -0{,}33 \\ -0{,}33 & -0{,}33 & 0{,}67 \end{pmatrix}$$

Vela a pena impor vínculos? Depende. Incertezas diminuem, mas surgem correlações.

P. ex., se o que queremos é apenas um dos ângulos, sim.

Breve resumo

- O MMQ deve ser usado:
  - Sempre que as incertezas dos dados sejam conhecidas (ou todas sejam iguais)
  - Há uma relação linear entre os dados e os parâmetros a serem ajustados

d)Média de dados correlacionados

Considere uma situação na qual temos *n* dados correspondentes à uma mesma grandeza, y_1, y_2, ... y_n.

Para simplificar, vamos supor que todos tenham o mesmo desvio padrão σ.

Caso os dados fossem não correlacionados, o valor a ser adotado seria a média simples dos dados e o desvio padrão dessa média seria $\sigma/\sqrt{n}$

Suponha, entretanto, que a matriz de covariância dos dados seja

$$\mathbf{V} = \sigma^2 \begin{bmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \ddots & \vdots \\ & & & \vdots \\ \rho & \rho & \cdots & 1 \end{bmatrix}$$

Todas as variâncias são iguais a σ^2.

Todas as covariâncias todas iguais a $\rho \cdot \sigma^2$.

O problema pode ser colocado dentro do esquema do MMQ, com

$$\begin{aligned} y_1 &= y_0 + e_1 \\ y_2 &= y_0 + e_2 \\ &\vdots \\ y_n &= y_0 + e_n \end{aligned} \qquad \mathbf{X} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

Os resultados para o valor adotado e a variância são

$$\tilde{y} = \frac{1}{n} \sum_{i=1}^{n} y_i \qquad \sigma_{\tilde{y}}^2 = \frac{(\rho n - \rho + 1)\sigma^2}{n}$$

Como no caso de dados não correlacionados, o valor adotado é a média simples dos dados

$$\tilde{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

Resultado esperado, pois todas as variâncias e covariâncias são iguais e, portanto, nenhum dado poderia ter peso maior que os demais.

A variância depende da correlação

$$\sigma_{\tilde{y}}^2 = \frac{(\rho n - \rho + 1)\sigma^2}{n}$$

Note que se ρ=0, a variância desse valor é $\sigma/\sqrt{n}$, como esperado, e tende a zero quando *n* tende a infinito

Se ρ não é nulo, entretanto, mesmo que *n* tenda a infinito, ρ·σ². Conclusão: a covariância é um limite para a variância do valor adotado

- O ajuste é não tendencioso
- Os dados não precisam obedecer a nenhum f.d.p. particular
- Se há uma quantidade grande de dados, as f.d.p.s dos parâmetros ajustados tendem a formas gaussianas
- Não precisa haver uma relação funcional do tipo $y=y(x)$ para usar o MMQ
- As covariâncias são tão importantes quanto as variâncias