Introdução à Física Experimental LicFís & EngFís 2021 / 2022

- Método dos mínimos quadrados
  - Ajuste de uma proporcionalidade (reta que passa pela origem)
  - Ajuste de uma reta
  - Condições de aplicação e incertezas nos parâmetros do ajuste
  - Exemplos de aplicação
  - Significado geométrico do ajuste
- Uso do Excel para o ajuste de uma reta pelo método dos mínimos quadrados
- Distinção entre *método dos mínimos quadrados* e *regressão linear*
  - O coeficiente de correlação
  - O bug do Excel

Método dos mínimos quadrados

O método dos mínimos quadrados é usado para determinar a o "melhor ajuste" de uma função a um conjunto de pontos experimentais. Por exemplo, podemos estar a ajustar um polinómio do segundo grau às coordenadas *x,y* [$y(x) = a_0 + a_1 x + a_2 x^2$] de um conjunto de pontos experimentais ou estar a ajustar uma sinusoide ao som de uma nota emitida por uma flauta [$p(t) = b_0 \cdot \sin(b_1 t + b_2)$].

Para o caso do ajuste de uma reta, admitindo que as abcissas não têm incerteza ou que a sua incerteza é desprezável e considerando ainda que as incertezas nas ordenadas é gaussiana e não correlacionadas entre pontos experimentais, é possível mostrar que minimizar a soma dos quadrados dos desvios coincide com maximizar a probabilidade desse conjunto de pontos experimentais.

Método dos mínimos quadrados: $y = ax$

Consideremos um conjunto de pontos $\{(x_i, y_i)\}$ com $i = 1, 2, \ldots, N$ que seguem uma relação de proporcionalidade $y = ax$, sendo a a constante de proporcionalidade a ajustar. Devemos ainda considerar que não há incerteza nos x mas apenas nos y e que esta incerteza é gaussiana. Para determinar a minimizamos a soma do quadrado dos desvios:

$\chi^2 = \sum_{i=1}^{N} \left[\frac{y_i - ax_i}{\sigma_i}\right]^2$ (a esta soma chama-se o chi-quadrado)

O χ^2 é uma função de a e o mínimo verifica-se quando a derivada de χ^2 em ordem a a for nula. Se supusermos que todos os desvios padrões são iguais ($\sigma_i = \sigma$) obtemos:

$\hat{a} = \frac{\sum x_i y_i}{\sum x_i^2} = \frac{\overline{xy}}{\overline{x^2}}$ onde $\overline{xy} = \frac{\sum x_i y_i}{N}$ e $\overline{x^2} = \frac{\sum x_i^2}{N}$

Reescrevendo a equação na forma:

$\hat{a} = \sum \frac{x_i}{N\overline{x^2}} y_i$

Tendo em conta que não há incerteza nos x e aplicando a propagação de incertezas para uma soma obtemos:

$\widehat{\sigma_{\hat{a}}} = \sqrt{\sum \left(\frac{x_i}{N\overline{x^2}}\right)^2 \sigma_y^2} = \frac{\sigma_y}{\sqrt{N\overline{x^2}}}$

Método dos mínimos quadrados: $y = ax$

Se o desvio padrão dos y, σ_y, não for conhecido, o seu valor pode ser estimado a partir da qualidade do ajuste, uma vez ajustado $\hat{a}$:

$$\widehat{\sigma_y} = \sqrt{\frac{\sum(y_i - \hat{a}x_i)^2}{N-1}}$$

Nota: este procedimento é habitual, fornecer ao software os dados experimentais sem as incertezas associadas e o software ajusta a função (uma reta que passa pela origem, neste caso) e, a partir da qualidade do ajuste, estima a incerteza nos *y* bem como nos parâmetros ajustados, o *a*, neste caso.

Se as incertezas forem diferentes para cada ponto temos de fazer um ajuste pesado, dando pesos diferentes a cada ponto:

$$\hat{a} = \frac{\sum\left({x_i y_i}/{\sigma_i^2}\right)}{\sum\left({x_i^2}/{\sigma_i^2}\right)} = \frac{\overline{xy}}{\overline{x^2}}; \quad \text{onde} \quad \overline{xy} = \left.\sum\left(\frac{x_i y_i}{\sigma_i^2}\right)\right/\sum\left(\frac{1}{\sigma_i^2}\right) \quad \text{e} \quad \overline{x^2} = \left.\sum\left(\frac{x_i^2}{\sigma_i^2}\right)\right/\sum\left(\frac{1}{\sigma_i^2}\right)$$ são médias pesadas

$$\widehat{\sigma_{\hat{a}}} = \frac{\sigma_y}{\sqrt{N\overline{x^2}}} \quad \text{onde} \quad \sigma_y = \sqrt{\frac{N}{\sum\left({1}/{\sigma_i^2}\right)}} \quad \text{e} \quad \overline{x^2} = \left.\sum\left(\frac{x_i^2}{\sigma_i^2}\right)\right/\sum\left(\frac{1}{\sigma_i^2}\right)$$ são médias pesadas

Método dos mínimos quadrados: $y = ax$ → exemplo

Considerem os dados à direita relativos à determinação do módulo de Young. Admitindo que a relação entre as duas variáveis é uma proporcionalidade direta, podemos ajustar a reta que passa pela origem e melhor se ajusta a estes dados:

- A incerteza nas forças (massa) dividida pela gama de valores é menor do que para a deformação, por isso esta última será o *y*.
- $\sum x_i^2 = 140.22$ N^2; $\sum x_i y_i = 12.040$ mm·N
  $\hat{a} = 0.08587$ mm/N
  $\widehat{\sigma_y} = 0.00685$ mm
  $\widehat{\sigma_a} = 0.000578$ mm/N
- Verificamos que a incerteza padrão estimada para os *y* (deformação) é 0.7×10^{-2} mm, isto é, cerca de 68% dos *y* experimentais desviam-se menos do que 0.7 da menor divisão do valor verdadeiro. **Esta estimativa parece fazer sentido.**
- O valor final seria escrito como $a = (0.0859 \pm 0.0006)$ mm/N

N	mm
F	**vidro**
0.00	0
0.98	0.088
1.96	0.17
2.94	0.255
3.92	0.339
4.90	0.418
5.88	0.493
4.90	0.42
3.92	0.341
2.94	0.26
1.96	0.175
0.98	0.095
0.00	0.012

Método dos mínimos quadrados: $y = ax + b$

Consideremos um conjunto de pontos $\{(x_i, y_i)\}$ com $i = 1, 2, \ldots, N$ que seguem uma relação de proporcionalidade $y = ax + b$, sendo a e b as constantes a ajustar. Devemos ainda considerar que não há incerteza nos x mas apenas nos y e que esta incerteza é gaussiana. Para determinar a e b minimizamos a soma do quadrado dos desvios. Iríamos obter um sistema de duas equações cuja solução é:

$\chi^2 = \sum_{i=1}^{N} \left[\frac{y_i - ax_i - b}{\sigma_i}\right]^2$ (a esta soma chama-se o chi-quadrado)

O χ^2 é uma função de a e b, o mínimo verifica-se quando as derivadas parciais de χ^2 em ordem a a e a b forem nulas.

Se supusermos que todos os desvios padrões são iguais ($\sigma_i = \sigma$) obtemos:

$\hat{a} = \frac{N \cdot \sum x_i y_i - \sum x_i \cdot \sum y_i}{N \cdot \sum x_i^2 - (\sum x_i)^2} = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - \bar{x}^2}$ e $\widehat{\sigma_{\hat{a}}} = \frac{\sigma_y}{\sqrt{N \cdot (\overline{x^2} - \bar{x}^2)}}$ nota: $\overline{xy} = \frac{\sum x_i y_i}{N}$ e $\bar{x}\bar{y} = \frac{\sum x_i}{N} \cdot \frac{\sum y_i}{N}$

$\hat{b} = \frac{\sum y_i}{N} - \hat{a} \cdot \frac{\sum x_i}{N} = \bar{y} - \hat{a}\bar{x}$ e $\widehat{\sigma_{\hat{b}}} = \sqrt{\frac{\overline{x^2}}{N \cdot (\overline{x^2} - \bar{x}^2)}} \cdot \sigma_y = \sqrt{\overline{x^2}} \cdot \widehat{\sigma_{\hat{a}}}$

$\text{covariância}(\hat{a}, \hat{b}) = -\frac{\bar{x} \cdot \sigma_y^2}{N \cdot (\overline{x^2} - \bar{x}^2)}$

Método dos mínimos quadrados: $y = ax + b$ → exemplo

Considerem os dados à direita relativos à determinação do módulo de Young. Desta vez vamos ajustar uma reta que pode não passar pela origem:

- A incerteza nas forças (massa) dividida pela gama de valores é menor do que para a deformação, por isso esta última será o *y*.
- $\sum x_i^2 = 140.22$ N^2; $\sum x_i y_i = 12.040$ mm·N ; $\sum x_i = 35.25$ N ; $\sum y_i = 3.066$ mm
  $\hat{a} = 0.08364$ mm/N ; $\hat{b} = 0.00887$ mm
  $\widehat{\sigma_y} = 0.00465$ mm
  $\widehat{\sigma_a} = 0.000698$ mm/N ; $\widehat{\sigma_b} = 0.00229$ mm
- Verificamos que a incerteza padrão estimada para os *y* (deformação) é 0.5×10^{-2} mm, ligeiramente inferior à calculada anteriormente, como era de esperar. **Esta estimativa parece fazer sentido.**
- O valor final seria escrito como $a = (0.0836 \pm 0.0007)$ mm/N ; $b = (0.009 \pm 0.002)$ mm

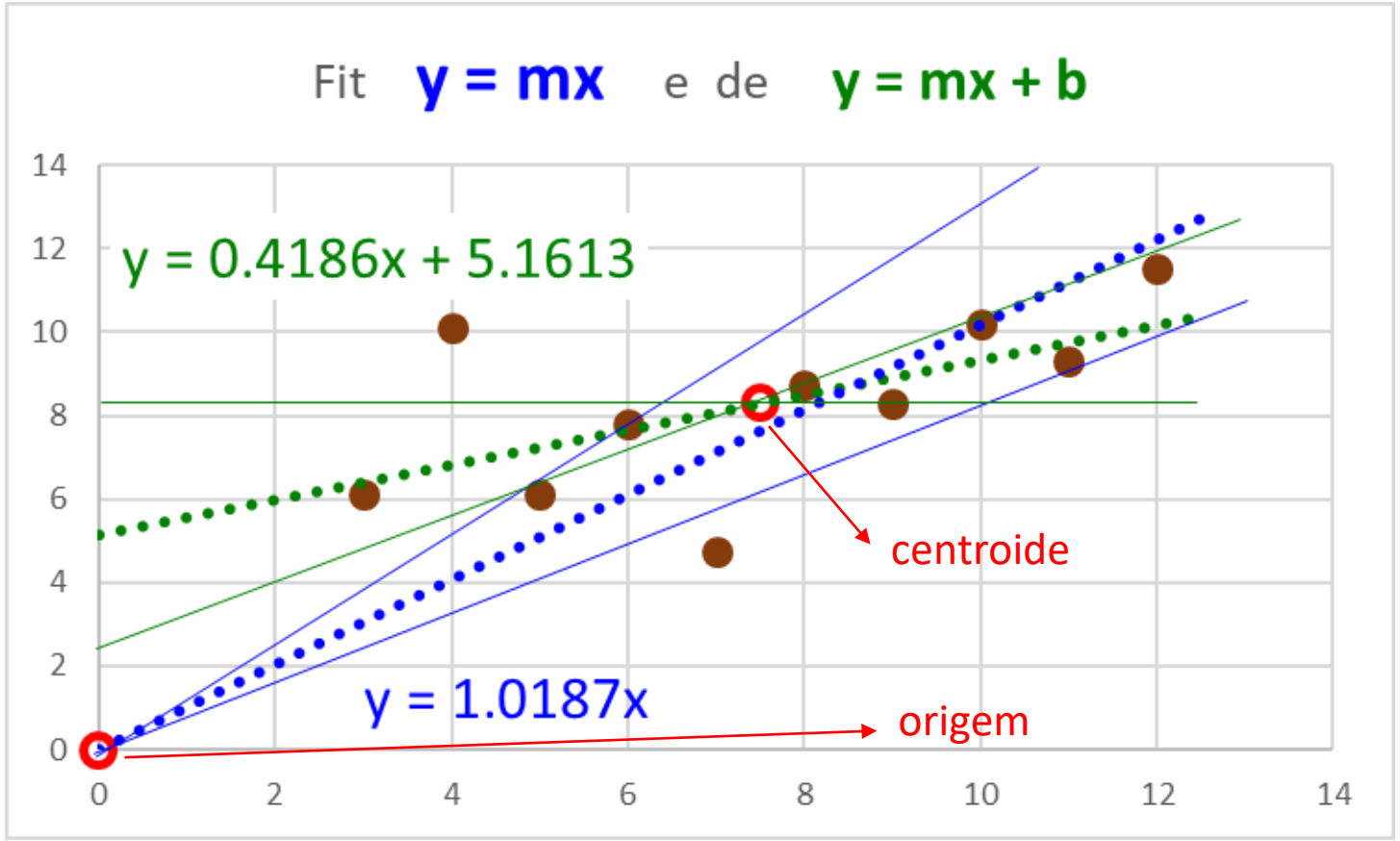
N	mm
F	**vidro**
0.00	0
0.98	0.088
1.96	0.17
2.94	0.255
3.92	0.339
4.90	0.418
5.88	0.493
4.90	0.42
3.92	0.341
2.94	0.26
1.96	0.175
0.98	0.095
0.00	0.012

$y = ax + b$ versus $y = ax$

Os pontos a castanho foram simulados (tabela e gráfico)

- As retas $y = ax$ passam pela origem. Quanto mais distantes da origem, mais influência os pontos experimentais têm na reta ajustada.
- As retas $y = ax + b$ passam pelo centroide dos pontos experimentais. Quanto mais distantes do centroide, mais influência os pontos têm na reta ajustada.

x	y
3	6.111868
4	10.12389
5	6.106789
6	7.810005
7	4.760439
8	8.732711
9	8.313216
10	10.20147
11	9.319167
12	11.52821



$y = ax + b$: extrapolação e incerteza na ordenada

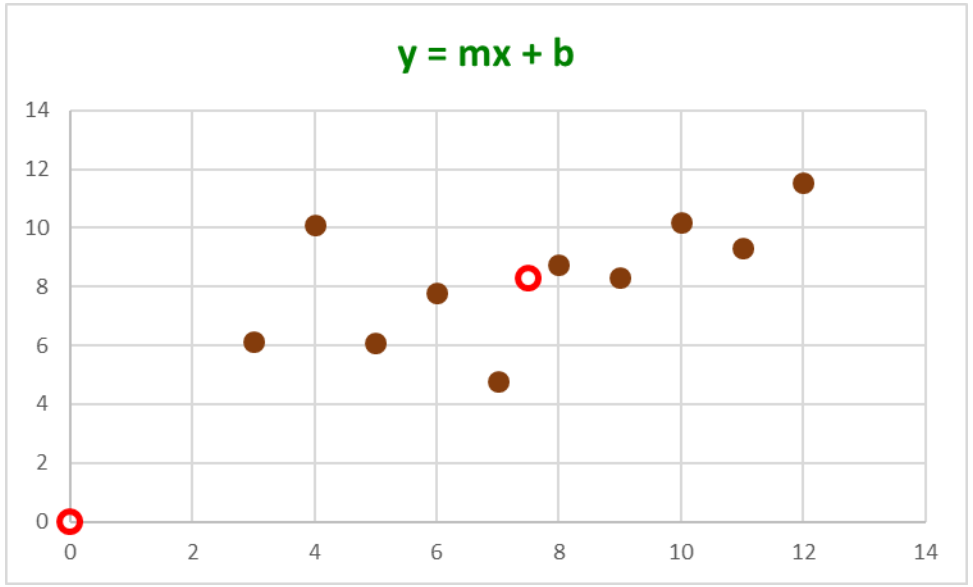
Suponhamos que pretendemos estimar o valor de *y* para um *x* que não medimos usando o nosso ajuste. Qual a incerteza no *y* estimado?

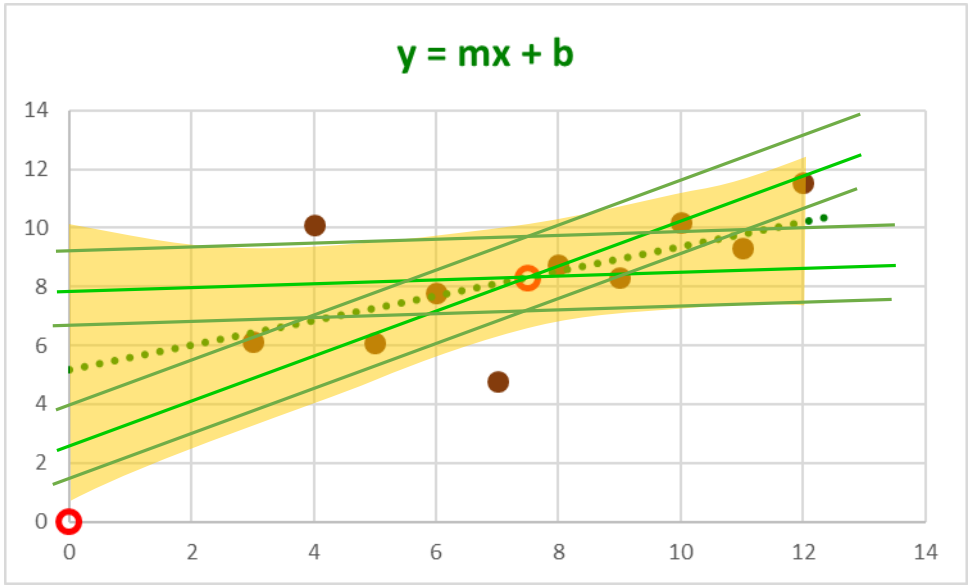
Correto ☺

$$y = a \cdot x + b \rightarrow \sigma_y = \sqrt{\left(\frac{\partial y}{\partial a} \cdot \sigma_a\right)^2 + \left(\frac{\partial y}{\partial b} \cdot \sigma_b\right)^2} = \sqrt{(x \cdot \sigma_a)^2 + \sigma_b^2}$$ → **Errado !!!**

$$y = a \cdot x + b \rightarrow \sigma_y = \sqrt{(x \cdot \sigma_a)^2 + \sigma_b^2 + 2 \cdot x \cdot cov(a,b)} \rightarrow \sigma_y = \sqrt{[(x - \bar{x}) \cdot \sigma_a]^2 + \frac{\sigma_y^2}{N}}$$

A covariância é nula no centroide (mais precisamente para uma mudança de referencial $x' = x - \bar{x}$). Nesse caso, $\sigma_b = {}^{\sigma_y}/_{\sqrt{N}}$. Notem que a incerteza padrão no declive não altera quando se muda o referencial.





Fit $y = ax + b$ e $y = ax$ usando o Excel

O seguinte site explica uma função do Excel que nos fornece muitos destes dados já calculados:
https://pages.mtu.edu/~fmorriso/cm3215/UncertaintySlopeInterceptOfLeastSquaresFit.pdf
Esta função calcula os parâmetros mostrados abaixo (caso $y = ax + b$).
Para as aplicações em física, os parâmetros mais importantes são

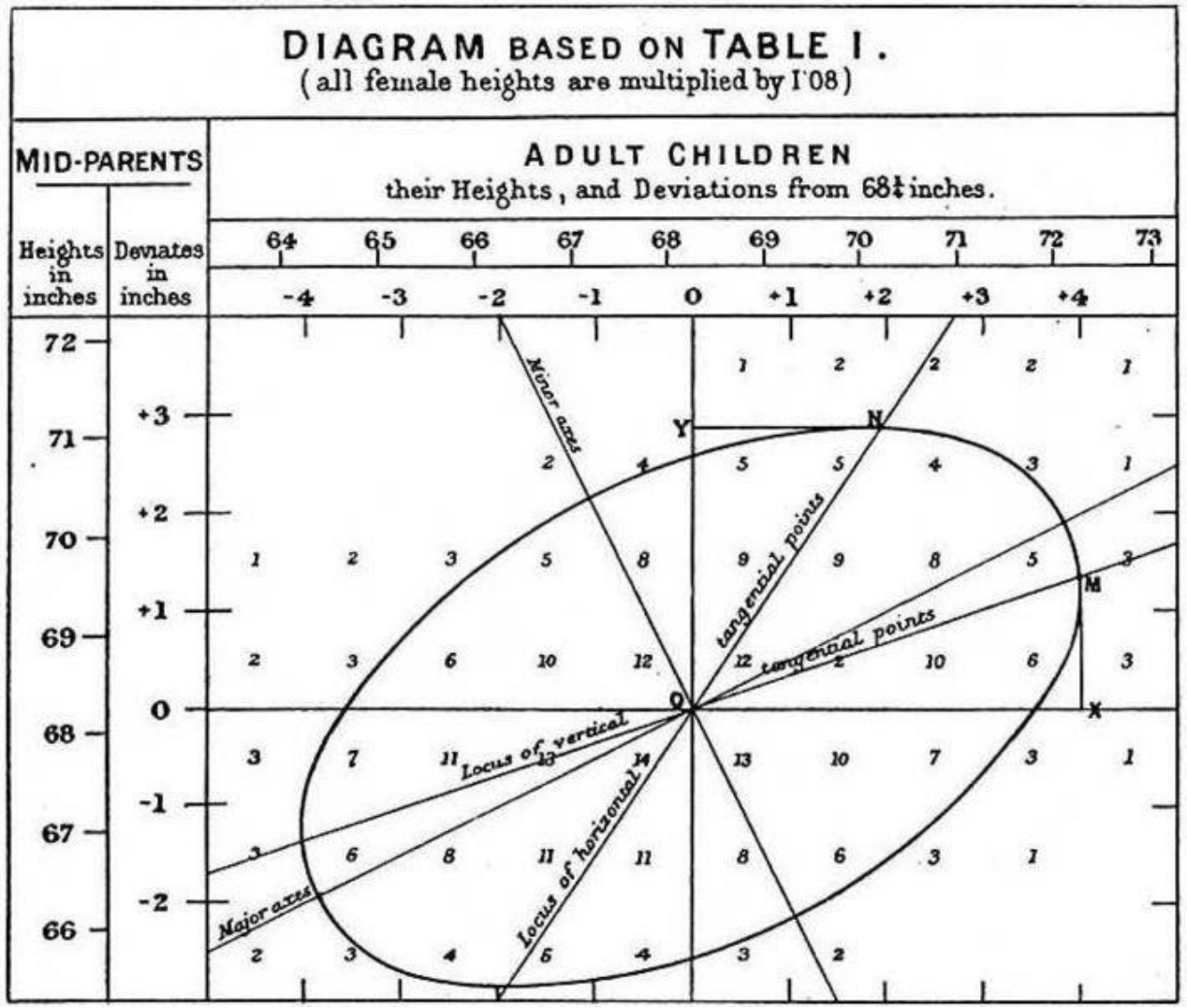
- O declive
- O desvio padrão do declive
- A ordenada na origem
- O desvio padrão na ordenada na origem
- O desvio padrão estimado dos *y*

A	B	C	D
slope, m	0.043931479	22.76492191	intercept, b
std dev of slope	0.000131455	0.111823909	std dev of intercept
$R^2=SS_R/SS_T$	0.997748762	0.153614504	$s_{y,x}$, std dev of y
Fisher F statistic, $SS_R(n-2)/SS_E$	111686.4212	252	degrees of freedom, n-2
regression ss (SS_R); explained variation	2635.510932	5.946548808	errors ss (SS_E); unexplained variation

Regressão linear – curiosidade histórica

O termo “regressão” terá sido usado no século XIX por Francis Galton no contexto de um estudo sobre a altura dos filhos relativamente à dos pais. Ele verificou que pais altos tendem a ter filhos altos, mesmo assim mais baixos do que eles próprios. Uma situação simétrica se passava com pais baixos. A essa tendência dos filhos terem alturas que se aproximavam mais da altura média da população ele chamou “regressão para a média”. No estudo usou um método que hoje se chama “regressão linear”, nome que, aparentemente, se deve a ele.

(https://en.wikipedia.org/wiki/Francis_Galton#Correlation_and_regression)



Método dos mínimos quadrados ou *regressão linear* ?

O termo *método dos mínimos quadrados* é mais geral e refere-se a uma técnica que pode ser aplicada ao ajuste de retas ou de outros tipos de funções.

Regressão linear está limitado a retas, mas há autores que fazem uma outra distinção:

★6.2.6 Regression

The data on the left in Figure 6.2 are measurements of the temperature and pressure of a gas at constant volume. The line through the points is the

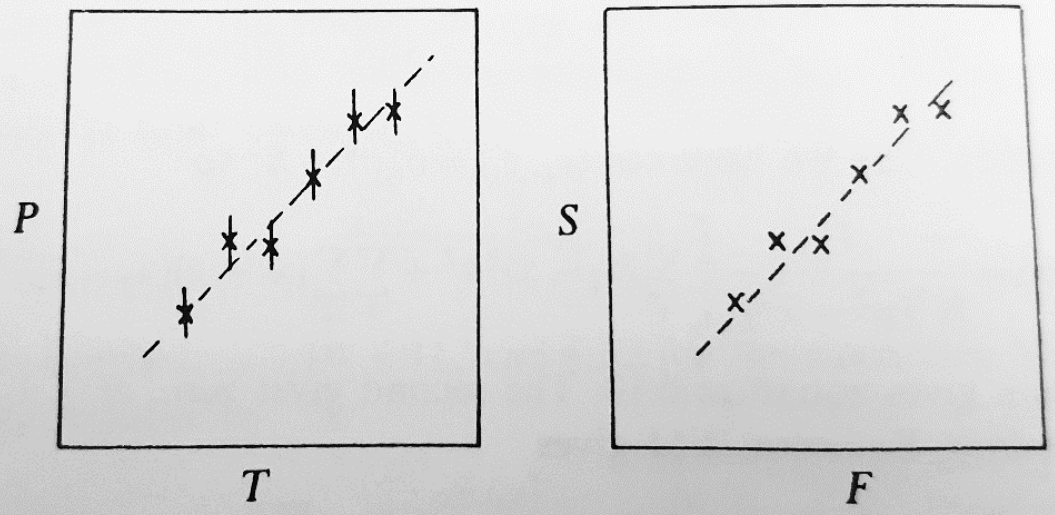


Fig. 6.2. Two sets of data and two straight lines.

Método dos mínimos quadrados ou *regressão linear* ?

'straight line fit'. The data on the right are measurements of the heights of some fathers and their (adult) sons: the line through the points is the 'line of regression'. What is the difference? Mathematically there is none—both lines were evaluated by equations 6.6 to 6.9. However, there is a profound difference in their meaning.

Notice how the left-hand plot has error bars on its measurements and the right-hand one has not. The real difference between the two plots lies here. If P had been measured with greater precision, using a more accurate barometer, then the measured points would move closer to their true relationship. (For an ideal gas this would be a straight line, and for a real gas should be reasonably close to a straight line). If the heights were remeasured using, say, a laser interferometer instead of an ordinary ruler, then there would be no visible effect on the measured points shown.

The true values of pressure and temperature are believed to be related by an exact law, and the line drawn through those points is an estimate of that law. The heights of father and son are related by a trend—if you know that John Smith, Snr, is 6′6″ tall, then you might expect John Smith, Jnr, to be on the tall side, but you could not predict his height very accurately. Thus, although there are similarities, regression is a part of descriptive statistics, like correlation, whereas straight line fitting is a form of estimation.

The name 'regression', which is a bit misleading, came from Galton's original work on just this problem. He found that tall fathers did tend to have tall sons, but that as the correlation is not perfect, a tall father will tend to have a son shorter than himself. He called this 'regression towards the norm', and the name stuck.

O termo *regressão linear* é mais usado na estatística descritiva, o termo *método dos mínimos quadrados* é mais usado quando as variáveis seguem leis fundamentais.

Notem a referência à correlação. Embora seja tradicional acrescentar o parâmetro R^2 na caraterização dos ajustes lineares, esse parâmetro tem pouco interesse na maioria dos ajustes em física.

A utilização mais habitual é a da estatística descritiva. A maioria dos livros, sites da internet e pacotes de software estão direcionados para ela e o seu uso em física deve aplicado com cautela.

O coeficiente de correlação e o bug do Excel

O excerto do livro do slide anterior classificava o coeficiente de correlação como um parâmetro descritivo. Na verdade, é muito raro ele ser usado numericamente, o uso costuma ser mais qualitativo.

O Excel tinham um bug e calculava mal alguns parâmetros da regressão linear. Na década de 90 (século passado) a Microsoft descreveu o problema e corrigiu as fórmulas… bem, as usadas na folha de cálculo. Uma, usada no gráfico ficou "esquecida" até hoje:

O R^2 é um número real ao quadrado e deve ser sempre positivo.

Contudo, com o exemplo sugerido pela própria Microsoft à direita, o valor de R^2 calculado no gráfico está errado.

E ninguém dá por nada !!!

x	y
1	11
2	12
3	13

5.28571	0
1.23718	#N/A
R^2 = 0.90125	4.6291
18.2533	2
391.143	42.8571