Física Experimental IV

https://edisciplinas.usp.br/course/view.php?id=90535

2º Semestre 2021

Revisão

Método dos Mínimos Quadrados

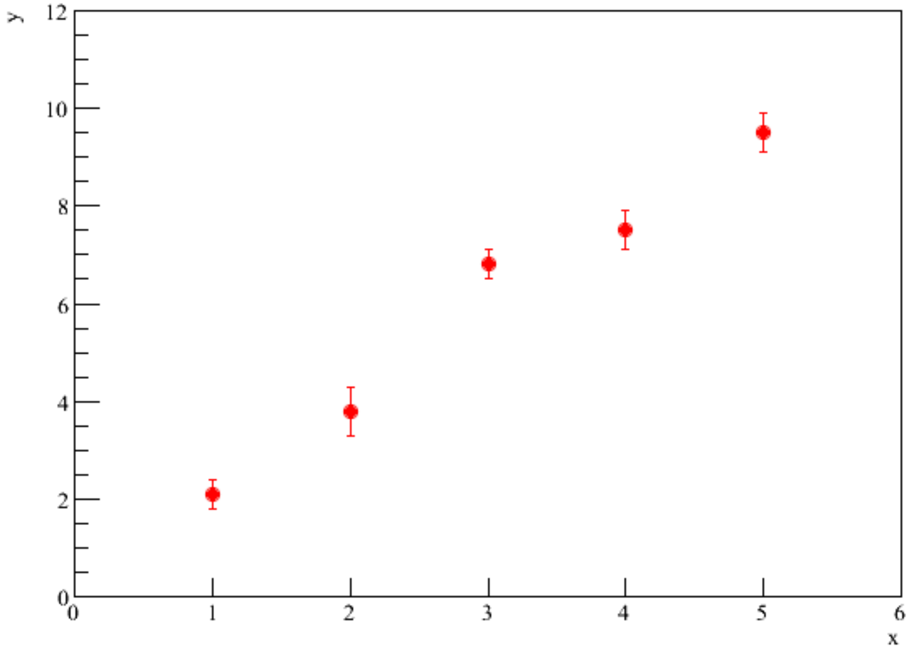
Prof. Henrique Barbosa

hbarbosa@if.usp.br

http://www.fap.if.usp.br/~hbarbosa

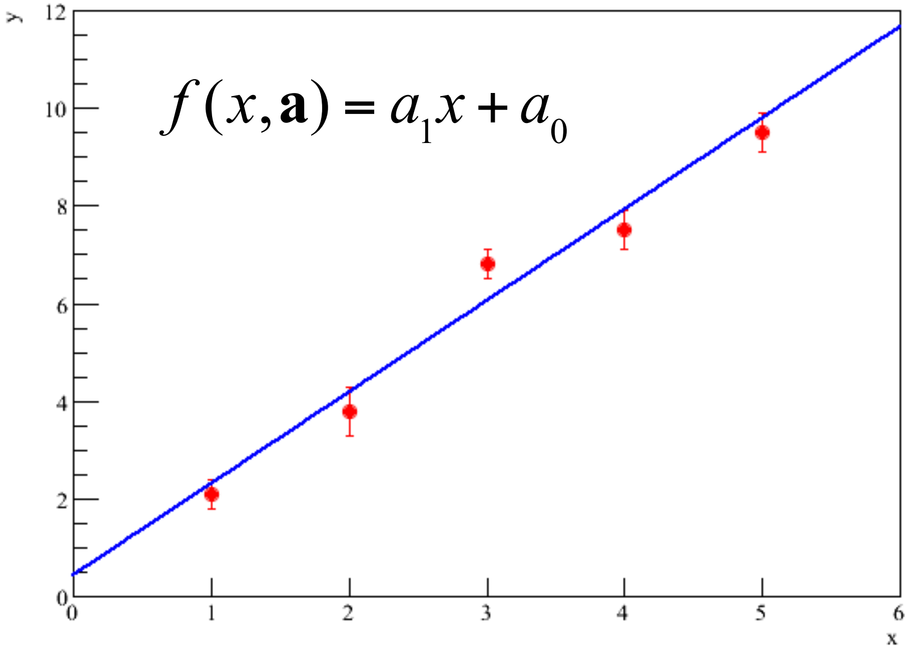
Problema geral

- Conjunto de dados $\{x,y,\sigma\}$
- Há uma relação entre x e y
- Como encontramos a função $y=f(x,\mathbf{a})$ que <u>melhor</u> descreve (estatisticamente) os dados?



Parâmetros?

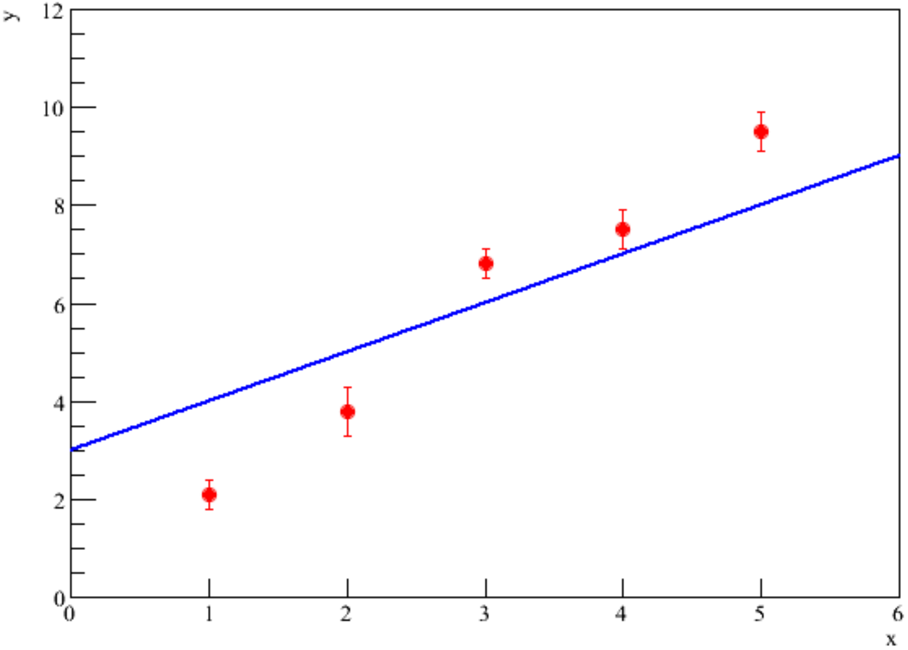
- Conjunto de dados $\{x,y,\sigma\}$
- Há uma relação entre x e y
- Como encontramos a função $y=f(x,\mathbf{a})$ que <u>melhor</u> descreve (estatisticamente) os dados?
- Como definimos (estatisticamente) os parâmetros $\mathbf{a}=(a_0,\ldots,a_k)$?



Máxima verossimilhança

- Não servem parâmetros **a**=(a_0,...,a_k) quaisquer... Então como fazemos?

- **MMQ:** encontra coeficientes que maximizam a probabilidade da função ser compatível com os dados



Probabilidades e Hipótese

- Dados independentes
  - $\{x_1,y_1,s_1\}, \ldots, \{x_n,y_n,s_n\}$
- Medidas y_i:
  - gaussiana com média μ_i e desvio padrão σ_i
- A probabilidade de medirmos y_i é

$$P(y_i) \propto \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[-\frac{1}{2}\left(\frac{y_i - \mu(x_i)}{\sigma_i}\right)^2\right]$$

A função $f(x_i,\mathbf{a})$ é a nossa estimativa dos valores verdadeiros

Probabilidades dos dados

- Portanto, a probabilidade para uma medida é:

$$P_i \propto \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[-\frac{1}{2}\left(\frac{y_i - f(x_i, \mathbf{a})}{\sigma_i}\right)^2\right]$$

- E probabilidade de medirmos o nosso conjunto de dados:

$$P_{dados} = P_1 P_2 \ldots P_n = \prod_{i=1}^{n} P_i$$

Probabilidades dos dados

- Portanto, a probabilidade total é:

$$P_{dados} \propto \frac{1}{\left(\sqrt{2\pi}\right)^n \sigma_1 \ldots \sigma_n} \exp\left[\underbrace{-\frac{1}{2}\overbrace{\sum_{i=1}^{n}\left(\frac{y_i - f(x_i,\mathbf{a})}{\sigma_i}\right)^2}^{\chi^2 > 0}}\right]$$

- A probabilidade é máxima quando o X^2 for mínimo!
- IMPORTANTE: X^2 é uma função de $\mathbf{a}=(a_0,\ldots,a_k)$

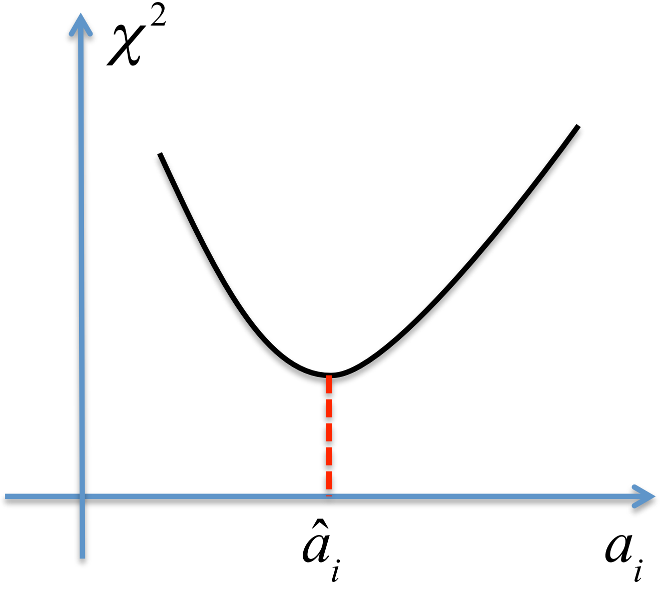
Ajuste

- Consiste em encontrar os parâmetros $\mathbf{a}=(a_0,\ldots,a_k)$, da função $f(x,\mathbf{a})$, de modo a minimizar o X^2.

- Hipóteses:
  - As medidas y_i tem uma distribuição gaussiana
  - As medidas são independentes

Ajuste

- Dado $f(x,\mathbf{a})$, o que significa minimizar o X^2?
- Significa que qualquer variação de qualquer a_i faz o X^2 aumentar!
- Portanto: $$\frac{\partial \chi^2}{\partial a_i} = 0$$



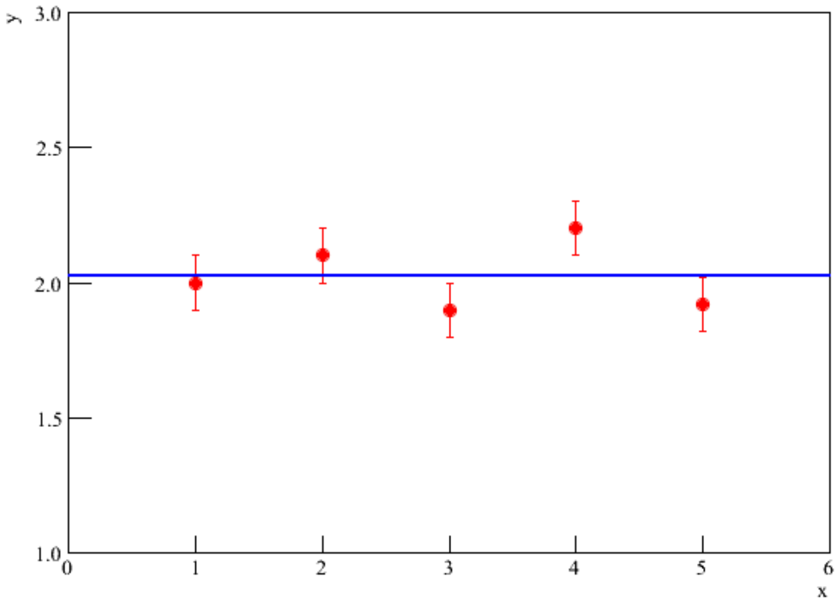
Constante

- O chi2 é:

$$\chi^2 = \sum_{i=1}^{n} \left(\frac{y_i - a_0}{\sigma_i} \right)^2$$

- E precisamos calcular:

$$\frac{\partial}{\partial a_0} \chi^2$$



Constante

$$\frac{\partial}{\partial a_0}\chi^2 = \frac{\partial}{\partial a_0}\sum_{i=1}^{n}\frac{\left[y_i - a_0\right]^2}{\sigma_i^2} = \sum_{i=1}^{n}\frac{2\left[y_i - a_0\right](-1)}{\sigma_i^2}$$

$$0 = -2\sum_{i=1}^{n}\frac{y_i}{\sigma_i^2} + 2a_0\sum_{i=1}^{n}\frac{1}{\sigma_i^2}$$

Se $\sigma_i = \sigma =$ constante

$$\sum_{i=1}^{n}\frac{y_i}{\sigma_i^2} = a_0\sum_{i=1}^{n}\frac{1}{\sigma_i^2}$$

$$\frac{S_y}{S_\sigma} = a_0$$

$$= \frac{\Sigma y_i / \sigma^2}{\Sigma 1 / \sigma^2} = \frac{1}{n}\Sigma y_i$$

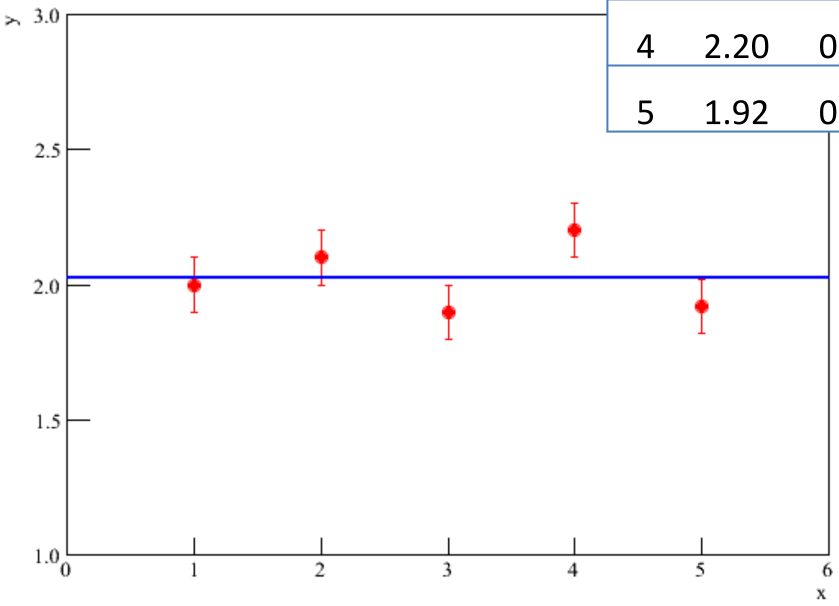
Exemplo

$$S_\sigma = \sum_{i=1}^{n} \frac{1}{\sigma_i^2} = \sum_{i=1}^{5} \frac{1}{0.1^2} = \frac{5}{0.01} = 500$$

$$S_y = \sum_{i=1}^{n} \frac{y_i}{\sigma_i^2} = \frac{10.12}{0.01} = 1012$$

$$a_0 = \frac{S_y}{S_\sigma} = \frac{10.12/0.01}{5/0.01} = 2.024$$

X	Y	s
1	2.00	0.1
2	2.10	0.1
3	1.90	0.1
4	2.20	0.1
5	1.92	0.1



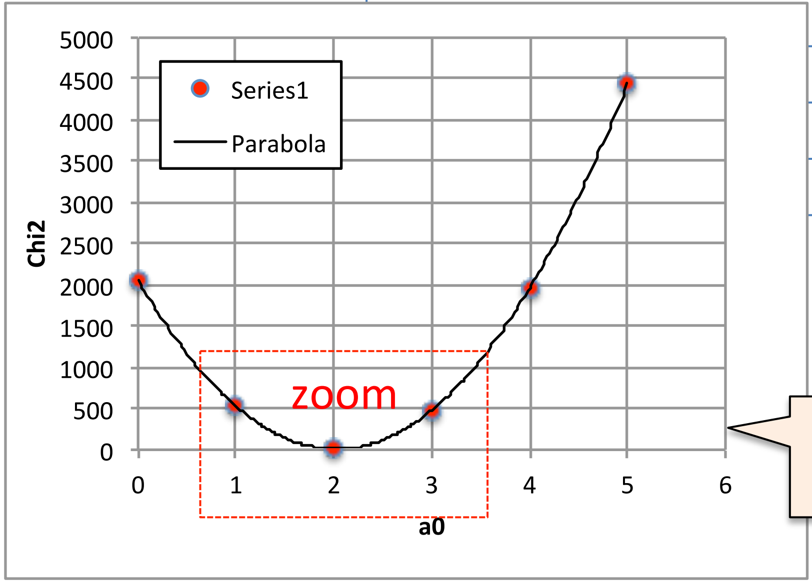
$$\chi^2 = \sum_{i=1}^{n} \left(\frac{y_i - a_0}{\sigma_i} \right)^2$$

Graficamente

Tentamos diversos valores para o parâmetro

X	Y	sigma	a0=	0	1	2	3	4	5
1.00	2.00	0.1	(yi-Fi)^2/si^2=	400	100	0.0	100	400	900
2.00	2.10	0.1		441	121	1.0	81	361	841
				361	81	1.0	121	441	961
				484	144	4.0	64	324	784
				369	85	0.6	117	433	949
			chi2=	2055	531	6.6	483	1959	4435

Para cada tentativa, calculamos o chi2

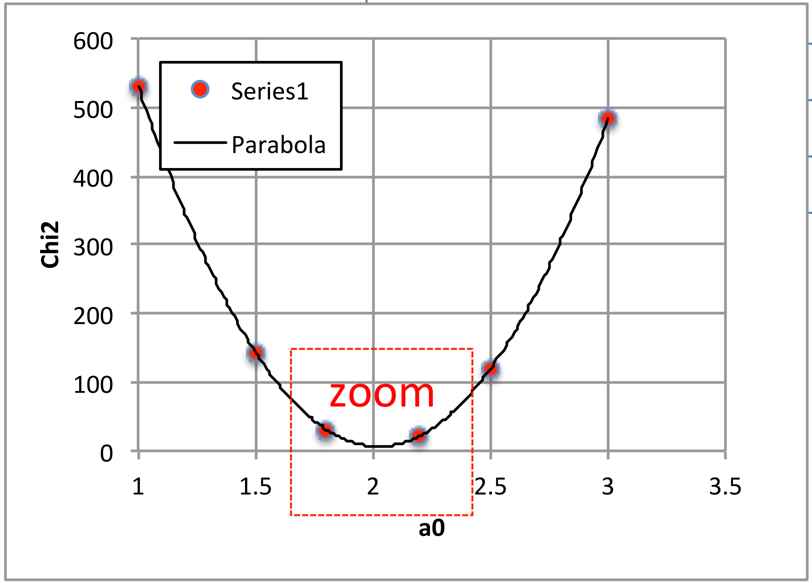


Fazemos um gráfico, que é uma parábola

Graficamente

Zoom na região onde temos o mínimo

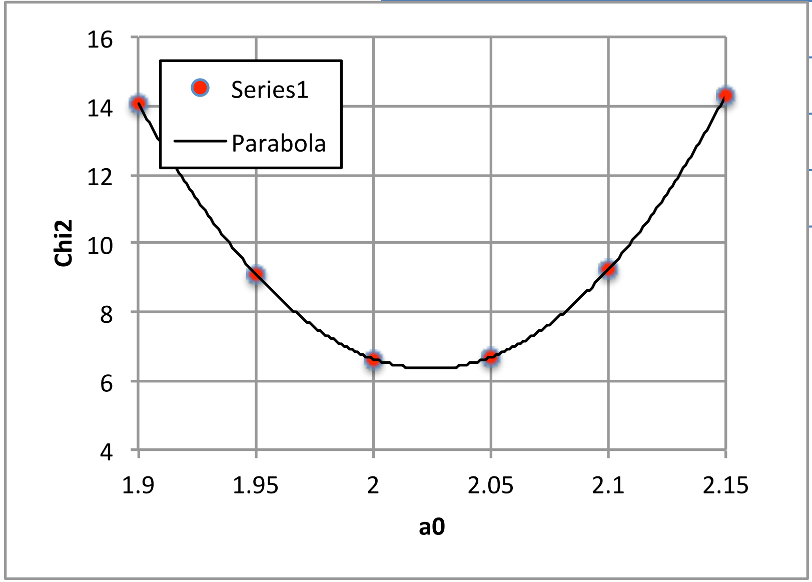
X	Y	sigma	a0=	1	1.5	1.8	2.2	2.5	3
1.00	2.00	0.1	(yi-Fi)^2/si^2=	100	25	4.0	4	25	100
2.00	2.10	0.1		121	36	9.0	1	16	81
				81	16	1.0	9	36	121
				144	49	16.0	0	9	64
				85	18	1.4	8	34	117
			chi2=	531	144	31.4	22	120	483



Graficamente

Zoom na região onde temos o mínimo

X	Y	sigma	a0=	1.6	1.7	1.9	2	2.2	2.4
1.00	2.00	0.1	(yi-Fi)^2/si^2=	16	9	1.0	0	4	16
2.00	2.10	0.1		25	16	4.0	1	1	9
				9	4	0.0	1	9	25
				36	25	9.0	4	0	4
				10	5	0.0	1	8	23
			chi2=	96	59	14.0	7	22	77



Repete o processo até que a redução do Chi2, ou variação do parâmetro, seja menor que um certo limiar (ex. 1e-6)

Ajuste de Reta

- O chi2 é: $$\chi^2 = \sum_{i=1}^{n} \left(\frac{y_i - (a_1 x_i + a_0)}{\sigma_i} \right)^2$$

- E precisamos calcular: $$\frac{\partial}{\partial a_1}\chi^2 \qquad \frac{\partial}{\partial a_0}\chi^2$$

Ajuste de Reta, a_0

$$\frac{\partial}{\partial a_0}\chi^2 = \frac{\partial}{\partial a_0}\sum_{i=1}^{n}\frac{\left[y_i-(a_1x_i+a_0)\right]^2}{\sigma_i^2} = \sum_{i=1}^{n}\frac{2\left[y_i-(a_1x_i+a_0)\right](-1)}{\sigma_i^2}$$

$$0 = -2\sum_{i=1}^{n}\frac{y_i}{\sigma_i^2} + 2a_1\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2} + 2a_0\sum_{i=1}^{n}\frac{1}{\sigma_i^2}$$

$$\sum_{i=1}^{n}\frac{y_i}{\sigma_i^2} = a_1\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2} + a_0\sum_{i=1}^{n}\frac{1}{\sigma_i^2}$$

$$S_y = a_1S_x + a_0S_\sigma$$

Ajuste de Reta, a_1

$$\frac{\partial}{\partial a_1}\chi^2 = \frac{\partial}{\partial a_1}\sum_{i=1}^{n}\frac{\left[y_i-(a_1x_i+a_0)\right]^2}{\sigma_i^2} = \sum_{i=1}^{n}\frac{2\left[y_i-(a_1x_i+a_0)\right](-x_i)}{\sigma_i^2}$$

$$0 = -2\sum_{i=1}^{n}\frac{y_ix_i}{\sigma_i^2} + 2a_1\sum_{i=1}^{n}\frac{x_i^2}{\sigma_i^2} + 2a_0\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}$$

$$\sum_{i=1}^{n}\frac{y_ix_i}{\sigma_i^2} = a_1\sum_{i=1}^{n}\frac{x_i^2}{\sigma_i^2} + a_0\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}$$

$$S_{xy} = a_1S_{x^2} + a_0S_x$$

Resolvendo o sistema

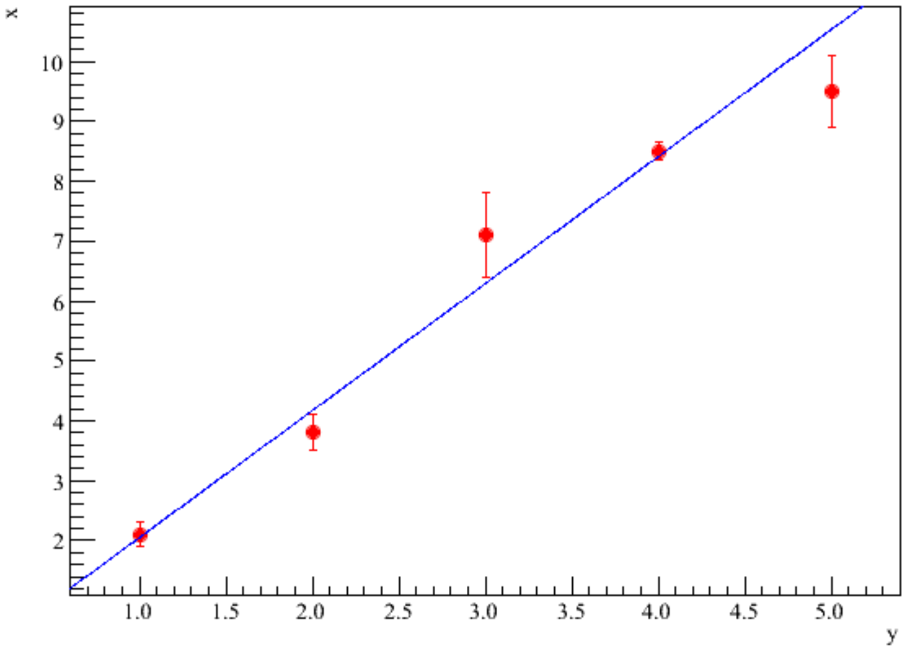
$$\begin{cases} S_y = a_1 S_x + a_0 S_\sigma \\ S_{xy} = a_1 S_{x^2} + a_0 S_x \end{cases} \Rightarrow \begin{pmatrix} S_y \\ S_{xy} \end{pmatrix} = \begin{pmatrix} S_\sigma & S_x \\ S_x & S_{x^2} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix}$$

Podemos resolver explicitamente:

$$\boxed{\frac{S_{x^2} S_y - S_{xy} S_x}{S_{x^2} S_\sigma - S_x S_x} = a_0} \qquad \boxed{\frac{S_{xy} S_\sigma - S_y S_x}{S_{x^2} S_\sigma - S_x S_x} = a_1}$$

Exemplo

X	Y	sigma
1	2.2	0.12
2	3.8	0.29
3	7.1	0.69
4	8.5	0.17
5	9.5	0.58

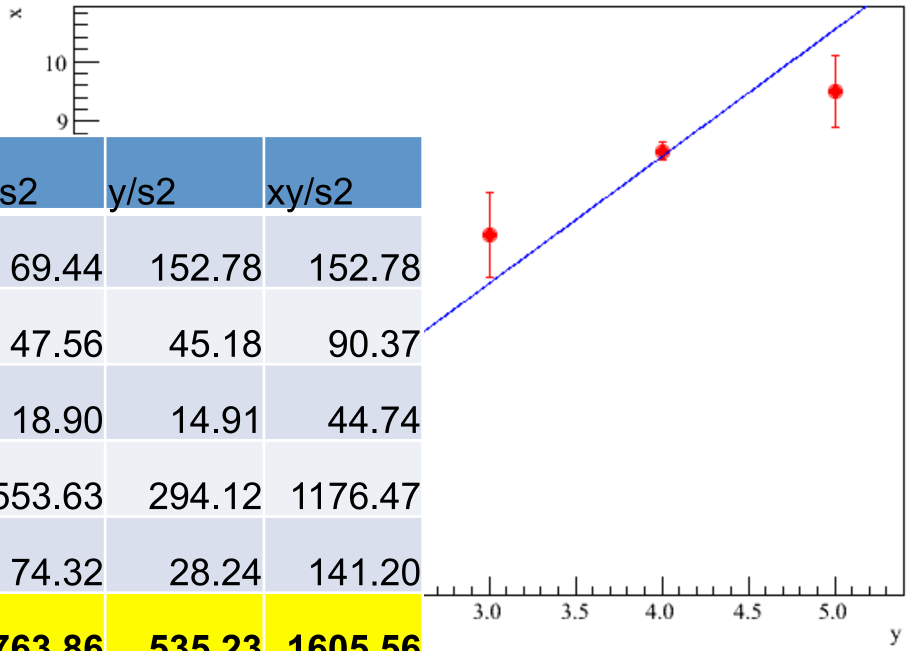


Exemplo

$$a_0 = 0.104$$

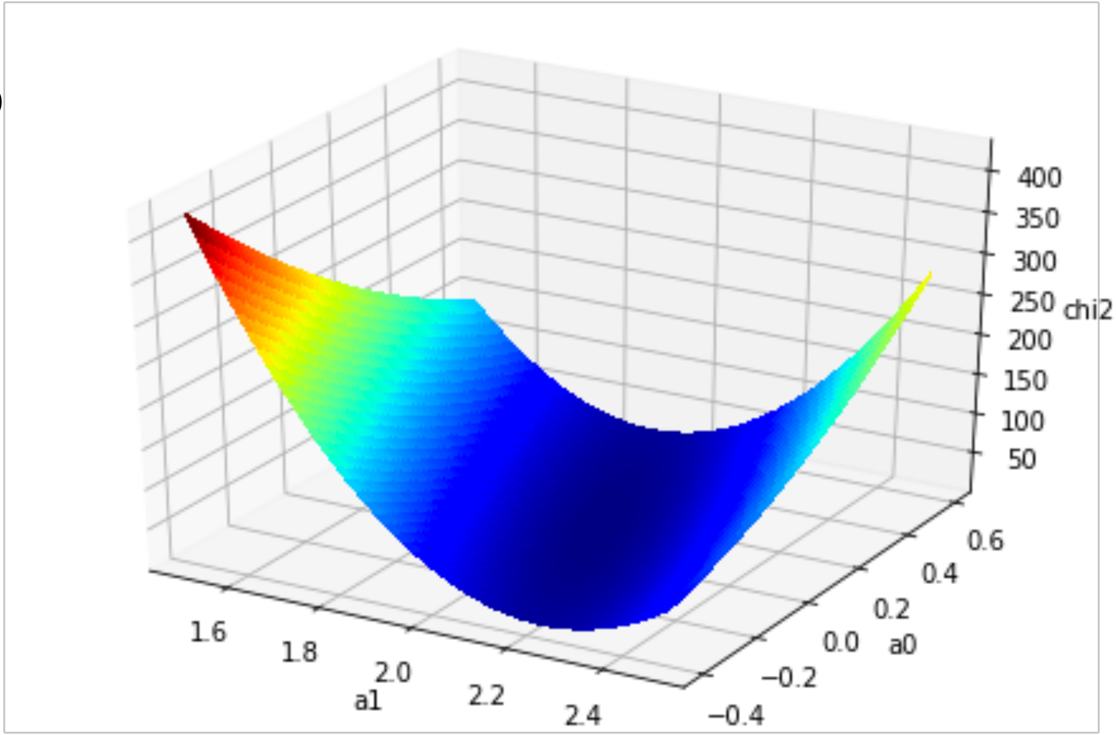
$$a_1 = 2.068$$

X	Y	sigma	1/s2	x/s2	x2/s2	y/s2	xy/s2
1	2.2	0.12	69.44	69.44	69.44	152.78	152.78
2	3.8	0.29	11.89	23.78	47.56	45.18	90.37
3	7.1	0.69	2.10	6.30	18.90	14.91	44.74
4	8.5	0.17	34.60	138.41	553.63	294.12	1176.47
5	9.5	0.58	2.97	14.86	74.32	28.24	141.20
			121.01	**252.80**	**763.86**	**535.23**	**1605.56**
			Ssig	**Sx**	**Sx2**	**Sy**	**Sxy**



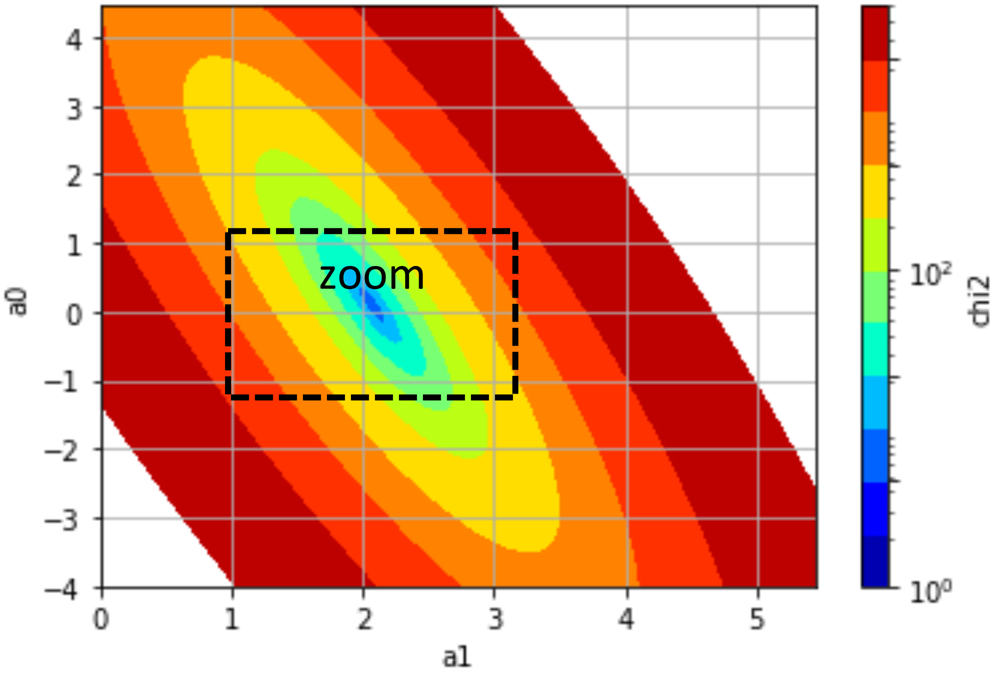
Graficamente

- Agora temos 2 parâmetros, portanto gráfico de chi2 será uma superfície



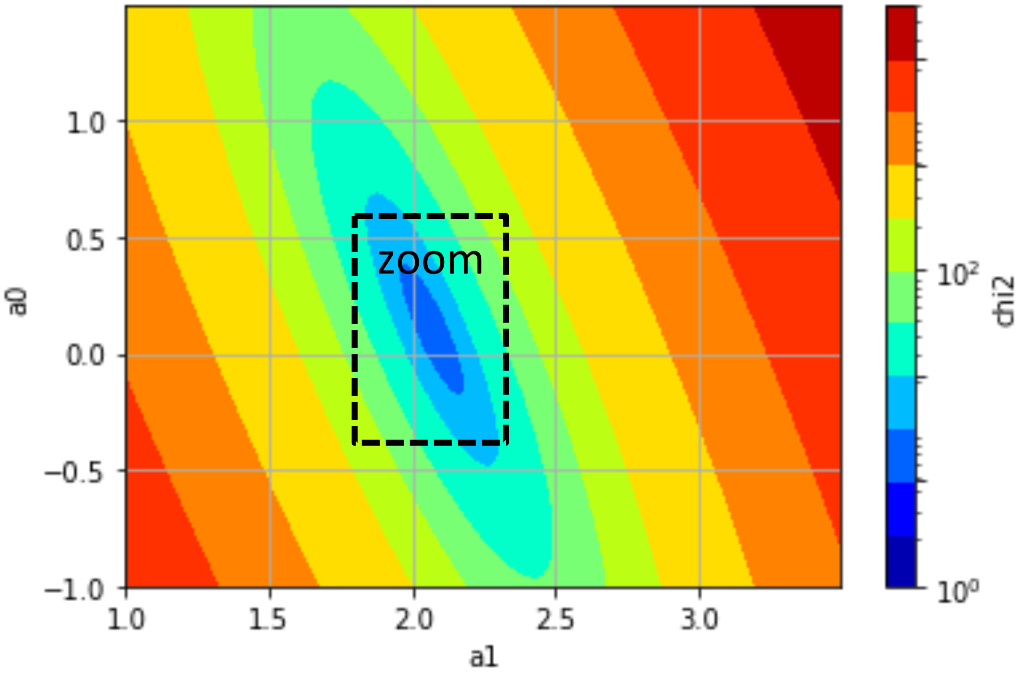
Graficamente

- Podemos também fazer um gráfico de contornos



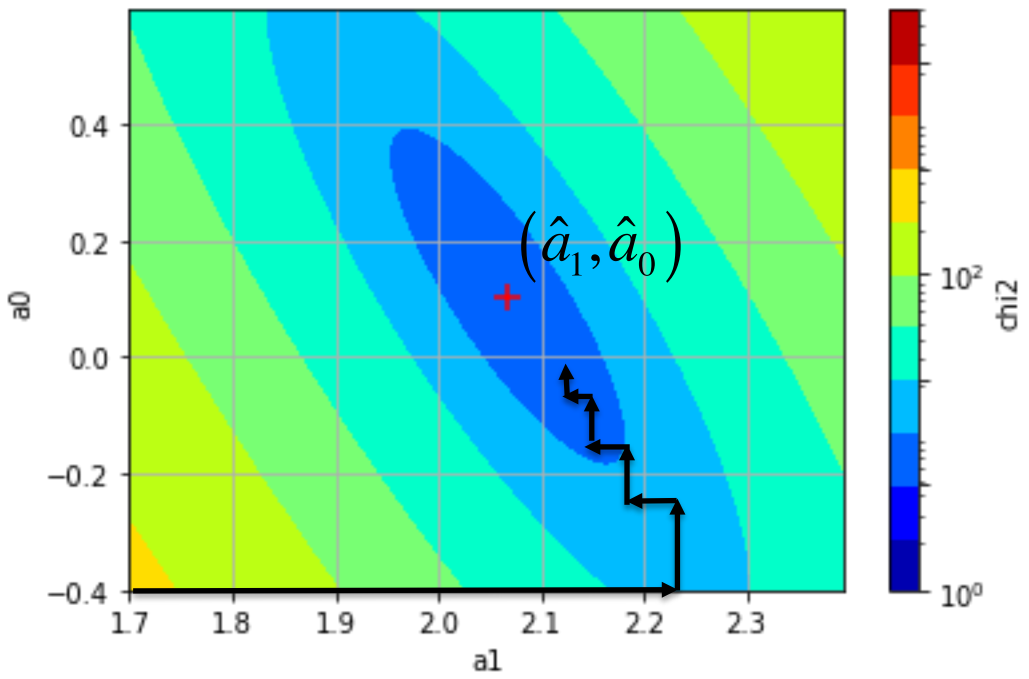
Graficamente

- A região ampliada se parece com a região anterior



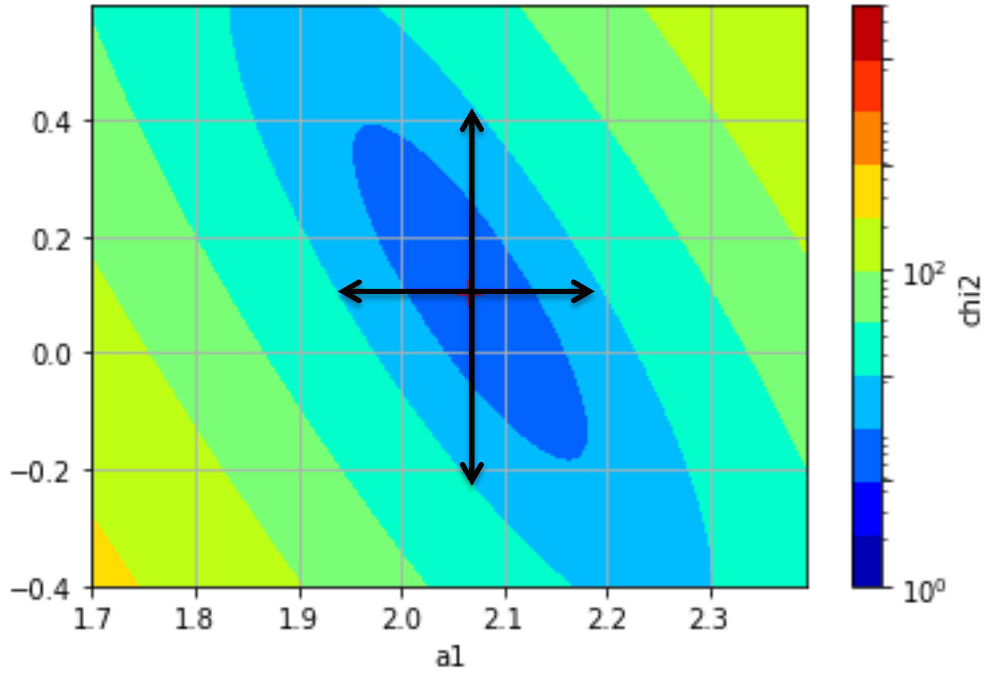
Numericamente

- Podemos procurar, alternadamente (em x,y), pelo mínimo, até convergir.

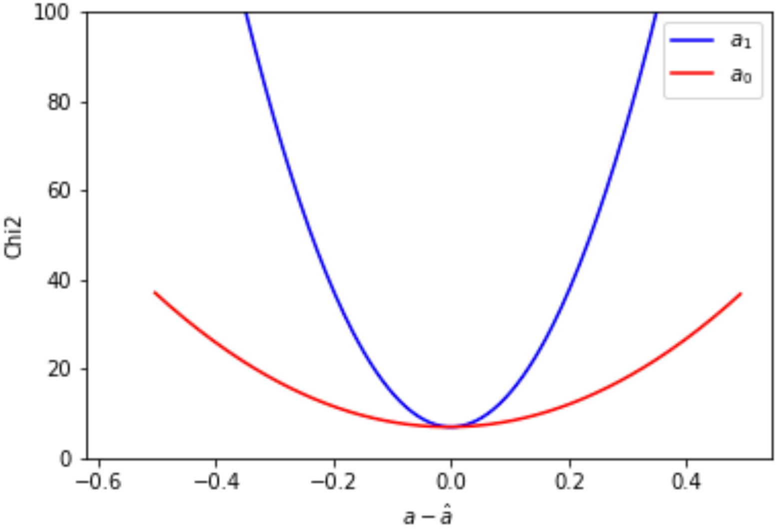


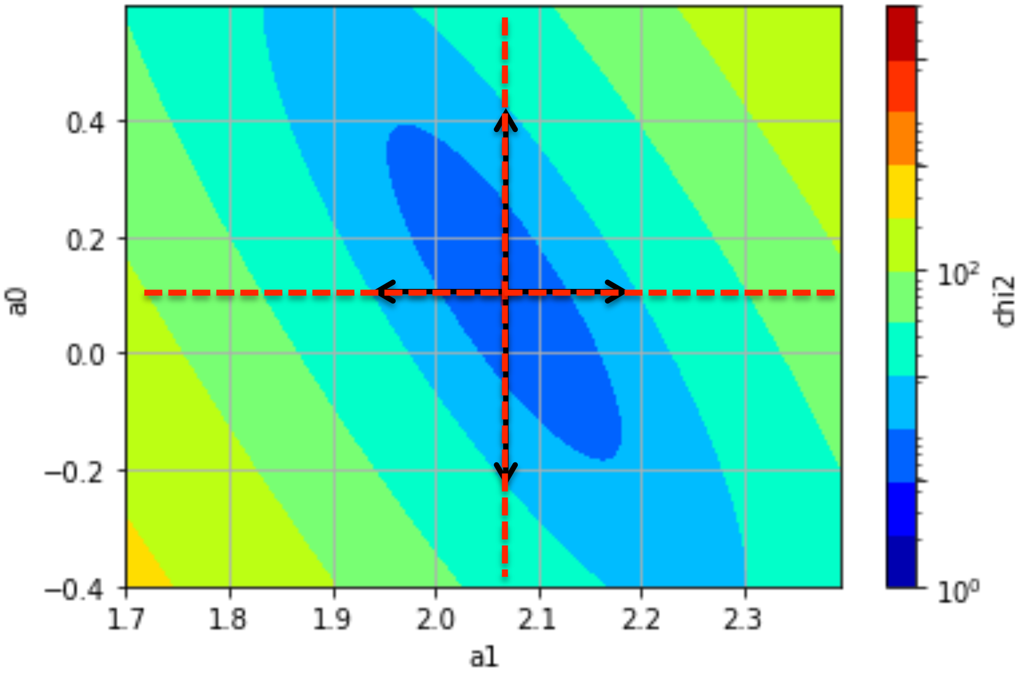
Incerteza dos parâmetros

- **a**=(a_0,...,a_k) é estimado dos dados e é afetado pelas flutuações nos valores y_i
- Tem uma incerteza associada a estimativa de **a**=(a_0,...,a_k)
    - Depende do comportamento da função X^2 em torno do mínimo



Incerteza dos parâmetros





A incerteza em a_1 será menor que em a_0

Propagação

- Seja uma função $g(y_j)$, a incerteza devido aos erros em y_j é

$$\sigma_g^2 = \sum_{j=1}^{n} \left(\frac{\partial g}{\partial y_j} \sigma_j \right)^2$$

- É a propagação de incerteza que estamos acostumados!
- Precisamos aplicá-la para os parâmetros estimados...

Ajuste linear, incertezas

$$a_0 = \frac{S_{x^2}S_y - S_{xy}S_x}{S_{x^2}S_\sigma - S_xS_x}$$

$$a_1 = \frac{S_{xy}S_\sigma - S_yS_x}{S_{x^2}S_\sigma - S_xS_x}$$

$$\left.\begin{aligned} S_\sigma &= \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \\ S_x &= \sum_{i=1}^{n} \frac{x_i}{\sigma_i^2} \\ S_{x^2} &= \sum_{i=1}^{n} \frac{x_i^2}{\sigma_i^2} \end{aligned}\right\}$$

Não dependem de y_j, por isso

$$\frac{\partial}{\partial y_j} \Rightarrow 0$$

$$\frac{\partial S_y}{\partial y_j} = \frac{\partial}{\partial y_j} \sum_{i=1}^{n} \frac{y_i}{\sigma_i^2} = \frac{1}{\sigma_j^2}$$

$$\frac{\partial S_{xy}}{\partial y_j} = \frac{\partial}{\partial y_j} \sum_{i=1}^{n} \frac{y_i x_i}{\sigma_i^2} = \frac{x_j}{\sigma_j^2}$$

Ajuste linear, incertezas

$$a_0 = \frac{S_{x^2} S_y - S_{xy} S_x}{S_{x^2} S_\sigma - S_x S_x}$$

$$a_1 = \frac{S_{xy} S_\sigma - S_y S_x}{S_{x^2} S_\sigma - S_x S_x}$$

$$\sigma_{a0}^2 = \sum_{j=1}^{n} \left(\frac{\partial a_0}{\partial y_j} \sigma_j \right)^2$$

$$\sigma_{a1}^2 = \sum_{j=1}^{n} \left(\frac{\partial a_1}{\partial y_j} \sigma_j \right)^2$$

Fazendo a conta, chegamos em:

$$\sigma_{a_0}^2 = \frac{S_{x^2}}{S_{x^2} S_\sigma - S_x S_x}$$

$$\sigma_{a_1}^2 = \frac{S_\sigma}{S_{x^2} S_\sigma - S_x S_x}$$

A conta para a_0...

$$\boxed{a_0 = \frac{S_{x^2}S_y - S_{xy}S_x}{S_{x^2}S_\sigma - S_xS_x}} \qquad \frac{\partial a_0}{\partial y_j} = \frac{S_{x^2}\frac{\partial S_y}{\partial y_j} - \frac{\partial S_{xy}}{\partial y_j}S_x}{S_{x^2}S_\sigma - S_xS_x} = \frac{S_{x^2}\frac{1}{\sigma_j^2} - \frac{x_j}{\sigma_j^2}S_x}{\Delta}$$

$$\sigma_{a0}^2 = \sum_{j=1}^{n}\left(\frac{\partial a_0}{\partial y_j}\sigma_j\right)^2 = \frac{1}{\Delta^2}\sum_{j=1}^{n}\left(\left[S_{x^2}\frac{1}{\sigma_j^{\cancel{2}}} - \frac{x_j}{\sigma_j^{\cancel{2}}}S_x\right]\cancel{\sigma_j}\right)^2$$

$$= \frac{1}{\Delta^2}\left[\left(S_{x^2}\right)^2\underbrace{\sum_{j=1}^{n}\frac{1}{\sigma_j^2}}_{S_\sigma} - \cancel{2}S_{x^2}S_x\underbrace{\sum_{j=1}^{n}\frac{x_j}{\sigma_j^2}}_{S_x} + \cancel{(S_x)^2\underbrace{\sum_{j=1}^{n}\frac{x_j^2}{\sigma_j^2}}_{S_{x^2}}}\right] = \frac{S_{x^2}}{\Delta^{\cancel{2}}}\cancel{\left[S_{x^2}S_\sigma - (S_x)^2\right]}$$

$$\boxed{\sigma_{a0}^2 = \frac{S_{x^2}}{\Delta}}$$

A conta para a_1...

$$\boxed{a_1 = \frac{S_{xy}S_\sigma - S_y S_x}{S_{x^2}S_\sigma - S_x S_x}}$$

$$\frac{\partial a_1}{\partial y_j} = \frac{\frac{\partial S_{xy}}{\partial y_j}S_\sigma - \frac{\partial S_y}{\partial y_j}S_x}{S_{x^2}S_\sigma - S_x S_x} = \frac{\frac{x_j}{\sigma_j^2}S_\sigma - \frac{1}{\sigma_j^2}S_x}{\Delta}$$

$$\sigma_{a1}^2 = \sum_{j=1}^{n}\left(\frac{\partial a_1}{\partial y_j}\sigma_j\right)^2 = \frac{1}{\Delta^2}\sum_{j=1}^{n}\left(\left[\frac{x_j}{\cancel{\sigma_j^2}}S_\sigma - \frac{1}{\cancel{\sigma_j^2}}S_x\right]\cancel{\sigma_j}\right)^2$$

$$\boxed{\sigma_{a0}^2 = \frac{S_\sigma}{\Delta}}$$

$$= \frac{1}{\Delta^2}\left[(S_\sigma)^2\underbrace{\sum_{j=1}^{n}\frac{x^2}{\sigma_j^2}}_{S_{x^2}} - \cancel{2}S_\sigma S_x\underbrace{\sum_{j=1}^{n}\frac{x_j}{\sigma_j^2}}_{S_x} + \cancel{(S_x)^2\underbrace{\sum_{j=1}^{n}\frac{1}{\sigma_j^2}}_{S_\sigma}}\right] = \frac{S_\sigma}{\Delta^{\cancel{2}}}\left[\cancel{S_{x^2}S_\sigma - (S_x)^2}\right]$$

Ajuste linear, covariância

- Os parâmetros ajustados, em geral, não são independentes.
  - Reta: se mudarmos o a_0, teremos que mexer o a_1 para manter o X^2 baixo.

- No ajuste linear, a covariância é: $$\mathrm{cov}(a_0, a_1) = \frac{S_x}{S_{x^2} S_\sigma - S_x S_x}$$

- E a matriz de covariância é: $$M = \begin{bmatrix} \sigma_{a_0}^2 & \mathrm{cov}(a_0, a_1) \\ \mathrm{cov}(a_1, a_0) & \sigma_{a_1}^2 \end{bmatrix}$$

Generalizando

Se a função f(x,**a**) depende e é linear em k+1 parâmetros, **a**=(a_0,...,a_k), é fácil generalizar:

$$\begin{pmatrix} S_y \\ S_{xy} \\ \vdots \\ S_{x^k y} \end{pmatrix} = \begin{pmatrix} S_\sigma & S_x & & S_{x^k} \\ S_x & S_{x^2} & & \\ & & & \\ S_{x^k} & & & S_{x^{2k}} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{pmatrix}$$

Resolvemos multiplicando pela inversa da matriz quadrada dos dois lados.

Mas note que é preciso calcular os termos "S" para construir a matriz. Não é tão fácil...