ME170b Lecture 7

3/10/23

Experimental Techniques

Last time:

> Rejection data

> Weighted Averages

> Least Squares

Today:

> Notes of least squares

> Ch. 9

Correlation & covariance

Another look at least squares (from optimization perspective)

$F = a z + b$ (model)

measurements

parameters

$F_i = a z_i + b$

F_1, z_1

$\vdots$

F_N, z_N

$b = Ax$

as our parameters

$x = \begin{bmatrix} a \\ b \end{bmatrix}$

$$\begin{bmatrix} F_1 \\ \vdots \\ F_N \end{bmatrix} = \begin{bmatrix} z_1 & 1 \\ \vdots & \vdots \\ z_N & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$$

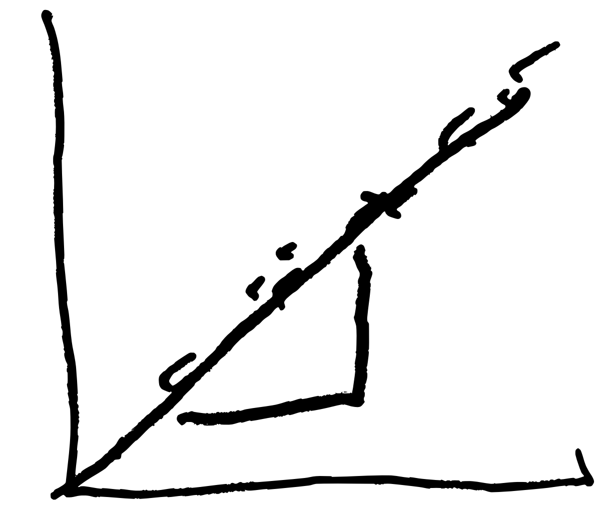
$$\boxed{b = A x}$$

Another look at least squares (from optimization perspective)

$$b = Ax$$

$$\boxed{\min \quad \|b - Ax\|_2^2}$$

$$N\times 1 \quad - \quad (N\times 2)\,(2\times 1)$$



$$\|\cdot\|_2^2 = \left(\sqrt{(\cdot)^2}\right)^2 = (\cdot)^2$$

$$\min \quad \underbrace{(b - Ax)^T (b - Ax)}$$

$$J = b^T b - b^T A x - (Ax)^T b + x^T A^T A x$$

Another look at least squares (from optimization perspective)

$$\frac{\partial J}{\partial x} = 0$$

$$J = b^T b - x^T A^T b - b^T A x + x^T A$$

$$\frac{\partial J}{\partial x} = 0 - A^T b - b^T A + 2A^T A x$$

$$= -2A^T b + 2A^T A x \Rightarrow 2A^T A x = 2A^T b$$

$$\boxed{x^* = \underbrace{(A^T A)^{-1} A^T}\, b}$$

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

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

$$F = az + bz^2 + c$$

can you rearrange into

$Ax = b$?

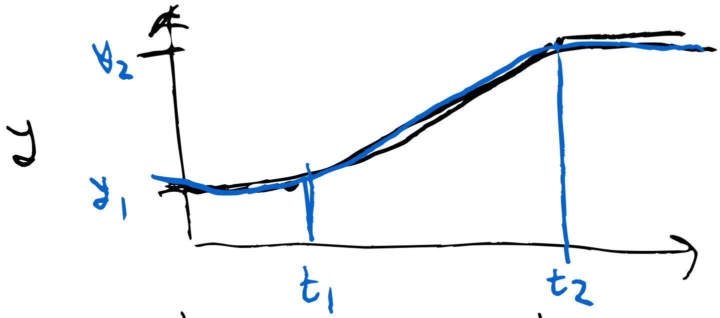
$$x = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

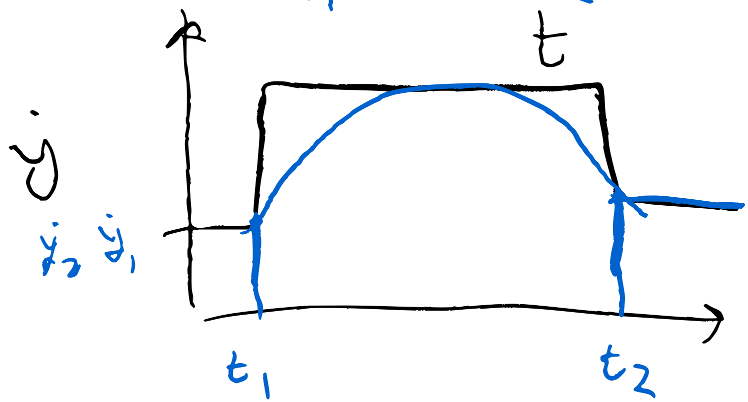
F_i	x_i

$$\begin{bmatrix} F_1 \\ \vdots \\ F_N \end{bmatrix} = \begin{bmatrix} x_1 & x_1^2 & 1 \\ \vdots & \vdots & \vdots \\ x_N & x_N^2 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

$$F = A \cdot \sin(z) + B$$

$$\begin{bmatrix} F_1 \\ \vdots \\ F_N \end{bmatrix} = \begin{bmatrix} \sin(z) & 1 \\ \vdots & \vdots \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix}$$





$$y = a_0 + a_2 t + a_3 t^2$$

$$y' = a_2 + 2a_3 t$$

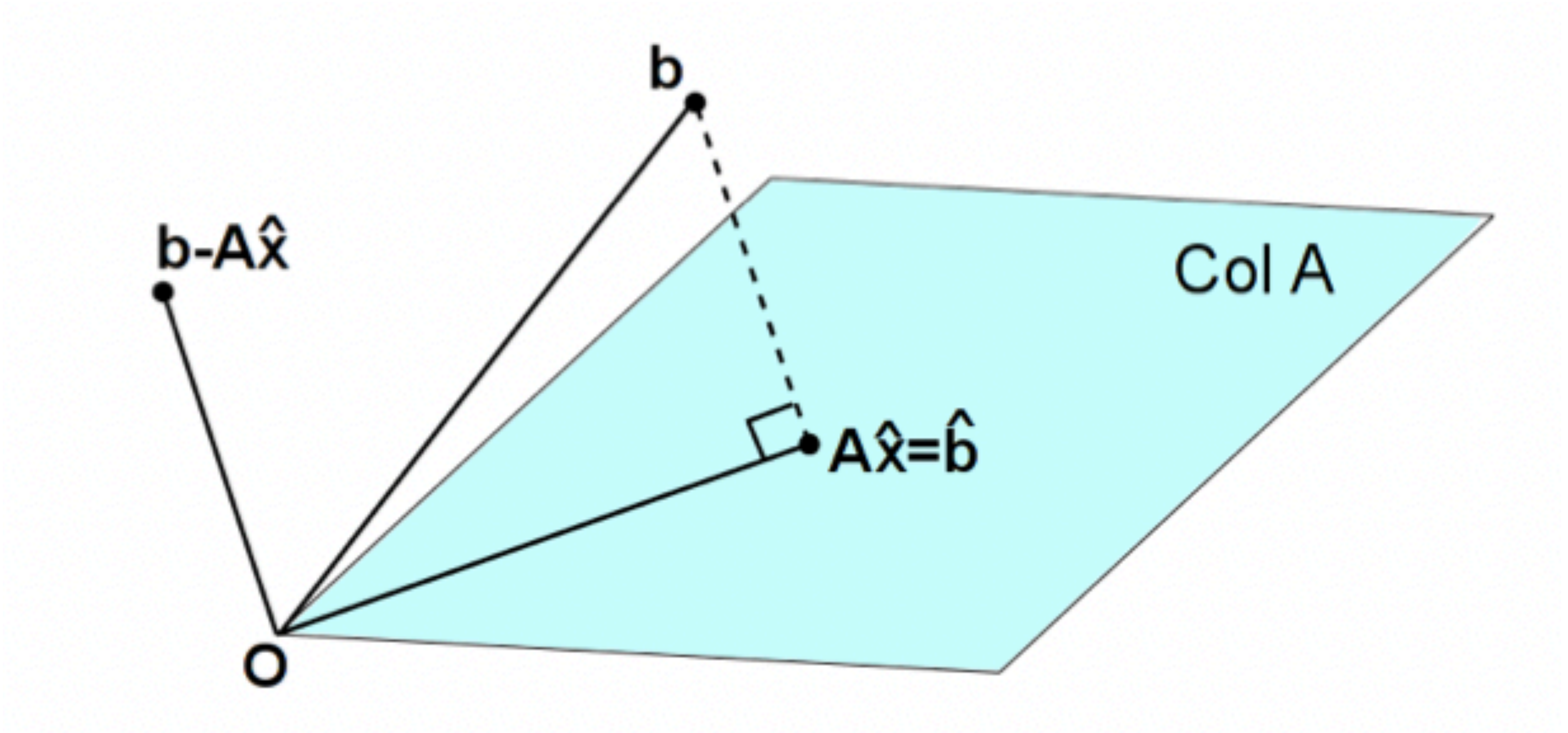
$$\left.\begin{aligned} y(t_1) &= y_1 \\ y(t_2) &= y_2 \\ \dot{y}(t_1) &= 0 \\ \dot{y}(t_2) &= 0 \end{aligned}\right]$$

$$\begin{bmatrix} y_1 \\ y_2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & t_1 & t_1^2 \\ 1 & t_2 & t_2^2 \\ 0 & 1 & 2t_1 \\ 0 & 1 & 2t_2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_2 \\ a_3 \end{bmatrix}$$

Another look at least squares (from optimization perspective)

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

Solution is the ‘projection’ of the on the space that matrix A spans



Covariance and Correlation

First let's review the principles of error propogation:

if we measure x ; y, to calculate $q(x, y)$:

$$\delta q \simeq \left|\frac{\partial q}{\partial x}\right| \delta x + \left|\frac{\partial q}{\partial y}\right| \delta y$$

Naive uncertainty measure

there could be canselations of errors in x ; y.

$$\delta q = \sqrt{\left(\frac{\partial q}{\partial x} \delta x\right)^2 + \left(\frac{\partial q}{\partial y} \delta y\right)^2}$$

we can prove this is better if Gaussians.

Covariance and Correlation

$$\delta q = \sqrt{\left(\frac{\partial q}{\partial x}\,\delta x\right)^2 + \left(\frac{\partial q}{\partial y}\,\delta y\right)^2}$$

$$\Updownarrow$$

$$\sigma_q = \sqrt{\left(\frac{\partial q}{\partial x}\,\sigma_x\right)^2 + \left(\frac{\partial q}{\partial y}\,\sigma_y\right)^2} \quad (1)$$

Assumption is x & y are independent & only have random error.

What if we cannot satisify this assumption? How should we update our estimates?

Claim: Eq. (1) is <u>always</u> an upper bound on uncertainty even if not independent!

Recall STD

$$\sigma_x^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2 \quad ✓$$

if x are Gaussian, then $N \to \infty$

$\sigma_x \to$ can be used as the width parameter (limiting Distribution)

$$\frac{1}{\sigma_x\sqrt{2\pi}} e^{-(x-\bar{X})^2/2\sigma_x^2}$$

σ_x is still STD, but we can no longer use the Gaussian as our limiting distribution.

Covariance Propagation

Suppose we measure $(x_1, y_1) \dots (x_N, y_N)$
and we want to calculate $q(x, y)$

We can calculate the following:

1. mean $\bar{x}$, σ_x
2. mean $\bar{y}$, σ_y
3. mean $\bar{q}$, σ_q

Covariance Propagation

$$(x_i, y_i) \quad i = 1 \cdots N$$

$$q_i = q(x_i, y_i)$$

$$\simeq q(\bar{x}, \bar{y}) + \frac{\partial q}{\partial x}(x_i - \bar{x}) + \frac{\partial q}{\partial y}(y_i - \bar{y})$$

$$\bar{q} = \frac{1}{N}\sum_{i=1}^{N} q_i$$

$$= \frac{1}{N}\sum_i \Big[\underbrace{q(\bar{x}, \bar{y})} + \overbrace{\frac{\partial q}{\partial x}\underbrace{(x_i - \bar{x})}_{0} + \frac{\partial q}{\partial y}\underbrace{(y_i - \bar{y})}_{0}} \Big]$$

$$\boxed{\bar{q} = q(\bar{x}, \bar{y})}$$

$$\sigma_q^2 = \frac{1}{N}\sum (q_i - \bar{q})^2$$

Covariance Propagation

$$\sigma_q^2 = \frac{1}{N}\sum (q_i - \bar{q})^2$$

$$= \frac{1}{N}\sum \left[\frac{\partial q}{\partial x}(x_i - \bar{x}) + \frac{\partial q}{\partial y}(y_i - \bar{y}) \right]^2$$

$$\sigma_q^2 = \left(\frac{\partial q}{\partial x}\right)^2 \overbrace{\frac{1}{N}\sum (x_i - \bar{x})^2}^{\sigma_x} + \left(\frac{\partial q}{\partial y}\right)^2 \overbrace{\frac{1}{N}\sum (y_i - \bar{y})^2}^{\sigma_y}$$

$$+ \underbrace{2 \frac{\partial q}{\partial x}\frac{\partial q}{\partial y}\frac{1}{N}\sum (x_i - \bar{x})(y_i - \bar{y})}_{\sigma_{xy}}.$$

Covariance Propagation

$$\sigma_{xy} = \frac{1}{N}\sum_{i}^{N} (x_i - \bar{x})(y_i - \bar{y})$$

$$\boxed{\sigma_z^2 = \left(\frac{\partial z}{\partial x}\right)^2 \sigma_x^2 + \left(\frac{\partial z}{\partial y}\right)^2 \sigma_y^2 + 2\frac{\partial z}{\partial x}\frac{\partial z}{\partial y}\sigma_{xy}}$$

Covariance Propagation

(1) if x & y are not independent, then the σ_{xy} is non-zero

(2) if x & y are independent $\sigma_{xy} = 0$

(3) when σ_{xy} is non-zero we say "errors in x are correlated with errors in y"

Example: Two Angles with a Negative Covariance

Each of five students measures the same two angles α and β and obtains the results shown in the first three columns of Table 9.1.

Table 9.1. Five measurements of two angles α and β (in degrees).

Student	α	β	$(\alpha - \overline{\alpha})$	$(\beta - \overline{\beta})$	$(\alpha - \overline{\alpha})(\beta - \overline{\beta})$
A	35	50	2	−2	−4
B	31	55	−2	3	−6
C	33	51	0	−1	0
D	32	53	−1	1	−1
E	34	51	1	−1	−1

$$\sigma_{\alpha\beta} = \frac{1}{N}\Sigma(\alpha - \overline{\alpha})(\beta - \overline{\beta}) = \frac{1}{5} \times (-12) = -2.4$$

Upper limit on sigma_q

Schwarz inequality

$$|\sigma_{xy}| \le \sigma_x \sigma_y$$

$$\sigma_q^2 \le \left(\frac{\partial q}{\partial x}\right)^2 \sigma_x^2 + \left(\frac{\partial q}{\partial y}\right)^2 \sigma_y^2 + 2\left|\frac{\partial q}{\partial x}\frac{\partial q}{\partial y}\right| \sigma_x \sigma_y$$

$$\left[\left|\frac{\partial q}{\partial x}\right| \sigma_x + \left|\frac{\partial q}{\partial y}\right| \sigma_y\right]^2$$

$$\boxed{\sigma_q \le \left|\frac{\partial q}{\partial x}\right| \sigma_x + \left|\frac{\partial q}{\partial y}\right| \sigma_y} \quad ✓$$

naive estimate is still upper bound

Main Results on Covariance

Coefficient of Linear Correlation

Q: Given a set of measurements

$(x_1, y_1) \cdots (x_n, y_n)$, how well do

they support the hypothesis that x & y

are linearly related?

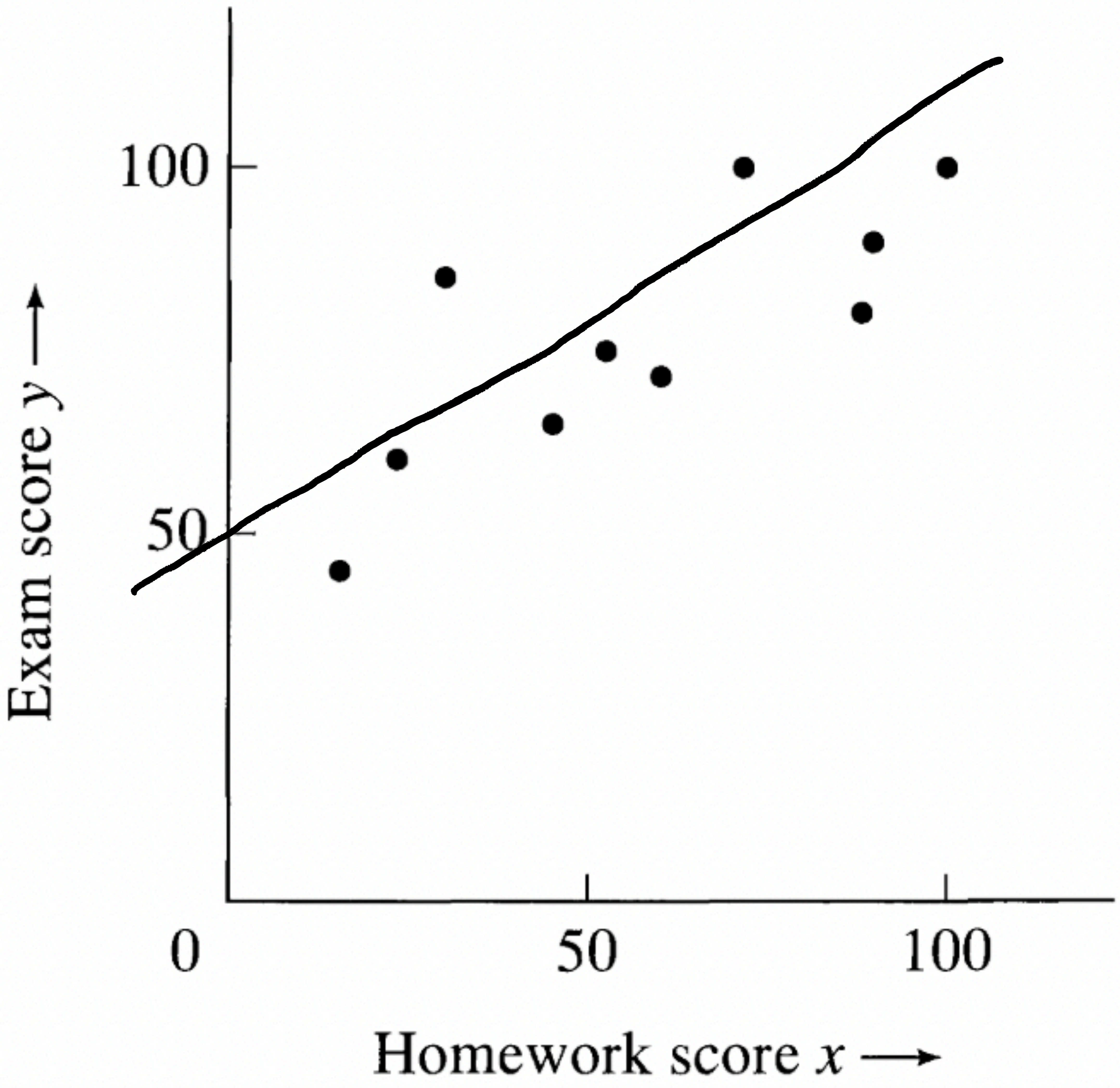
$$y = A + Bx$$

we fit with least squares

How to determine "goodness of fit"

Coefficient of Linear Correlation

Example



HW vs Exam scores

the problem:

we have no sense of the uncertainties

$$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

Correlation Coefficient

$$|\sigma_{xy}| \leq \sigma_x \sigma_y$$

$$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

$$-1 \leq r \leq 1$$

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

Correlation Coefficient - perfect correlation

$$y_i = A + Bx_i$$

$$\bar{y} = A + B\bar{x}$$

$$y_i - \bar{y} = B(x_i - \bar{x})$$

$$r = \frac{B\sum(x_i - \bar{x})^2}{\sqrt{\sum(x_i - \bar{x})^2 B^2 \sum(x_i - \bar{x})^2}} = \frac{B}{|B|} = \pm 1$$

Correlation Coefficient revisit data

$r = 0.8$

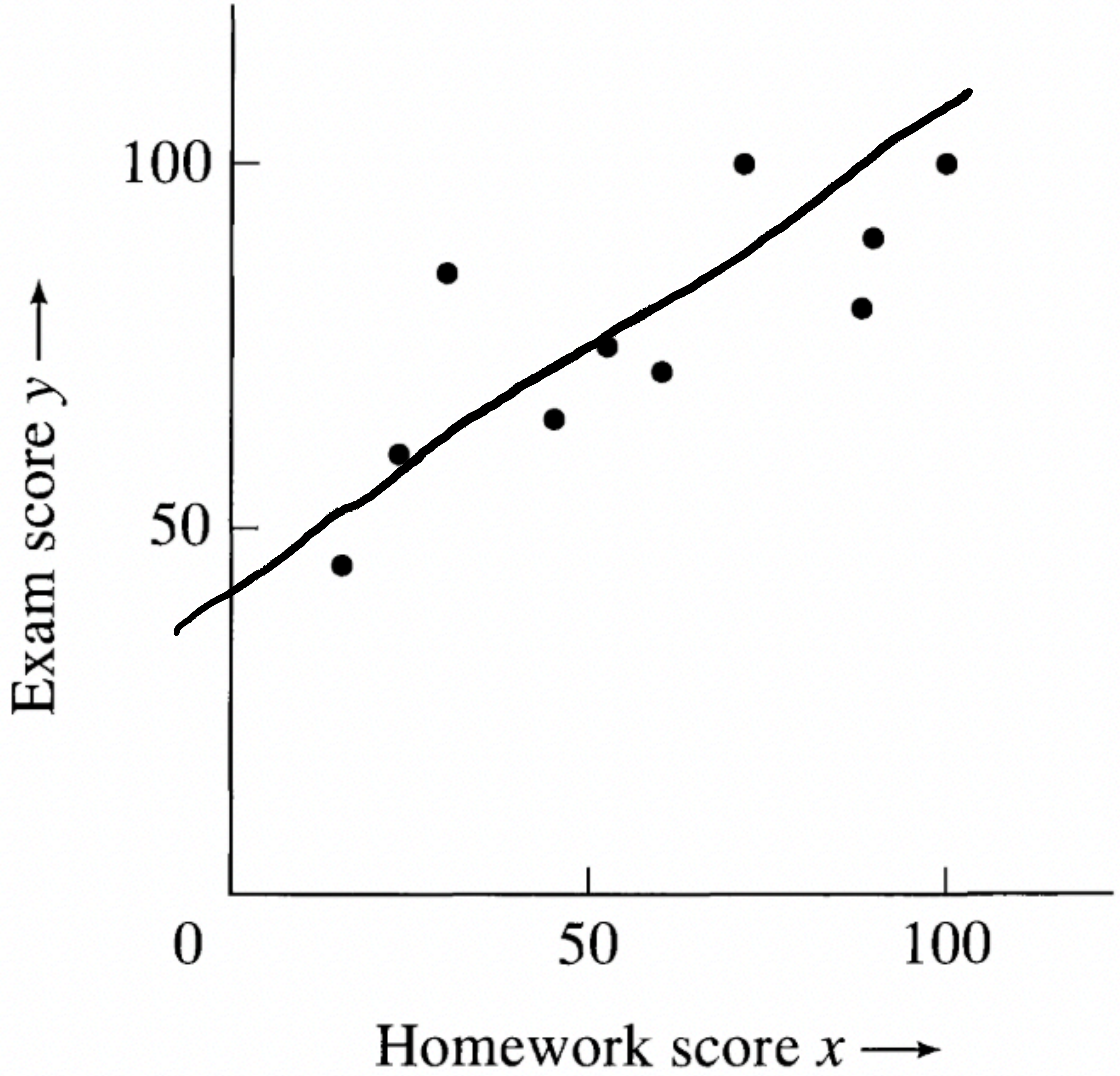


Table 9.3. Students' scores.

Student i	1	2	3	4	5	6	7	8	9	10
Homework x_i	90	60	45	100	15	23	52	30	71	88
Exam y_i	90	71	65	100	45	60	75	85	100	80

Quantitative Significance of r

Q: How to decide objectively what is a good r, given a certain data set (N-samples)

X & y uncorrelated $r \longrightarrow 0$

$$\text{Prob}_N(|r| \geq r_0) \Rightarrow \text{Prob}_N(|r| \geq 0.8)$$

Not straight forward calculation

Prob N measurements of two uncorrelated variables x and y would produce a correlation coefficient with

	r_o										
N	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
3	100	94	87	81	74	67	59	51	41	29	0
6	100	85	70	56	43	31	21	12	6	1	0
10	100	78	58	40	25	14	7	2	0.5		0
20	100	67	40	20	8	2	0.5	0.1			0
50	100	49	16	3	0.4						0

Back to our original question

Q: Given a set of measurement $(x_1 \cdots x_N)$
$(y_1 \cdots y_N)$

how well do they support
the hypothesis that they are linearly related.

1. calculate r
2. $\text{Prob}_N(|r| \geq r_0)$ with uncorrelated data.
3. if the Prob. is sufficiently small,
   the the data supports your hypothesis.

"significant" 5%
"highly significant" 1%