PHYS 232 Mar. 6, 2024

- Assignment #4 due Wed., Mar. 13
- Sign up for Experiment #5

Last Time:

Weighted Mean:

$$\boxed{\mu' = \frac{\sum_{i=1}^{N} \frac{x_i}{\sigma_i^2}}{\sum_{i=1}^{N} \frac{1}{\sigma_i^2}} \qquad \sigma_{\mu'}^2 = \frac{1}{\sum_{i=1}^{N} \frac{1}{\sigma_i^2}}}$$

If all σ_i equal to σ, then $\sum_{i=1}^{N} \frac{1}{\sigma_i^2} = \frac{N}{\sigma^2}$

& μ' & $\sigma_{\mu'}$ reduce to:

$$\boxed{\mu' = \frac{1}{N} \sum_{i=1}^{N} x_i \qquad \sigma_{\mu'} = \frac{\sigma}{\sqrt{N}}}$$

Today: "Least-Squares" Fit to a Straight Line

Often meas. a quantity y as a fcn of x.

Instead of repeated meas. of a value x, make a series of meas. of pairs (x_i, y_i).

Consider linear fcns ...

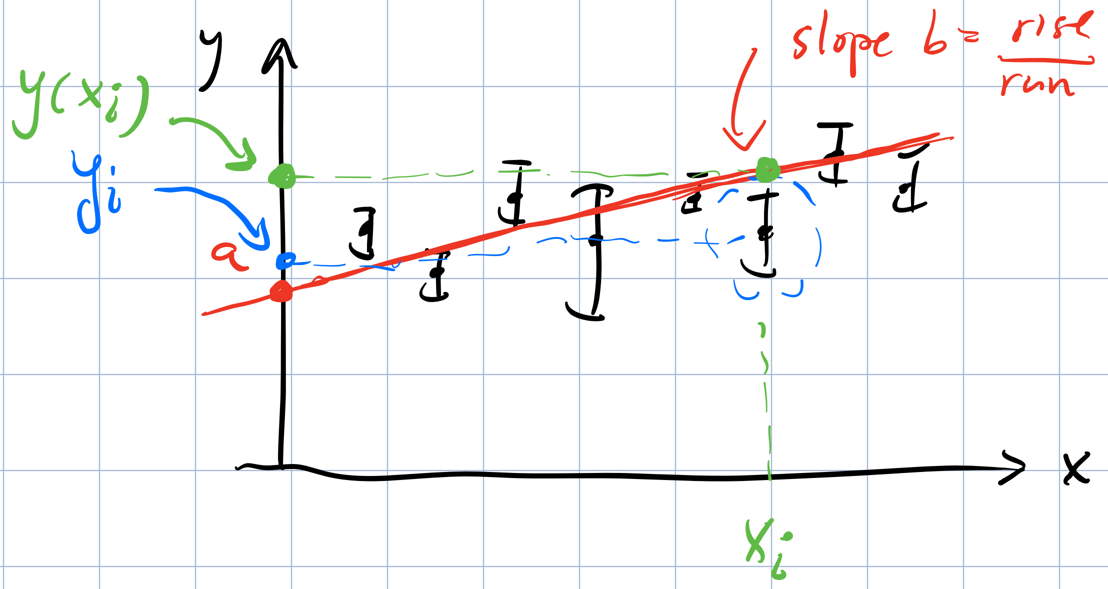
$$\text{eg.} \quad \ln I = \ln I_0 + \left(\frac{e}{k_B T}\right) V$$

$y = \ln I$ depends linearly on V

$$y = a + bx$$

$a = \ln I_0$ (intercept)

$b = \dfrac{e}{k_B T}$ (slope)



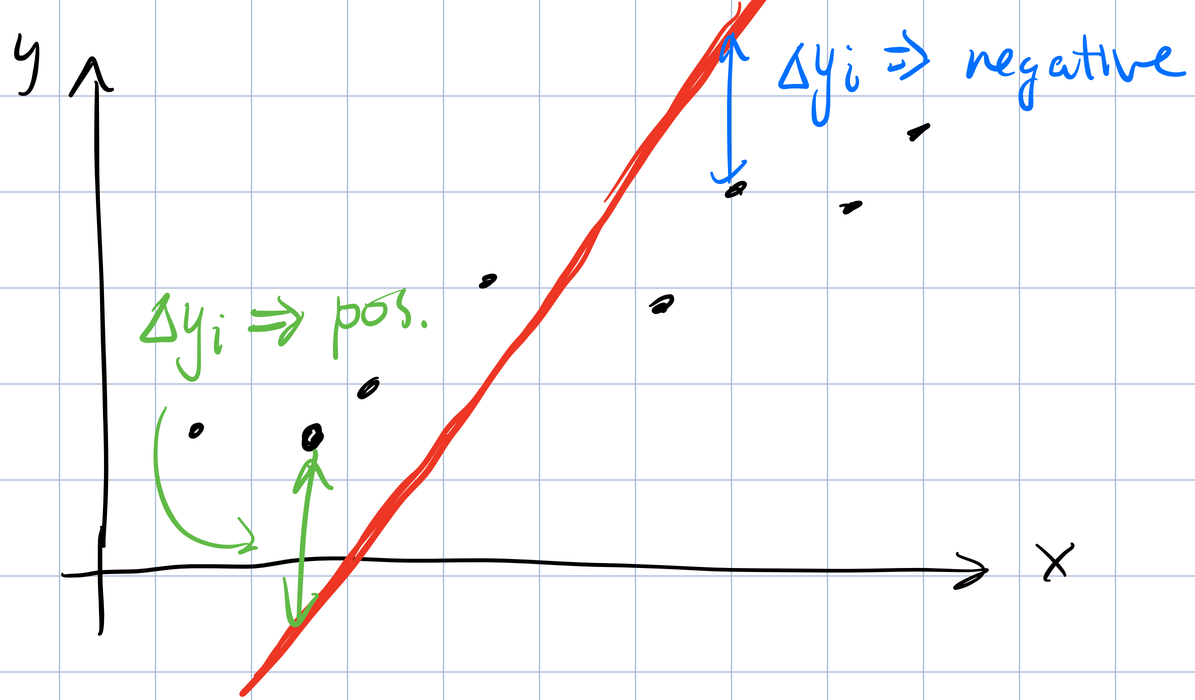
given a set of data $(x_i, y_i \pm \sigma_i)$, what are the most probable values of a & b?

In our approach, we will assume uncertainty in x_i values negligible wrt. uncertainty in y_i measurements.

We will examine the deviations between meas. y_i & y calculated from x_i:

$$\Delta y_i = y_i - y(x_i) = y_i - (a + b x_i)$$

Expect Δy_i to be small for good choices of a & b values.



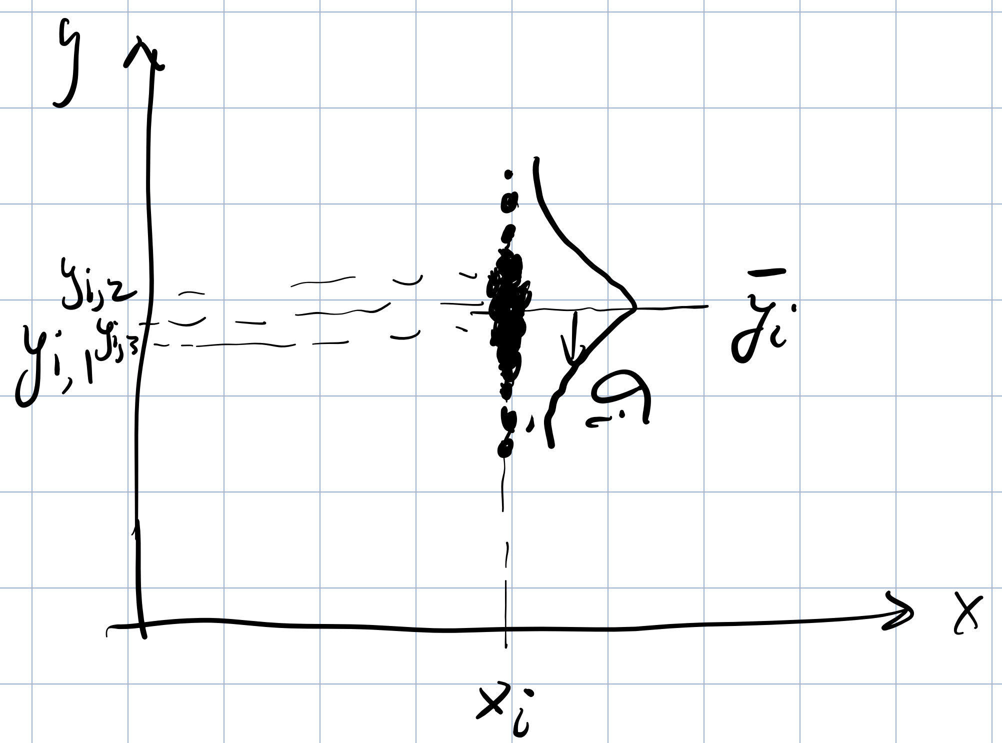
Δy_i on its own is <u>not</u> a good meas. of quality of fit.

In the example above,

$$\sum_{i=1}^{N} \Delta y_i \text{ would be small}$$

even though the "fit" is poor.

Better approach: Assume that y_i meas. drawn from a Gaussian dist'n.



Our Gaussian dist'n of y_i values would have a mean $y_0(x_i) = a_0 + b_0 x_i$

where a_0 & b_0 are true parameters that describe our line (unknown). The dist'n also has a st. dev. σ_i

Prob. of meas. a value of y between

y_i and $y_i + \Delta y$ is :

$$P_i = \frac{\Delta y}{\sigma_i \sqrt{2\pi}} \exp\left\{ -\frac{1}{2}\left[\frac{y_i - y_0(x_i)}{\sigma_i} \right]^2 \right\}$$

Prob. of making set of meas. $(x_1, y_1 + \sigma_1)$
$(x_2, y_2 \pm \sigma_2)$
$\vdots$
$(x_N, y_N \pm \sigma_N)$

is :

$$P(a_0, b_0) = \prod_{i=1}^{N} P_i$$

$$= \left[\prod_{i=1}^{N} \left(\frac{\Delta y}{\sigma_i \sqrt{2\pi}} \right) \right] \exp\left\{ -\frac{1}{2} \sum_{i=1}^{N} \left[\frac{y_i - y_0(x_i)}{\sigma_i} \right]^2 \right\}$$

To find the best estimates of a_0 & b_0 from meas. data, we will maximize $P(a_0, b_0)$ by minimizing

$$\chi^2 \equiv \sum_{i=1}^{N}\left[\frac{y_i - y(x_i)}{\sigma_i}\right]^2$$

Chi-squared → a "goodness of fit" parameter

$$\boxed{\chi^2 = \sum_{i=1}^{N}\left[\frac{y_i - a - b x_i}{\sigma_i}\right]^2}$$

Minimize χ^2 w.r.t. a & b.

2 eq'ns & 2 unknowns (a,b)

$$\frac{\partial \chi^2}{\partial a} = 0 = -2\sum_{i=1}^{N}\left[\frac{y_i - a - b x_i}{\sigma_i^2}\right]$$

$$\frac{\partial \chi^2}{\partial b} = 0 = -2\sum_{i=1}^{N}\left[\frac{x_i}{\sigma_i^2}(y_i - a - b x_i)\right]$$

$$\therefore \quad \underbrace{\sum \frac{y_i}{\sigma_i^2}}_{u} = a \underbrace{\sum \frac{1}{\sigma_i^2}}_{u_a} + b \underbrace{\sum \frac{x_i}{\sigma_i^2}}_{u_b}$$

✓

$$\underbrace{\sum \frac{x_i y_i}{\sigma_i^2}}_{v} = a \underbrace{\sum \frac{x_i}{\sigma_i^2}}_{v_a} + b \underbrace{\sum \frac{x_i^2}{\sigma_i^2}}_{v_b}$$

solve system of 2 eq'ns for 2 unknowns a & b.

Of the form:

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u_a & u_b \\ v_a & v_b \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}$$

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} a u_a + b u_b \\ a v_a + b v_b \end{pmatrix}$$ ✓

If $$\begin{pmatrix} u \\ v \end{pmatrix} = \overbrace{\begin{pmatrix} u_a & u_b \\ v_a & v_b \end{pmatrix}}^{A} \begin{pmatrix} a \\ b \end{pmatrix}$$

then $\begin{pmatrix} a \\ b \end{pmatrix} = A^{-1}\begin{pmatrix} u \\ v \end{pmatrix}$

For 2×2 matrix

$$A^{-1} = \frac{1}{\det A}\begin{pmatrix} V_b & -U_b \\ -V_a & U_a \end{pmatrix}$$

$$\therefore \begin{pmatrix} a \\ b \end{pmatrix} = \frac{1}{\det A}\begin{pmatrix} V_b & -U_b \\ -V_a & U_a \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix}$$

$$\det A = (U_a V_b - V_a U_b)$$

$$\boxed{a = \frac{V_b u - U_b v}{U_a V_b - V_a U_b}}$$

$$\boxed{b = \frac{-V_a U + U_a V}{U_a V_b - V_a U_b}}$$

Subbing in for our definitions of u, V
U_a, V_a, ~~ϵ~~, U_b, V_b we get:

$$a = \frac{\sum \frac{x_i^2}{\sigma_i^2} \sum \frac{y_i}{\sigma_i^2} - \sum \frac{x_i}{\sigma_i^2} \sum \frac{x_i y_i}{\sigma_i^2}}{\sum \frac{1}{\sigma_i^2} \sum \frac{x_i^2}{\sigma_i^2} - \left(\sum \frac{x_i}{\sigma_i^2}\right)^2}$$

$$\boxed{\begin{gathered} a = \frac{1}{\Delta}\left(\sum \frac{x_i^2}{\sigma_i^2} \sum \frac{y_i}{\sigma_i^2} - \sum \frac{x_i}{\sigma_i^2} \sum \frac{x_i y_i}{\sigma_i^2}\right) \\ \Delta = \sum \frac{1}{\sigma_i^2} \sum \frac{x_i^2}{\sigma_i^2} - \left(\sum \frac{x_i}{\sigma_i^2}\right)^2 \end{gathered}}$$

In the same way, find

$$\boxed{b = \frac{1}{\Delta}\left(\sum \frac{1}{\sigma_i^2} \sum \frac{x_i y_i}{\sigma_i^2} - \sum \frac{x_i}{\sigma_i^2} \sum \frac{y_i}{\sigma_i^2}\right)}$$

Best-fit parameters for a weighted linear fit.

Next time, use prop. of errors to find σ_a & σ_b.