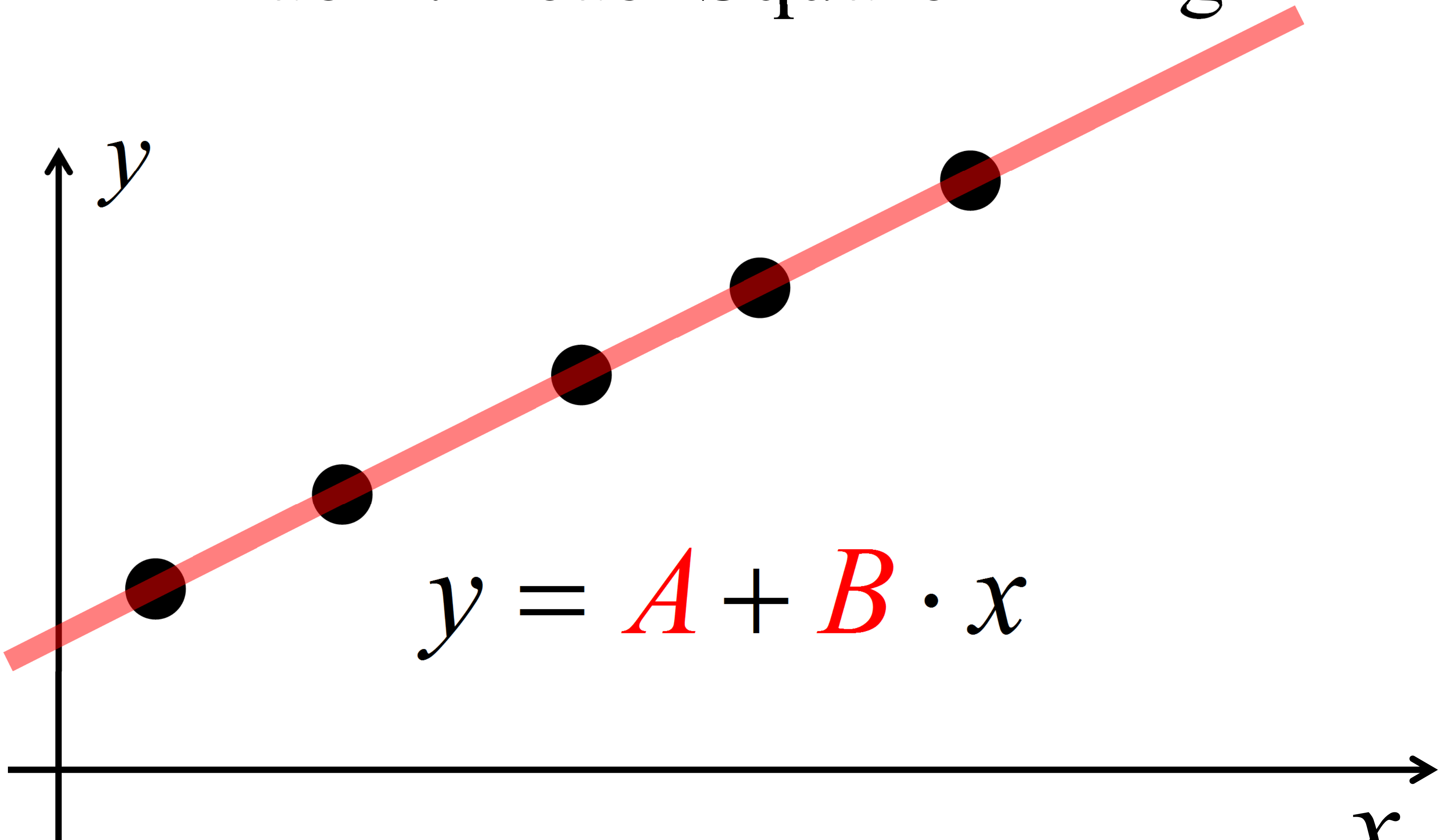
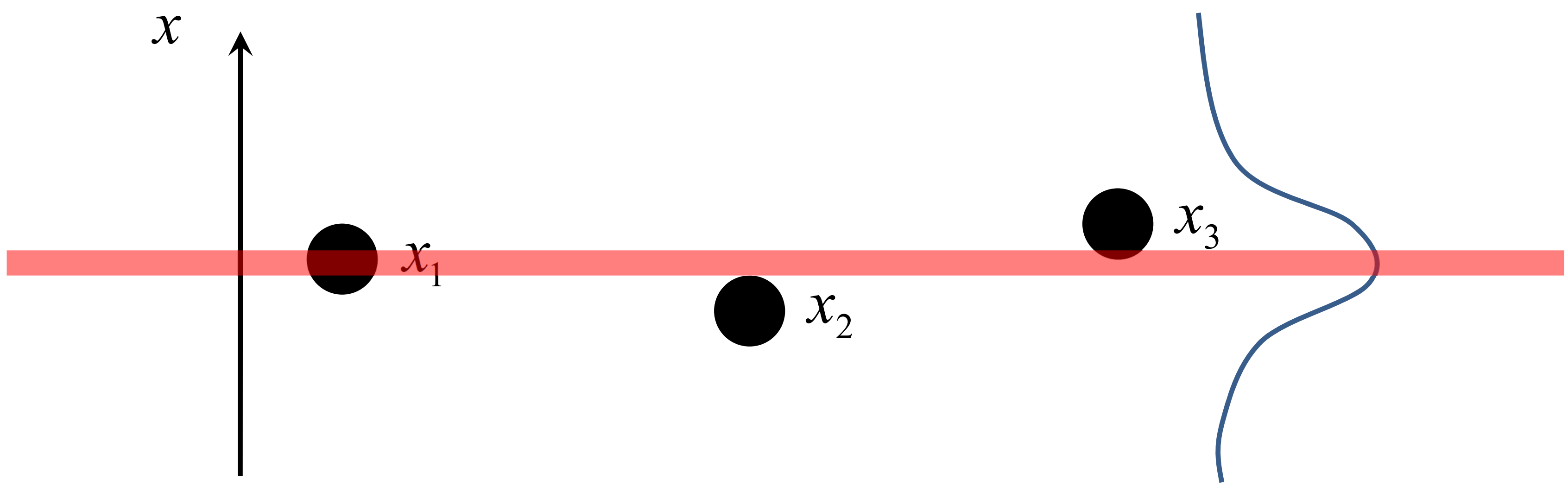
Lab 4: Least Square Fitting



Lab 4: Least Square Fitting

- The most popular approach of linear regression ($y=A+Bx$)
  - Linear regression is widely used in biological, behavioral and social sciences to describe possible relationships between variables. It is ranked as one of the most important tools used in these disciplines.
- Based on a set of measurements (x_i, y_i)
  - Calculate parameters: A and B.
  - Evaluate the quality of the fitting
  - Principle of maximum likelihood

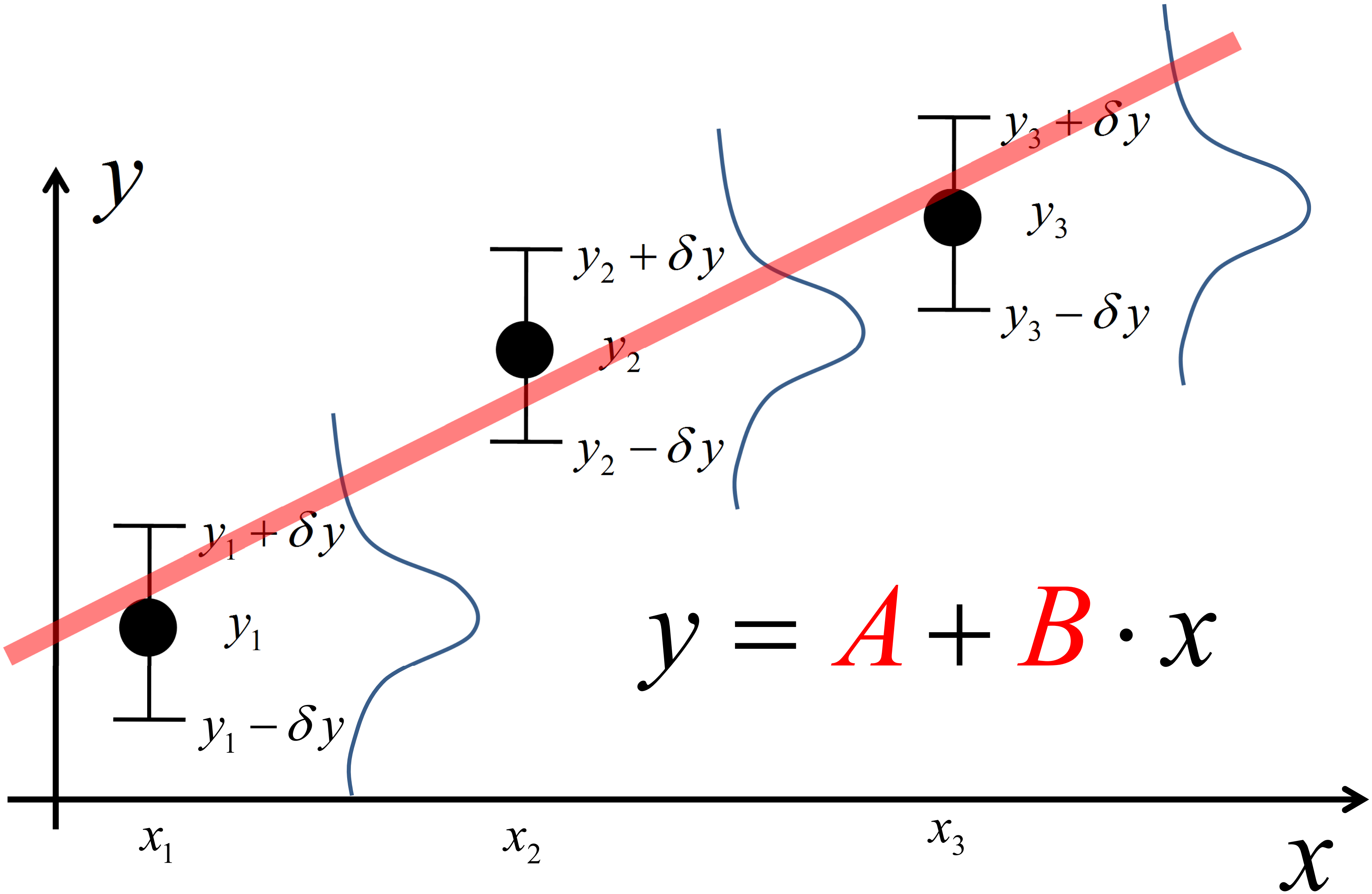
Statistics of a single quantity



Principle of maximum likelihood

$$X \approx \overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \qquad \sigma^2 \approx \sigma_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(x_i - \overline{x}\right)^2$$

Statistics of the relationship between multiple quantities



Calculate *A* and *B* with least square fitting

The simplest case

Assumptions:

for a set of measurements $[(x_i, y_i), i=1, \ldots, N]$

1. Ignore uncertainties of x_i (correlated with y_i);
2. Uncertainty of y follows a Gaussian distribution w/ true value $Y_i = A + B \cdot x_i$, and the **same** standard deviation σ_y (no need to know its value *a priori*)
3. Principle of maximum likelihood.

Calculate *A* and *B* with least square fitting

The simplest case

The probability of one measurement (x_i, y_i) is:

$$P_{A,B,\sigma_y}(x_i, y_i) \propto \frac{1}{\sigma_y} e^{-(y_i - Y_i)^2/2\sigma_y^2} = \frac{1}{\sigma_y} e^{-(y_i - A - Bx_i)^2/2\sigma_y^2}$$

The probability of $[(x_i, y_i), i=1, \ldots, N]$ is:

$$\Rightarrow \quad P_{A,B,\sigma_y}(\text{set}) \propto \prod_{i=1}^{N} \frac{1}{\sigma_y} e^{-(y_i - Y_i)^2/2\sigma_y^2} = \frac{1}{\sigma_y^N} e^{-\chi^2/2}$$

Chi square: $\chi^2 = \frac{1}{\sigma_y^2} \sum_{i=1}^{N} (y_i - A - Bx_i)^2$

χ^2 is a measure of how well the fitting is.

The simplest case (cont.)

Principle of maximum likelihood: $P_{A,B}(\text{set}) = \max$

$$\Rightarrow \quad \frac{\partial P_{A,B,\sigma_y}(\text{set})}{\partial A} = 0 \quad \text{and} \quad \frac{\partial P_{A,B,\sigma_y}(\text{set})}{\partial B} = 0$$

$$\Rightarrow \quad \frac{\partial}{\partial A}\left[\frac{1}{\sigma_y^N} e^{-\frac{\chi^2}{2}}\right] = 0 \quad \text{and} \quad \frac{\partial}{\partial B}\left[\frac{1}{\sigma_y^N} e^{-\frac{\chi^2}{2}}\right] = 0$$

$$\Rightarrow \quad \begin{cases} \dfrac{\partial \chi^2}{\partial A} = -\dfrac{2}{\sigma_y^2}\displaystyle\sum_{i=1}^{N}\left(y_i - A - Bx_i\right) = 0 \\ \dfrac{\partial \chi^2}{\partial B} = -\dfrac{2}{\sigma_y^2}\displaystyle\sum_{i=1}^{N}x_i\left(y_i - A - Bx_i\right) = 0 \end{cases}$$

The simplest case (cont.)

$$\begin{cases} \sum_{i=1}^{N} (y_i - A - Bx_i) = 0 \\ \sum_{i=1}^{N} x_i (y_i - A - Bx_i) = 0 \end{cases} \Rightarrow \begin{cases} A \bullet N + B \bullet \sum_{i=1}^{N} x_i = \sum_{i=1}^{N} y_i \\ A \bullet \sum_{i=1}^{N} x_i + B \bullet \sum_{i=1}^{N} x_i^2 = \sum_{i=1}^{N} x_i y_i \end{cases}$$

$$M_{11} = N,\ M_{12} = M_{21} = \sum_{i=1}^{N} x_i,\ M_{22} = \sum_{i=1}^{N} x_i^2,\ V_1 = \sum_{i=1}^{N} y_i,\ V_2 = \sum_{i=1}^{N} x_i y_i$$

$$\Rightarrow \begin{cases} A \bullet M_{11} + B \bullet M_{12} = V_1 \\ A \bullet M_{21} + B \bullet M_{22} = V_2 \end{cases}$$

$$\Leftrightarrow \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$$

The simplest case (cont.)

The solution is:

$$\begin{pmatrix} \hat{A} \\ \hat{B} \end{pmatrix} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}^{-1} \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = \frac{1}{\det \mathbf{M}} \begin{pmatrix} M_{22} & -M_{12} \\ -M_{21} & M_{11} \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$$

$$\begin{cases} \hat{A} = \dfrac{1}{\Delta}\left(M_{22} \bullet V_1 - M_{12} \bullet V_2\right) \\ \hat{B} = \dfrac{1}{\Delta}\left(M_{11} \bullet V_2 - M_{21} \bullet V_1\right) \end{cases}$$

$$M_{11} = N,\ M_{12} = M_{21} = \sum_{i=1}^{N} x_i,\ M_{22} = \sum_{i=1}^{N} x_i^2,\ V_1 = \sum_{i=1}^{N} y_i,\ V_2 = \sum_{i=1}^{N} x_i y_i$$

where $\Delta = \det \mathbf{M} = N \bullet \sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i\right)^2$ No need to know σ_y!

Estimate the uncertainty of y

Similar to N measurement of the **same** quantity: (if we know the true values of A and B):

$$\sigma_y = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_i - Y_i\right)^2} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_i - A - Bx_i\right)^2}$$

However, we don't really know the true value of A and B. Instead, we use the best estimates for A and B, which reduce the value of above formula and need to be compensated.

$$\sigma_y = \sqrt{\frac{1}{N-\boxed{2}}\sum_{i=1}^{N}\left(y_i - \hat{A} - \hat{B}x_i\right)^2}$$

One can always find a line that perfectly passes through 2 points.

Uncertainties of *A* and *B*

$$A=\left[\left(\sum_{i=1}^{N}x_i^2\right)\bullet\left(\sum_{i=1}^{N}y_i\right)-\left(\sum_{i=1}^{N}x_i\right)\bullet\left(\sum_{i=1}^{N}x_iy_i\right)\right]\Big/\Delta$$

$$B=\left[N\bullet\sum_{i=1}^{N}x_iy_i-\left(\sum_{i=1}^{N}x_i\right)\bullet\left(\sum_{i=1}^{N}y_i\right)\right]\Big/\Delta$$

where $\Delta = N\bullet\sum_{i=1}^{N}x_i^2-\left(\sum_{i=1}^{N}x_i\right)^2$

Using error propagation formula:

$$\sigma_A^2=\sum_{i=1}^{N}\left(\frac{\partial A}{\partial y_i}\sigma_y\right)^2=\sigma_y^2\sum_{i=1}^{N}\left(\frac{\partial A}{\partial y_i}\right)^2$$

$$=\frac{\sigma_y^2}{\Delta^2}\sum_{i=1}^{N}\left[\left(\sum_{j=1}^{N}x_j^2\right)-x_i\left(\sum_{j=1}^{N}x_j\right)\right]^2$$

$$=\frac{\sigma_y^2}{\Delta^2}\left[N\left(\sum_{i=1}^{N}x_i^2\right)^2-\left(\sum_{i=1}^{N}x_i^2\right)\left(\sum_{i=1}^{N}x_i\right)^2\right]$$

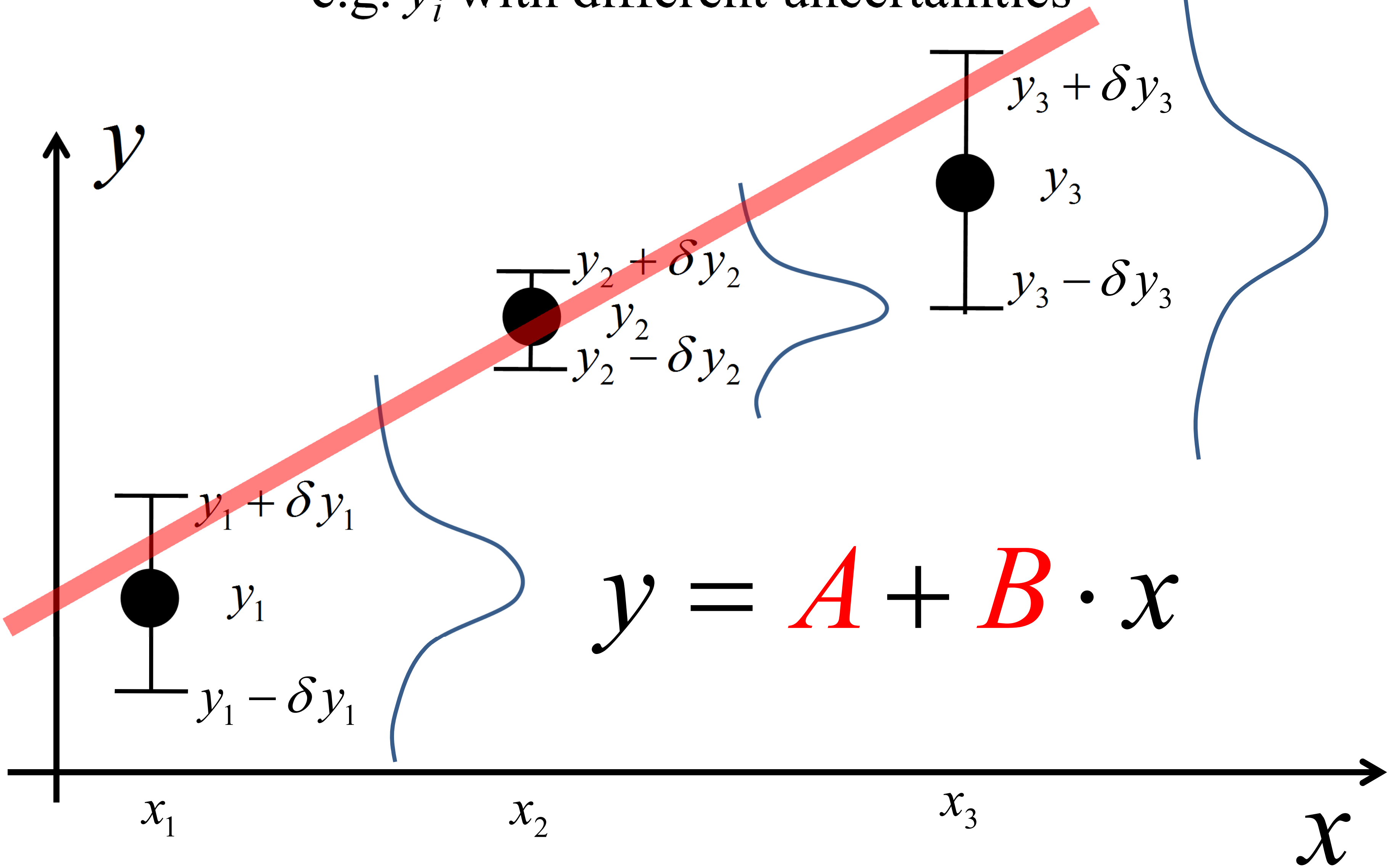
$$=\frac{\sigma_y^2}{\Delta}\sum_{i=1}^{N}x_i^2=\frac{\sigma_y^2}{\Delta}M_{22}$$

similarly,

$$\sigma_B^2=\frac{\sigma_y^2}{\Delta}N=\frac{\sigma_y^2}{\Delta}M_{11}$$

Weighted Least Square fitting (this lab)

e.g. y_i with different uncertainties



Weighted Least Square fitting

More general case

Assumptions:

for a set of measurements [(x_i, y_i), i=1, ..., N]

1. Ignore the uncertainties of x_i;
2. Uncertainties of y_i's follow Gaussian distribution w/ true values $Y_i=A+Bx_i$, and standard deviations $\boldsymbol{\sigma_i}$ (which are needed for fitting).
3. Principle of maximum likelihood.

Weighted Least Square fitting

More general case

The probability of one measurement (x_i, y_i) is:

$$P_{A,B}\left(x_i, y_i\right) \propto \frac{1}{\sigma_i} e^{-\left(y_i - Y_i\right)^2 / 2\sigma_i^2}$$

➡ The probability of [(x_i, y_i), i=1, ..., N] is:

$$P_{A,B}\left(\text{set}\right) \propto \prod_{i=1}^{N} \frac{1}{\sigma_i} e^{-\left(y_i - Y_i\right)^2 / 2\sigma_i^2} = \left(\prod_{i=1}^{N} \frac{1}{\sigma_i}\right) \times e^{-\frac{\chi^2}{2}}$$

Chi square: $$\chi^2 = \sum_{i=1}^{N} \frac{\left(y_i - A - Bx_i\right)^2}{\boxed{\sigma_i^2}}$$

Different for every i.

More general case (cont.)

Principle of maximum likelihood: $P_{A,B}(\text{set}) = \max$

⇒ $$\frac{\partial P_{A,B}(\text{set})}{\partial A} = 0 \quad \text{and} \quad \frac{\partial P_{A,B}(\text{set})}{\partial B} = 0$$

⇒ $$\frac{\partial}{\partial A}\left[\left(\prod_{i=1}^{N}\frac{1}{\sigma_i}\right)\times e^{-\frac{\chi^2}{2}}\right] = 0 \quad \text{and} \quad \frac{\partial}{\partial B}\left[\left(\prod_{i=1}^{N}\frac{1}{\sigma_i}\right)\times e^{-\frac{\chi^2}{2}}\right] = 0$$

⇒ $$\begin{cases} \dfrac{\partial \chi^2}{\partial A} = -2\displaystyle\sum_{i=1}^{N}\frac{(y_i - A - Bx_i)}{\sigma_i^2} = 0 \\ \dfrac{\partial \chi^2}{\partial B} = -2\displaystyle\sum_{i=1}^{N}\frac{x_i(y_i - A - Bx_i)}{\sigma_i^2} = 0 \end{cases}$$

$$\chi^2 = \sum_{i=1}^{N}\frac{(y_i - A - Bx_i)^2}{\sigma_i^2}$$

More general case (cont.)

$$\begin{cases} -2\sum_{i=1}^{N}\frac{\left(y_i - A - Bx_i\right)}{\sigma_i^2} = 0 \\ -2\sum_{i=1}^{N}\frac{x_i\left(y_i - A - Bx_i\right)}{\sigma_i^2} = 0 \end{cases} \Rightarrow \begin{cases} A\bullet\sum_{i=1}^{N}\frac{1}{\sigma_i^2} + B\bullet\sum_{i=1}^{N}\frac{x_i}{\sigma_i^2} = \sum_{i=1}^{N}\frac{y_i}{\sigma_i^2} \\ A\bullet\sum_{i=1}^{N}\frac{x_i}{\sigma_i^2} + B\bullet\sum_{i=1}^{N}\frac{x_i^2}{\sigma_i^2} = \sum_{i=1}^{N}\frac{x_i y_i}{\sigma_i^2} \end{cases}$$

$$\begin{cases} A\bullet M_{11} + B\bullet M_{12} = V_1 \\ A\bullet M_{21} + B\bullet M_{22} = V_2 \end{cases} \Rightarrow \boxed{\begin{cases} \hat{A} = \left(M_{22}\bullet V_1 - M_{12}\bullet V_2\right)/\Delta \\ \hat{B} = \left(M_{11}\bullet V_2 - M_{21}\bullet V_1\right)/\Delta \\ \Delta = M_{11}\bullet M_{22} - M_{12}\bullet M_{21} \end{cases}}$$

where: $M_{11} = \sum_{i=1}^{N}\frac{1}{\sigma_i^2}$, $M_{12} = M_{21} = \sum_{i=1}^{N}\frac{x_i}{\sigma_i^2}$, $V_1 = \sum_{i=1}^{N}\frac{y_i}{\sigma_i^2}$, $M_{22} = \sum_{i=1}^{N}\frac{x_i^2}{\sigma_i^2}$, $V_2 = \sum_{i=1}^{N}\frac{x_i y_i}{\sigma_i^2}$

Uncertainties of A and B

Pr. 8.19

Using error propagation formula:

$$\sigma_A^2 = \sum_{i=1}^{N}\left(\frac{\partial \hat{A}}{\partial y_i}\sigma_i\right)^2 \quad \text{and} \quad \sigma_B^2 = \sum_{i=1}^{N}\left(\frac{\partial \hat{B}}{\partial y_i}\sigma_i\right)^2$$

E.g. $\hat{A} = \left(M_{22} \bullet V_1 - M_{12} \bullet V_2\right) / \Delta$

$$\frac{\partial A}{\partial y_i}\sigma_i = \sigma_i\left(M_{22} \bullet \frac{\partial V_1}{\partial y_i} - M_{12} \bullet \frac{\partial V_2}{\partial y_i}\right) \Big/ \Delta$$

$$= \left(M_{22} \bullet \frac{1}{\sigma_i} - M_{12} \bullet \frac{x_i}{\sigma_i}\right) \Big/ \Delta$$

$$V_1 = \sum_{i=1}^{N}\frac{y_i}{\sigma_i^2}$$

$$V_2 = \sum_{i=1}^{N}\frac{x_i y_i}{\sigma_i^2}$$

Uncertainties of A and B (cont.)

$$\sigma_A^2 = \sum_{i=1}^{N}\left[\left(M_{22}\bullet\frac{1}{\sigma_i} - M_{12}\bullet\frac{x_i}{\sigma_i}\right)\Big/\Delta\right]^2$$

$$= \frac{1}{\Delta^2}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}\left(M_{22}^2 - 2M_{22}\bullet M_{12}x_i + M_{12}^2\bullet x_i^2\right)$$

$$= \frac{1}{\Delta^2}\left[M_{22}^2\bullet\left(\sum_{i=1}^{N}\frac{1}{\sigma_i^2}\right) - 2M_{22}\bullet M_{12}\left(\sum_{i=1}^{N}\frac{x_i}{\sigma_i^2}\right) + M_{12}^2\bullet\left(\sum_{i=1}^{N}\frac{x_i^2}{\sigma_i^2}\right)\right]$$

$$= \frac{1}{\Delta^2}\left[M_{22}^2\bullet M_{11} - 2M_{22}\bullet M_{12}\bullet M_{12} + M_{12}^2\bullet M_{22}\right]$$

$$= \frac{M_{22}}{\Delta}$$

similarly, $\sigma_B^2 = \dfrac{M_{11}}{\Delta}$

Summary of Least square fitting

Only linear algebra!

Assumptions:

for a set of measurements $[(x_i, y_i), i=1, \ldots, N]$

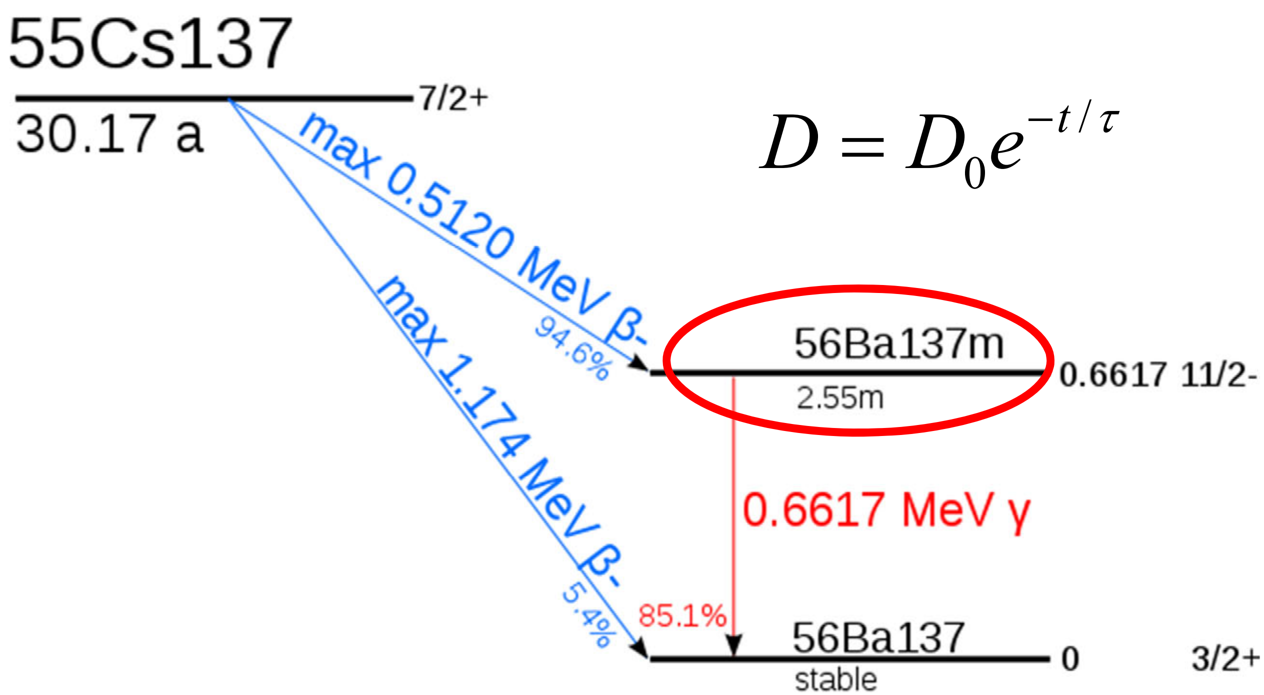
1. Uncertainty of y follows Gaussian distribution w/ true value $Y_i = A + Bx_i$, and standard deviation σ_i;
2. Principle of maximum likelihood.

$$\begin{cases} A = \left(M_{22} \bullet V_1 - M_{12} \bullet V_2\right) / \Delta \\ B = \left(M_{11} \bullet V_2 - M_{21} \bullet V_1\right) / \Delta \\ \Delta = M_{11} \bullet M_{22} - M_{12} \bullet M_{21} \end{cases} \qquad \begin{cases} \sigma_A^2 = \dfrac{M_{22}}{\Delta} \\ \sigma_B^2 = \dfrac{M_{11}}{\Delta} \end{cases}$$

$$M_{11} = \sum_{i=1}^{N} \frac{1}{\sigma_i^2}, \; M_{12} = \sum_{i=1}^{N} \frac{x_i}{\sigma_i^2}, \; V_1 = \sum_{i=1}^{N} \frac{y_i}{\sigma_i^2}$$

$$M_{21} = \sum_{i=1}^{N} \frac{x_i}{\sigma_i^2}, \; M_{22} = \sum_{i=1}^{N} \frac{x_i^2}{\sigma_i^2}, \; V_2 = \sum_{i=1}^{N} \frac{x_i y_i}{\sigma_i^2}$$

Lab 4: γ decay of ^{137}Ba



$$D = D_0 e^{-t/\tau}$$

Half life: $D = 0.5\, D_0$ ⟹ $\tau_{1/2} = \tau \ln 2$ Show derivation in your report

Linear regression: $y = A + B \bullet t$

$$\ln D = \ln D_0 - t/\tau \Rightarrow \begin{cases} A = \ln D_0 \\ B = -\dfrac{1}{\tau} \end{cases}$$

Uncertainty is not constant!

Uncertainty of decay count D_i (Poisson):

$$\sigma_{D_i} = \sqrt{D_i}$$

At time progresses, D_i is getting smaller and smaller.

What is the uncertainty of $\ln D_i$?

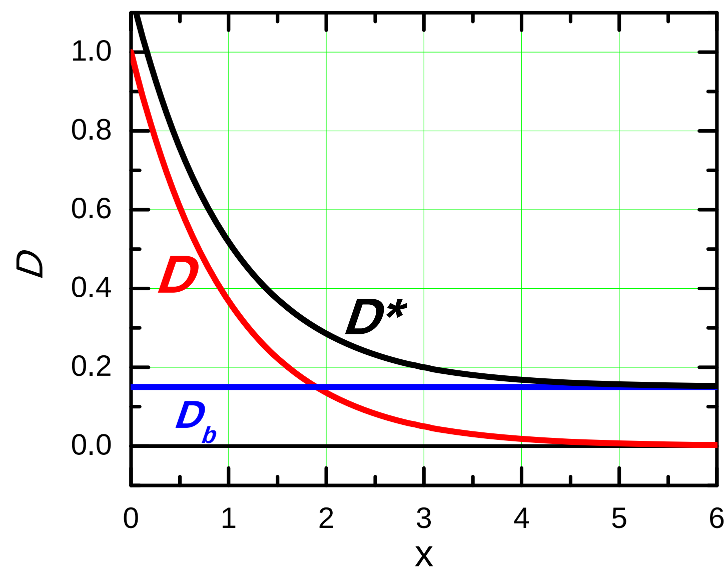
$$\sigma_{(\ln D_i)} = \left|\frac{\partial \ln D_i}{\partial D_i}\right| \bullet \sigma_{D_i} = \frac{\sigma_{D_i}}{D_i} = \frac{1}{\sqrt{D_i}}$$

The background radiation

Background radiation is the radiation constantly present in the natural environment of the Earth, which is emitted by natural and artificial sources.

$$D_i^* = D_i + D_b \Rightarrow D_i = D_i^* - D_b$$

- Sources in the Earth.
- Sources from outer space, such as cosmic rays.
- Sources in the atmosphere, such as the radon gas released from the Earth's crust.



Lab 4: Least Square Fitting of decay counts

1. One run of decay counts/interval (D) vs. time (t)
   a. Must start counting shortly after sample is loaded
   b. Sampling rate: 10 second/sample (Δt)
   c. Run time: 600 seconds (# of measurements: n=60).
2. Measurement of background radiation
   a. Wait 20-30 minutes
   b. Repeat the counting experiment in 1.
   c. Make sure no other radioactive sources near your counter
3. Analyze data
   – Subtract background $D=D^*-D_b$, error propagation.
   – plot D vs. t and ln$D$ vs. t
   – Least square fit and overlap your fitting curves with data plots.

* "Origin" (OriginLab®) is more convenient than Matlab.