Fitting data to models

1

$$y_0(x) \equiv y(\underset{\text{variables}}{x}, \underbrace{a_1 \cdots a_m}_{\text{parameters}}) \Rightarrow \underset{\text{observations}}{y_i}$$

$$\underbrace{\ln[A]_i}_{y_i} = \underbrace{\ln[A]_0}_{a_1} - \underbrace{k}_{a_2}\,\underbrace{t_i}_{x_i} + \underset{\text{error}}{\varepsilon_i}$$

Assumptions of normal distribution of errors:

$$P(y_i - y(x_i)) \approx \exp\left[-\frac{1}{2}\left(\frac{y_i - y(x_i)}{\sigma_i}\right)^2\right]$$

Maximum liklihood says minimize

$$\sum_i \left(\frac{y_i - y(x_i)}{\sigma_i}\right)^2 \equiv \chi^2(a_1, a_2)$$

i.e. do $\left\{\begin{matrix}\text{least squares} \\ \text{chi-square}\end{matrix}\right\}$ fitting

General linear regression

2

$$\underset{\text{observations}}{\underset{\sim}{y}} = \underset{\text{independent variables}}{\underset{\approx}{X}}\,\underset{\text{parameters}}{\underset{\sim}{a}} + \underset{\text{errors}}{\underset{\sim}{\varepsilon}}$$

e.g. straight line

$$y_i = a_1 + a_2 x_i + \varepsilon_i$$

$$\begin{bmatrix} y_1 \\ \vdots \\ \\ y_N \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_N \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_N \end{bmatrix}$$

\# of columns of $\underset{\approx}{X}$ = \# rows of $\underset{\sim}{a}$

= \# of ~~variables~~ unknown, or fitting parameters (including "constant")

"linear regression"

Least-squares procedure: minimize

$$\chi^2 = (\underset{\sim}{y} - \underset{\approx}{X}\underset{\sim}{a})^T (\underset{\sim}{y} - \underset{\approx}{X}\underset{\sim}{a})$$

$$\underset{\sim}{0} = \frac{\partial \chi^2}{\partial \underset{\sim}{a}} = -\underset{\approx}{X}^T(\underset{\sim}{y} - \underset{\approx}{X}\underset{\sim}{\hat{a}}) - (\underset{\sim}{y} - \underset{\approx}{X}\underset{\sim}{\hat{a}})^T \underset{\approx}{X}$$

$\underset{\sim}{\hat{a}}$ is the maximum likelihood estimator of the (unknown, "true") $\underset{\sim}{a}$

$$\underset{\sim}{0} = -\underset{\approx}{X}^T \underset{\sim}{y} + (\underset{\approx}{X}^T \underset{\approx}{X})\underset{\sim}{\hat{a}}$$

$$\boxed{\underset{\sim}{\hat{a}} = (\underset{\approx}{X}^T \underset{\approx}{X})^{-1} \underset{\approx}{X}^T \underset{\sim}{y}}$$

what if the inverse doesn't exist?

See section 15.4 of Numerical Recipes

$$\underset{\sim}{y} = \underset{\approx}{X}\underset{\sim}{a} + \underset{\sim}{\varepsilon} \qquad (1)$$

$$\underset{\sim}{\hat{a}} = (\underset{\approx}{X}^T \underset{\approx}{X})^{-1} \underset{\approx}{X}^T \underset{\sim}{y} \qquad (2)$$

$$\underset{\sim}{\hat{a}} = (\underset{\approx}{X}^T \underset{\approx}{X})^{-1} \underset{\approx}{X}^T \underset{\approx}{X}\underset{\sim}{a} + (\underset{\approx}{X}^T \underset{\approx}{X})^{-1} \underset{\approx}{X}^T \underset{\sim}{\varepsilon}$$

$$\overline{(\underset{\sim}{\hat{a}} - \underset{\sim}{a})} = \overline{(\underset{\approx}{X}^T \underset{\approx}{X})^{-1} \underset{\approx}{X}^T \underset{\sim}{\varepsilon}}$$

"unbiased"

Then if $\overline{\underset{\sim}{\varepsilon}} = \underset{\sim}{0} \Rightarrow \overline{(\underset{\sim}{\hat{a}} - \underset{\sim}{a})} = \underset{\sim}{0}$

$$\overline{(\underset{\sim}{\hat{a}} - \underset{\sim}{a})(\underset{\sim}{\hat{a}} - \underset{\sim}{a})^T} = (X^T X)^{-1} X^T \underbrace{\overline{\varepsilon \varepsilon^T}} X (X^T X)^{-1}$$

"variance-covariance matrix"

assume $= \sigma^2 \underset{\approx}{1}$ (constant exp. error)

$$= \sigma^2 (\underset{\approx}{X}^T \underset{\approx}{X})^{-1}$$

5

For example, for a straight line:

$$y = mx + b$$

The variance-covariance matrix would have:

uncertainty in the slope

$$\begin{bmatrix} \sigma_m^2 & \sigma_{mb} \\ \sigma_{mb} & \sigma_b^2 \end{bmatrix}$$

coupling of these uncertainties

uncertainty in the intercept

For an input x, best estimate of y is

$$y = mx + b$$

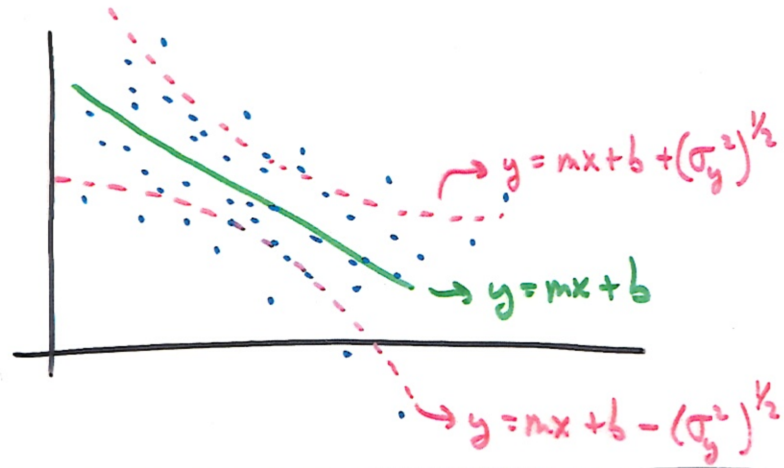
The expected uncertainty in this estimate:

$$\sigma_y^2 = \left(\frac{\partial y}{\partial m}\right)^2 \sigma_m^2 + \left(\frac{\partial y}{\partial b}\right)^2 \sigma_b^2 + 2\left(\frac{\partial y}{\partial m}\right)\left(\frac{\partial y}{\partial b}\right) * \sigma_{mb}$$

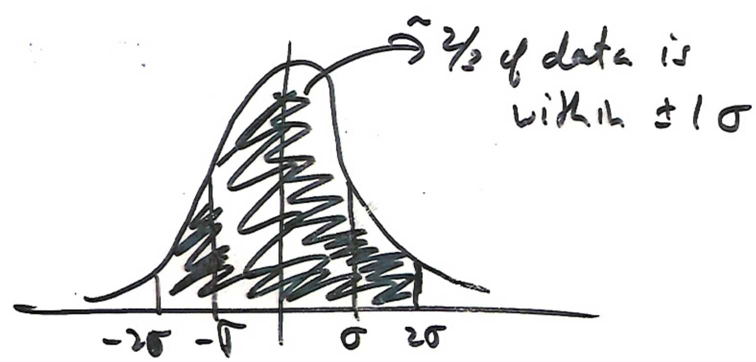
$$= x^2\sigma_m^2 + \sigma_b^2 + 2x\sigma_{mb}$$

6

This is a parabola in x:



rule of thumb: 95% confidence level is about $\pm 2\sigma$



"Robust" estimation

reduce impact of outliers $\rightarrow 1/\sigma_i^2$

1. least squares $\sum_i W_i(\text{calc} - \text{obs})_i^2$

   (Gaussian errors) $e^{-(x-\bar{x})^2/\sigma^2}$

2. mean abs. deviations: $\sum_i W_i \,|\text{calc} - \text{obs}|_i$

   (double exponential errors) $e^{-|x-\bar{x}|/\sigma}$

3. Cauchy: $\sum_i W_i \log\left(1 + \frac{1}{2}z^2\right) \quad z \equiv (\text{calc} - \text{obs})_i$

   (Lorentzian errors)

Numerical Recipes