Introduction to Statistics and Error Analysis III

Physics 116C, 4/26/06 rev. 1

D. Pellett

References:
Data Reduction and Error Analysis for the Physical Sciences by Bevington and Robinson
Particle Data Group notes on probability and statistics, etc.—online at http://pdg.lbl.gov/2004/reviews/contents_sports.html
(Reference is S. Eidelman et al, Phys. Lett. B 592, 1 (2004))

Any original presentations copyright D. Pellett 2006
Topics

• Least squares fit with variable errors on points
  • Likelihood function and chi-square (again)
  • Minimization of chi-square with respect to parameters
  • Formulas for parameters
  • Use of propagation of errors to find errors of parameters
  • Chi-square distribution and significance of fit
  • Use of numerical method to minimize chi-square directly
    • Levenberg-Marquardt method and LabVIEW
• Next - sampled signals and Finite Fourier Transform
Method of Maximum Likelihood (Again)

- Have data with Gaussian errors. Errors vary from point to point.
- Fit data with \( y(x) = a + bx \) using method of maximum likelihood.
- Probability density function for each point; expectation \( y(x_i) \) is on fitted line:

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

- Joint probability (unnormalized) of fitted curve (called likelihood):

\[
L(a, b) = \prod_{i=1}^{N} p_i = \left( \prod_{i=1}^{N} \frac{1}{\sigma_i \sqrt{2\pi}} \right) \exp \left[ -\frac{1}{2} \sum_{i=1}^{N} \left( \frac{y_i - y(x_i)}{\sigma_i} \right)^2 \right]
\]

- Maximize likelihood by minimizing sum in (negative) exponential w.r.t. \( a \) and \( b \):

\[
\chi^2 = \sum_{i=1}^{N} \left( \frac{y_i - y(x_i)}{\sigma_i} \right)^2 = \sum_{i=1}^{N} \left[ \frac{1}{\sigma_i} (y_i - a - bx_i) \right]^2
\]
Minimize $\chi^2$ by setting partial derivatives equal to zero:

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

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

Rearrange sums to get two equations in two unknowns:

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

$$\left( \sum \frac{x_i}{\sigma_i^2} \right) a + \left( \sum \frac{x_i^2}{\sigma_i^2} \right) b = \sum \frac{x_i y_i}{\sigma_i^2}.$$ 

The solution is:

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

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

where

$$\Delta = \sum \frac{1}{\sigma_i^2} \sum \frac{x_i^2}{\sigma_i^2} - \left( \sum \frac{x_i}{\sigma_i^2} \right)^2.$$
Use Propagation of Errors to Find $\sigma_a, \sigma_b$

The fitted parameters $a$ and $b$ depend on the measured points:

$$a = f_a (y_1, y_2, \ldots, y_N), \quad b = f_b (y_1, y_2, \ldots, y_N)$$

where $f_a$ and $f_b$ are given on the previous page, so

$$\sigma_a^2 = \sum \left[ \sigma_i^2 \left( \frac{\partial f_a}{\partial y_i} \right)^2 \right], \quad \sigma_b^2 = \sum \left[ \sigma_i^2 \left( \frac{\partial f_b}{\partial y_i} \right)^2 \right].$$

Now do the work:

$$\frac{\partial a}{\partial y_j} \equiv \frac{\partial f_a}{\partial y_j} = \frac{1}{\Delta} \left( \frac{1}{\sigma_j^2} \sum_{i=1}^{N} \frac{x_i^2}{\sigma_i^2} - \frac{x_j}{\sigma_j^2} \sum_{i=1}^{N} \frac{x_i}{\sigma_i^2} \right)$$

$$\frac{\partial b}{\partial y_j} \equiv \frac{\partial f_b}{\partial y_j} = \frac{1}{\Delta} \left( \frac{x_j}{\sigma_j^2} \sum_{i=1}^{N} \frac{1}{\sigma_i^2} - \frac{1}{\sigma_j^2} \sum_{i=1}^{N} \frac{x_i}{\sigma_i^2} \right).$$

Square, combine, sum, do algebra (exercise left for student...):

$$\sigma_a^2 = \frac{1}{\Delta} \sum \frac{x_i^2}{\sigma_i^2}, \quad \sigma_b^2 = \frac{1}{\Delta} \sum \frac{1}{\sigma_i^2}.$$

What about covariances?

What is the Error Matrix?
Recast in Matrix Notation (See Bevington, Ch.7)

Consider fitting functions linear in the parameters

→ \( y = a_1 \exp(-a_2 \, t) \) is not linear (in \( a_2 \)) but \( y = a_1 + a_2 \, x + a_3 \, x^2 \) is linear in \( a \)'s

as is Steinhart-Hart Equation for thermistor: \( y = 1/T = A + B \ln R + D(\ln R)^3 \)

Done in Ch. 7. Apply to our simple straight line fit to see how it works:

For example \( \chi^2 = \sum \frac{1}{\sigma_i^2} \left[ y_i - (a_1 - a_2 \, x_i) \right]^2 \)

\( a_1 \equiv a \), \( a_2 \equiv b \)

Fitting function \( y(x_i) = \sum_{k=1}^{m} a_k \, f_k(x_i); \, f_1 = 1, \, f_2 = x_i \) \( (m = 2) \)

Minimizing \( \chi^2 \) leads to a set of \( m \) coupled linear equations with \( l \) running from 1 to \( m \).

\[ \sum \frac{y_i \, f_k(x_i)}{\sigma_i^2} = \sum_{k=1}^{m} \left\{ a_k \sum_{i=1}^{m} \left[ \frac{1}{\sigma_i^2} \, f_k(x_i) \, f_k(x_i) \right] \right\} \]

In our case \( (m = 2) \)

\( l=1, \, f_1 = 1 \) \( \sum y_i \frac{1}{\sigma_i^2} = a_1 \sum \frac{1}{\sigma_i^2} + a_2 \sum \frac{1}{\sigma_i^2} \, x_i \)

\( l=2, \, f_2 = x_i \) \( \sum y_i \frac{x_i}{\sigma_i^2} = a_1 \sum \frac{x_i}{\sigma_i^2} + a_2 \sum \frac{1}{\sigma_i^2} \, x_i^2 \)

This can be written in matrix form:

\[ [\beta] = [A] [\alpha] \]

\[ [A] = \begin{bmatrix} a_1 & a_2 \\ 1 & x_i \end{bmatrix} \]

\[ [\alpha] = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \]

\( [\sigma] = \begin{bmatrix} \sigma_1 & \sigma_2 \\ \sigma_2 & \sigma_2 \end{bmatrix} \) (symmetric)
Matrix Solution (Continued)

\[ \beta_k = \sum \frac{1}{\sigma_i^2} y_i f_k(x_i) \]

\[ \alpha_{lk} = \sum \left[ \frac{1}{\sigma_i^2} f_l(x_i) f_k(x_i) \right] \]

\[ [\beta] = [\alpha] [\alpha]^{-1} \]

Solve by multiplying both sides by \([\epsilon] = [\alpha]^{-1}\) (inverse of \([\alpha]\))


\([\alpha]\) called curvature matrix

Note \[ \frac{\partial^2 \chi^2}{\partial \alpha_l \partial \alpha_k} = 2 \alpha_{lk} \]

(related to curvature of \(\chi^2\) in parameter space)
(see Bovingrani)
Error Matrix (Covariance Matrix)

\[ \Sigma = \mathbf{A}^{-1} \]
called error matrix

\[ \sigma_{ajal}^2 = \sum \left[ \sigma^2 \frac{\partial q_i}{\partial y_i} \frac{\partial q_l}{\partial y_i} \right] \]  (covariance matrix)

Can show \[ \sigma_{ajal}^2 = \epsilon_{jl} \]

Again, see Bevington, Ch. 7.

Elements of error matrix are covariances of the parameters.

- Can calculate chi-square, use chi-square distribution with \( V \) degrees of freedom to test goodness of fit, where \( V = (\text{no. points} - \text{no. parameters}) \)
Can Minimize Chi-Square Numerically (e.g., Lev-Mar VI) to Find a, b, Errors

- Cleverly searches parameter space for minimum chi-square → see Bevington, Ch. 8; Levenberg-Marquardt method in Sec. 8.6
- Evaluates curvature matrix and covariance (error) matrix numerically
- Calculates “mean square error” = chi-square/number of points
- Must provide fitting function (and derivatives w.r.t. parameters) as sub-VI:
Example: Levenberg-Marquardt Fit to Straight Line

• This is for the data of the first slide, fit $y = a_1 + a_2 \times$

• Finds $a_1 = -29, a_2 = 0.099$

• $\sigma_{11}^2 = 100.7, \sigma_{22}^2 = 7.176E-5$

• $\sigma_{a1} = 10, \sigma_{a2} = 0.0085$

• covariances non-negligible

• $\chi^2/N = 0.55 \ (N = \text{number of points})$

• $\chi^2/\nu = \langle \chi^2 \rangle = (6/4)(0.55) = 0.82 \ (\text{good fit})$

• Covariance Matrix found numerically may be misleading if the fit is not linear in the parameters or if the experimental errors are not Gaussian
Cautions on Covariance Matrix, $\chi^2$ from Numerical Fit for Non-Gaussian Errors

• Summary of comments in Press, et al., *Numerical Recipes*, Sec. 14.5:

• The procedure is fine for finding “best fit” parameters, finding contours of constant $\Delta \chi^2$ as a confidence region boundary (with further work needed to give actual confidence levels)

• **But** Covariance Matrix elements are **not** necessarily (co)variances of parameters $\sigma_{ij}$ and $\chi^2$ result does **not** necessarily follow $\chi^2$ distribution

• One **can** use $\sigma_{ij}$ and $\chi^2$ distribution if (i) the measurement errors are normally distributed and either (ii) the fitting function is linear in its parameters or (iii) the sample size is large enough that the uncertainties in the fitted parameters do not extend outside a region in which the model could be replaced by a suitable linearized model

• Hence the previous fit error and $\chi^2$ interpretations were **OK** (Gaussian errors, fitting function is linear in its parameters)

• *If conditions are not met*, one can say that the covariance matrix is the “formal covariance matrix of the fit on the assumption of normally distributed errors”