Outline

1. Errors
2. Distributions
3. Computing Distributions
4. Error Propagation
5. Errors with Gaussian Distributions
6. Fitting a Straight Line
7. Fitting a Linear Model

Overview

- fitting: compare measurements with model predictions
- fitting method depends on
  - errors in observations
  - how model depends on free parameters
  - definition of *best fit*
- to determine errors in observations, need to propagate errors through data acquisition and reduction
- need to define what

Accuracy and Precision

- *accuracy* of observation measures correctness of result, measures of how close observational result comes to true value
- *precision* of observation measures how reproducible result is, measures how exactly the result is determined without reference to what that result means
- *absolute precision*: magnitude of uncertainty in result in same units as result
- *relative precision*: uncertainty in terms of fraction of value of result
- both accuracy and precision need to be considered simultaneously
- useless to determine something with high precision but highly inaccurately
- observation cannot be considered accurate if precision is low

Random and Systematic Errors

- *systematic error*: reproducible inaccuracy introduced by faulty equipment, calibration, or technique
- accuracy generally depends on how well one can control or compensate for *systematic errors*
- *random error*: Indefiniteness of result due to *a priori* finite precision of observation, measures fluctuation in repeated observations
- precision depends on how well one can overcome or analyse *random errors*
- random errors require repeated trials to yield precise results
- given accuracy implies a precision at least as good $\Rightarrow$ depends somewhat on random errors

Characterizing Distributions

- *parent distribution*: infinite number of measurements $\Rightarrow$ observations distributed according to *true* probability distribution
- actual observations are *sample* of infinite number of possible measurements
- observations *estimate* parameters of parent distribution
- average:

$$\mu \simeq \overline{x} \equiv \frac{1}{N}\sum_{i=1}^{N} x_i$$

- variance:

$$\sigma^2 \simeq s^2 \equiv \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2 = \frac{1}{N-1}\left(\sum_{i=1}^{N} x_i{}^2 - N\overline{x}^2\right)$$

Independent Measurements

- estimate of average assumes that all measurements x_i are independent
- for variance, already used x_i values to estimate average $\Rightarrow N-1$ independent measurements left
- reason that sum of $(x_i - \overline{x})^2$ is divided by $N-1$
- single measurement ($N = 1$)
  - best value for average given by single measurement
  - no measure for variance

Sample Average

- consider $\overline{x}$ as variable in definition of variance
- try to find value of $\overline{x}$ for which s^2 is minimal
- set derivative of s^2 with respect to $\overline{x}$ to zero:

$$\frac{\partial s^2}{\partial \overline{x}} = \frac{-2}{N-1}\sum_{i=1}^{N}(x_i - \overline{x}) = \frac{-2}{N-1}\left(\sum_{i=1}^{N} x_i - N\overline{x}\right) = 0$$

- therefore

$$\overline{x} \equiv \frac{1}{N}\sum_{i=1}^{N} x_i$$

- definition for sample average minimizes sample variance

Average and Variance for Binned Data

- astronomical data often *binned* or discrete
- all digital data is binned
- definitions

  B number of bins, $b = 1, \ldots, B$
  x_b value of x in bin b
  N_b number of measurements in bin b
- normalize N_b by total number of measurements
- probability that measurement falls into bin b is $P_b \equiv N_b / \sum_{b=1}^{B} N_b$
- average:

$$\overline{x} = \sum_{b=1}^{B} P_b x_b$$

- variance:

$$s^2 = \frac{N}{N-1} \left(\sum_{b=1}^{B} P_b {x_b}^2 - \overline{x}^2 \right)$$

Higher-Order Moments of Distributions

- higher-order moments of distributions very sensitive to outliers
- large $x_i - \overline{x}$ value dominates much more in distribution of $(x_i - \overline{x})^2$ than in distribution of $|x_i - \overline{x}|$
- therefore do not use even higher moments such as
  - Skewness:

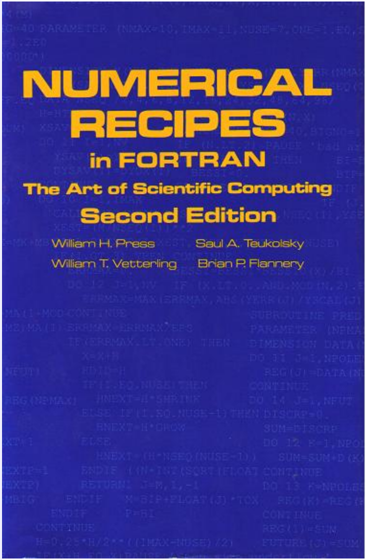
$$\equiv \frac{1}{N\sigma^3} \sum (x_i - \overline{x})^3$$

  - Kurtosis

$$\equiv \frac{1}{N\sigma^4} \sum (x_i - \overline{x})^4 - 3$$

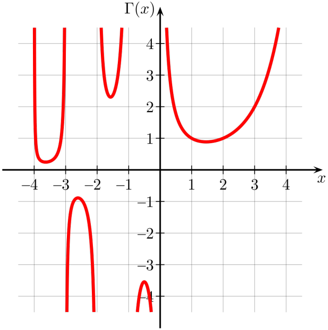
  - σ is standard deviation
  - subtraction of 3 in kurtosis makes kurtosis of Gaussian zero

Numerical Recipes



- funamental book on numerical algorithms
- exists for different programming languages
- 2nd edition available online at www.nr.com/oldverswitcher.html
- Chapter 6 contains algorithms to calculate special functions

Gamma Function



commons.wikimedia.org/wiki/File:Gamma-function.svg

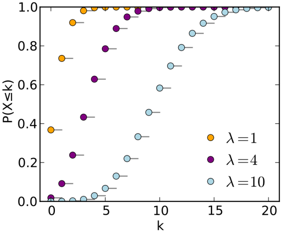
- $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$
- integer z: gamma function equals factorial with offset of one

$$n! = \Gamma(n+1)$$

Factorial

- small $n \Rightarrow$ compute factorial directly as $n \times (n-1) \times (n-2) \ldots$
- large values $\Rightarrow$ gamma-function
- large n, factorial larger than largest number allowed by computer
  - single precision (IEEE 754 32-bit decimal) has maximum exponents of -126, +127
  - double precision (EEE 754 64-bit declimal) has maximum exponents of -1022, 1023
  - calculate logarithm of factorial or logarithm of gamma function
- useful function routines from Numerical Recipes:
  - `function gammln(x)` returns $\ln \Gamma(x)$ for input x
  - `function factrl(n)` returns real $n!$ for input integer n
  - `function factln(n)` returns real $\ln n!$ for input integer n
  - `function bico(n,k)` returns real $\binom{n}{k}$ for integer inputs n, k
- useful to compute large number of frequently used distributions

Cummulative Poisson Distribution



en.wikipedia.org/wiki/Poisson_distribution

- cumulative Poisson probability describes probability that Poisson process will lead to result between 0 and $k-1$ inclusive:

$$P_x(<k) \equiv \sum_{n=0}^{k-1} P_P(k,x)$$

- *incomplete gamma function*:

$$P(a,x) \equiv \frac{1}{\Gamma(a)} \int_0^x t^{a-1} e^{-t} dt$$

Cummulative Poisson Distribution (continued)

- complement also called incomplete gamma function

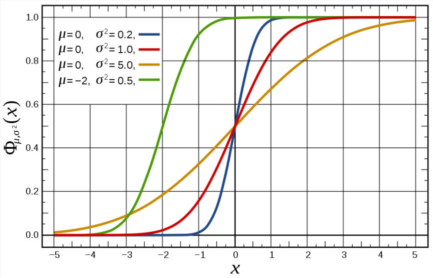
$$Q(a,x) \equiv 1 - P(a,x) \equiv \frac{1}{\Gamma(a)} \int_x^\infty t^{a-1} e^{-t} dt$$

- cumulative Poisson probability:

$$P_x(< k) = Q(k,x)$$

- corresponding routines in Numerical Recipes are:
  - `function gammp(a,x)` returns $P(a,x)$ for input a, x
  - `function gammq(a,x)` returns $Q(a,x)$ for input a, x

Cummulative Gauss Distribution



en.wikipedia.org/wiki/Normal_distribution

- integral probability of Gauss function from *error function*:

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

- *complementary error function*:

$$\mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt$$

Cummulative Gauss Distribution (continued)

- error functions given by incomplete gamma functions:

$$\mathrm{erf}(x) = P(1/2, x^2) \qquad (x \geq 0)$$
$$\mathrm{erfc}(x) = Q(1/2, x^2) \qquad (x \geq 0)$$

- corresponding routines in Numerical Recipes are:
  - `function erf(x)` returns erf(x) for input x, using `gammp`
  - `function erfc(x)` returns erfc(x) for input x, using `gammq`
  - `function erfcc(x)` returns erfc(x) based on direct series development

Error Propagation

Basics

- function f depends on variables $u, v, \ldots$:

$$f \equiv f(u, v, \ldots)$$

- estimate variance of f

$$\sigma_f{}^2 \equiv \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} (f_i - \overline{f})^2$$

  knowing variances $\sigma_u, \sigma_v, \ldots$ of variables $u, v, \ldots$

- assumption, usually only approximately correct, that average of f is well approximated by value of f for averages of variables:

$$\overline{f} = f(\overline{u}, \overline{v}, \ldots)$$

Basics (continued)

- Taylor expansion of f around average:

$$f_i - \overline{f} \simeq (u_i - \overline{u})\frac{\partial f}{\partial u} + (v_i - \overline{v})\frac{\partial f}{\partial v} + \ldots$$

- variance in f:

$$\sigma_f{}^2 \simeq \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N}\left[(u_i - \overline{u})\frac{\partial f}{\partial u} + (v_i - \overline{v})\frac{\partial f}{\partial v} + \ldots\right]^2$$

$$= \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N}\Big[\quad (u_i - \overline{u})^2\left(\frac{\partial f}{\partial u}\right)^2 + (v_i - \overline{v})^2\left(\frac{\partial f}{\partial v}\right)^2 +$$

$$2(u_i - \overline{u})(v_i - \overline{v})\frac{\partial f}{\partial u}\frac{\partial f}{\partial v} + \ldots\Big]$$

Basics (continued)

- variances of u and v

$$\sigma_u{}^2 \equiv \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} (u_i - \overline{u})^2; \qquad \sigma_v{}^2 \equiv \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} (v_i - \overline{v})^2$$

- covariance of u and v

$$\sigma_{uv}{}^2 \equiv \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} (u_i - \overline{u})(v_i - \overline{v})$$

- use these defintions to obtain

$$\sigma_f{}^2 = \sigma_u{}^2 \left(\frac{\partial f}{\partial u}\right)^2 + \sigma_v{}^2 \left(\frac{\partial f}{\partial v}\right)^2 + 2\sigma_{uv}{}^2 \frac{\partial f}{\partial u}\frac{\partial f}{\partial v} + \dots$$

Basics (continued)

- from before

$$\sigma_f{}^2 = \sigma_u{}^2 \left(\frac{\partial f}{\partial u}\right)^2 + \sigma_v{}^2 \left(\frac{\partial f}{\partial v}\right)^2 + 2\sigma_{uv}{}^2 \frac{\partial f}{\partial u}\frac{\partial f}{\partial v} + \dots$$

- if differences $u_i - \overline{u}$ and $v_i - \overline{v}$ not correlated $\Rightarrow$ sign of product as often positive as negative $\Rightarrow$ covariance small compared to other terms
- if differences are correlated $\Rightarrow$ most products $(u_i - \overline{u})(v_i - \overline{v})$ positive $\Rightarrow$ cross-correlation term can be large

Examples of Error Propagation

Weighted Sum: $f = au + bv$

- partial derivatives

$$\frac{\partial f}{\partial u} = a, \qquad \frac{\partial f}{\partial v} = b$$

- variance

$$\sigma_f{}^2 = a^2\sigma_u{}^2 + b^2\sigma_v{}^2 + 2ab\sigma_{uv}{}^2$$

- a and b can be positive or negative
- signs only affect cross-correlation term
- cross-correlation term can be negative $\Rightarrow$ makes variance smaller
- example: if each u_i is accompanied by a v_i such that $v_i - \overline{v} = -(b/a)(u_u - \overline{u})$, then $f = a\overline{u} + b\overline{v}$ for all u_i, v_i pairs, and $\sigma_f{}^2 = 0$

Product: $f = auv$

- partial derivative

$$\frac{\partial f}{\partial u} = av, \qquad \frac{\partial f}{\partial v} = au$$

- variance

$$\frac{{\sigma_f}^2}{f^2} = \frac{{\sigma_u}^2}{u^2} + \frac{{\sigma_v}^2}{v^2} + \frac{2{\sigma_{uv}}^2}{uv}$$

Division: $f = au/v$

- partial derivatives

$$\frac{\partial f}{\partial u} = \frac{a}{v}, \qquad \frac{\partial f}{\partial v} = -\frac{au}{v^2}$$

- variance

$$\frac{{\sigma_f}^2}{f^2} = \frac{{\sigma_u}^2}{u^2} + \frac{{\sigma_v}^2}{v^2} - \frac{2{\sigma_{uv}}^2}{uv}$$

Exponent: $f = ae^{bu}$

- partial derivatives

$$\frac{\partial f}{\partial u} = bf$$

- variance

$$\frac{\sigma_f}{f} = b\sigma_u$$

Power: $f = au^b$

- partial derivatives

$$\frac{\partial f}{\partial u} = \frac{bf}{u}$$

- variance

$$\frac{\sigma_f}{f} = b\frac{\sigma_u}{u}$$

Fitting Observations with Gaussian Error Distributions

Least Squares Method

- series of measurements y_i with associated errors distributed according to Gaussian with width σ_i
- same as each measurement drawn from Gaussian with width σ_i around model value y_m
- probability $P(y_i) \equiv P_i$ of obtaining a single measurement y_i in interval Δy given by

$$P_i \Delta y = \frac{1}{\sqrt{2\pi}\sigma_i} e^{\frac{-(y_i - y_m)^2}{2{\sigma_i}^2}} \Delta y$$

- different measurements have different associated errors σ_i

Least Squares Method (continued)

- probability P of obtaining series of N measurements

$$P(\Delta y)^N \equiv \prod_{i=1}^{N}(P_i \Delta y) = \frac{1}{(2\pi)^{N/2}\prod_i \sigma_i} \exp\left[-\frac{1}{2}\sum_{i=1}^{N}\frac{(y_i - y_m)^2}{{\sigma_i}^2}\right]\Delta y^N$$

- highest probability P for smallest

$$\chi^2 \equiv \sum_{i=1}^{N} {\chi_i}^2 \equiv \sum_{i=1}^{N}\frac{(y_i - y_m)^2}{{\sigma_i}^2}$$

- determine most probable model value for y_m for series of measurements y_i by finding value(s) for y_m for which sum of squares $(y_i - y_m)^2/{\sigma_i}^2$ is minimal
- *method of least squares*, first described by Gauss

Chi-Squared Distribution

- if errors are gaussian, each χ_i is random draw from normal distribution
- sum of $\chi_i{}^2$ is called *chi-square*
- N measurements fit by model with M parameters: $N - M$ independent measurements
- probability distribution for χ^2 is *chi-square distribution* for $N - M$ *degrees of freedom*
- distribution obtained by drawing $N - M$ random samples from normal distribution and add squares
- probability of given χ^2 from `gammq`-function
- probability that observed χ^2_{obs} or greater is reached for $N - M$ degrees of freedom is $P(\chi^2_{obs}) = \mathtt{gammq}(0.5(N - M), 0.5\chi^2_{obs})$

Chi-Squared Distribution (continued)

- if probability is very small $\Rightarrow$ something is wrong
  - wrong model
  - errors underestimated
  - errors not distributed as Gaussians
- probability of 0.05 is often acceptable
- wrong models produce much smaller probabilities (< 0.000001)
- probability of 5% occurs, on average, once every 20 trials
- finding 0.05 probability due to chance quite common
- consistently low probabilities must be investigated

Negative Results

- apparently significant results can arise from ignoring negative results
- lottery winner is person with correct six-digit number
- probability for one person to have winning number is one in a million
- if several million people participate we expect several to have guessed the correct number
- less obvious: repeated experiment, i.e. stock broker
- ten million people predict how stocks change in value
- after one year, select top 10% predictors
- repeat for total of six rounds $\Rightarrow$ (on average) ten brokers will have predicted among top 10% for six years in a row, even if prediction process is purely random!
- must know total number of brokers to decide wether they are better than random

Model Fitting

- fit model to data set, present 3 parts:
  1. best-fit values of parameters $a_1, a_2, \ldots$
  2. errors in these parameters
  3. probability that measured χ^2 is obtained by chance; i.e. the probability that model adequately describes measurements
- 5 cases for y_m
  1. constant (y_m same for all i)
  2. straight line $y_m = a + bx$ where y_m depends on variable x and model parameters a, b
  3. straight line $y_m = a + bx$ where y_m depends on variables x, y, model parameters a, b
  4. linear function $y_m = f(x, a, b, c, \ldots$ where y_m depends on variable x and linearly on model parameters a, b, $\ldots$
  5. general (non-linear) case where y_m depends on variable x and parameters $a_1, a_2, \ldots$: $y_m = y_m(x; \vec{a})$

Weighted Averages

- constant model $y_m = a \Rightarrow a$ best estimate for average $\overline{y}$ of y_i
- most probable value of a by minimizing χ^2 with respect to a

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

- least squares found for

$$\overline{y} \equiv a_{min} = \frac{\sum_{i=1}^{N}\frac{y_i}{{\sigma_i}^2}}{\sum_{i=1}^{N}\frac{1}{{\sigma_i}^2}}$$

Weighted Averages (continued)

- determine estimate of error in $\overline{y}$
- $\overline{y}$ is function of variables $y_1, y_2, \ldots$
- *if measurements y_i not correlated*: variance of $\overline{y}$ from error propagation:

$$\sigma_{\overline{y}}{}^2 = \sum_{i=1}^{N} \left[\sigma_i{}^2 \left(\frac{\partial \overline{y}}{\partial y_i} \right)^2 \right] = \sum_{i=1}^{N} \left[\sigma_i{}^2 \left(\frac{1/\sigma_i{}^2}{\sum_{k=1}^{N} (1/\sigma_k{}^2)} \right)^2 \right]$$

- therefore

$$\sigma_{\overline{y}}{}^2 = \frac{1}{\sum_{i=1}^{N} (1/\sigma_i{}^2)}$$

Identical Errors

- in case where all measurement errors are identical ($\sigma_i \equiv \sigma$)

$$\overline{y} = \frac{(1/\sigma^2)\sum_{i=1}^{N} y_i}{(1/\sigma^2)\sum_{i=1}^{N}(1)} = \frac{1}{N}\sum_{i=1}^{N} y_i$$

- therefore

$$\sigma_{\overline{y}}{}^2 = \frac{\sigma^2}{N}$$

Fitting Straight Line

- straight line: $y_m(x, a, b) = a + bx \Rightarrow y_m(x_i, a, b) = a + bx_i$
- fit a, b by minimizing χ^2 with respect to a, b
- set corresponding derivatives χ^2 to zero:

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

$$\sum_{i=1}^{N} \frac{y_i}{\sigma_i{}^2} - a \sum_{i=1}^{N} \frac{1}{\sigma_i{}^2} - b \sum_{i=1}^{N} \frac{x_i}{\sigma_i{}^2} = 0$$

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

$$\sum_{i=1}^{N} \frac{x_i y_i}{\sigma_i{}^2} - a \sum_{i=1}^{N} \frac{x_i}{\sigma_i{}^2} - b \sum_{i=1}^{N} \frac{x_i{}^2}{\sigma_i{}^2} = 0$$

Fitting Straight Line (continued)

- all sums can be evaluated without knowing *a* or *b*
- define the following sums

$$\sum_{i=1}^{N} \frac{1}{{\sigma_i}^2} \equiv S; \qquad \sum_{i=1}^{N} \frac{x_i}{{\sigma_i}^2} \equiv S_x; \qquad \sum_{i=1}^{N} \frac{{x_i}^2}{{\sigma_i}^2} \equiv S_{xx}$$

$$\sum_{i=1}^{N} \frac{y_i}{{\sigma_i}^2} \equiv S_y; \qquad \sum_{i=1}^{N} \frac{x_i y_i}{{\sigma_i}^2} \equiv S_{xy}; \qquad \Delta \equiv S S_{xx} - (S_x)^2$$

- rewrite as two equations for two unknowns *a* and *b*:

$$aS + bS_x - S_y = 0$$
$$aS_x + bS_{xx} - S_{xy} = 0$$

Fitting Straight Line (continued)

- solutions

$$a = \frac{S_{xx} S_y - S_x S_{xy}}{\Delta}; \quad b = \frac{S S_{xy} - S_x S_y}{\Delta}$$

- errors in a, b from considering a, b to depend on independent parameters y_i

$$\frac{\partial a}{\partial y_i} = \frac{S_{xx} - x_i S_x}{{\sigma_i}^2 \Delta}; \qquad \frac{\partial b}{\partial y_i} = \frac{x_i S - S}{{\sigma_i}^2 \Delta}$$

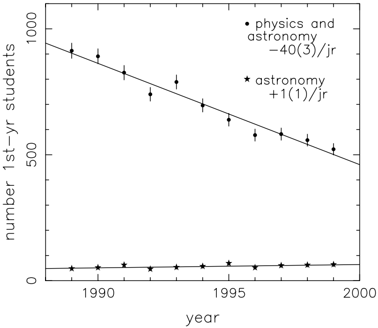
- use error propagation

$$ {\sigma_a}^2 = \frac{S_{xx}}{\Delta}; \qquad {\sigma_b}^2 = \frac{S}{\Delta}$$

- probability that good fit would produce observed ${\chi^2}_{obs}$ or bigger:

$$Q = \mathtt{gammq}\left(\frac{N-2}{2}, \frac{{\chi^2}_{obs}}{2}\right)$$

Straight Line Fitting Example



- number of new physics and astronomy students in Netherlands
- errors in measured integer numbers: square root of number
- actual number in year is drawn from distribution (here Poissonian) around expected value
- same in photon-counting observations

Example (continued)

- good choice of a, b
- number of students as $N(t) = a + bt$ where t is the year $\Rightarrow$ a gives number of students for year 0
- sums involving x_i-values are very large $\Rightarrow$ subtracting them from one another (as in computing Δ) easily leads to roundoff errors
- errors in a and b will be highly correlated: small change in b changes a dramatically in one direction
- prevent both problems by centering time interval around point of fitting, i.e. $N = a + b(t - 1994)$
- avoids round-off errors and correlation of variations are minimized
- good practice in astronomy to define time with respect to some fiducial point near middle of measurements

Linear Models with Errors in Both Coordinates

- x_i may also have errors $\Rightarrow$ minimize

$$\chi^2(a,b) = \sum_{i=1}^{N} \frac{(y_i - a - bx_i)^2}{\sigma_{y_i}^2 + b^2\sigma_{x_i}^2}$$

- weighted sum of variances in denominator from error propagation
- a determined from setting partial derivative to zero

$$a = \left[\sum_{i=1}^{N} \frac{(y_i - bx_i)^2}{\sigma_{y_i}^2 + b^2\sigma_{x_i}^2}\right] \Big/ \sum_{i=1}^{N} \frac{1}{\sigma_{y_i}^2 + b^2\sigma_{x_i}^2}$$

- b determination more complicated because equation becomes non-linear $\Rightarrow$ numerical solution to minimize with respect to b
- at each iteration ensure that minimum with respect to b is also minimized with respect to a
- complicated errors in parameter estimates $\Rightarrow$ use apprach to be discussed for general case

General Linear Least Squares Problem

Linear Models

- model y_i^m is linear combination of M given functions of x
- example: polynomial of degree $M-1$:
  $y(x) = a_1 + a_2 x + a_3 x^2 + \ldots + a_M x^M$
- general form

$$y(x) = \sum_{k=1}^{N} A_k X_k$$

- $X_1(x), \ldots, X_M(x)$ arbitrary (non-linear!) fixed functions of x
- minimize

$$\chi^2 = \sum_{i=1}^{N} \frac{y_i - \sum_{k=1}^{N} a_k X_k(x_i)}{\sigma_i^2}$$

- design matrix

$$A_{ij} = \frac{X_j(x_i)}{\sigma_i}$$

Linear Models (continued)

- in general, A has more rows than columns ($N > M$)
- vector $\vec{b}$ of length N:

$$b_i = \frac{y_i}{\sigma_i}$$

- minimum of chi-squared where derivatives with respect to all M parameters vanishes leads to

$$(A^T A)\vec{a} = A^T \vec{b}$$

- inverse matrix of positive definite matrix $A^T A$

$$C = (A^T A)^{-1}$$

- errors in parameters then given by

$$\sigma^2(a_j) = C_{jj}$$

- off-diagonal elements C_{jk} are covariances between a_j and a_k

Linearizing Models

- apparently nonlinear problems can be linearized
- example: $y(x) = ae^{bx}$ becomes $\log y(x) = c + bx$
- warning: transformations does not make Gaussian errors into Gaussian errors
- warning: watch out for degenerate parameters, e.g. $y(x) = ae^{bx+d}$