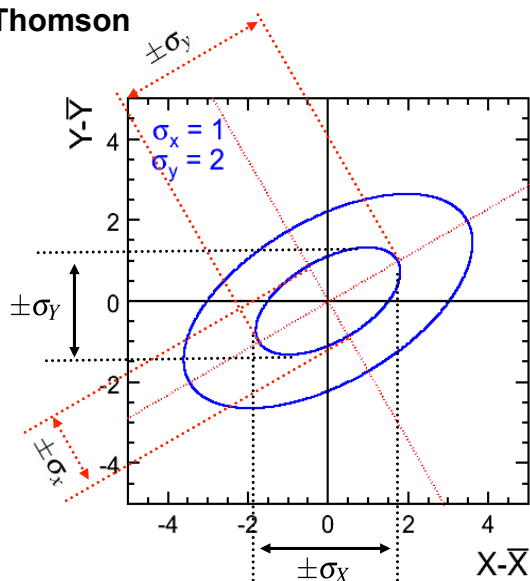
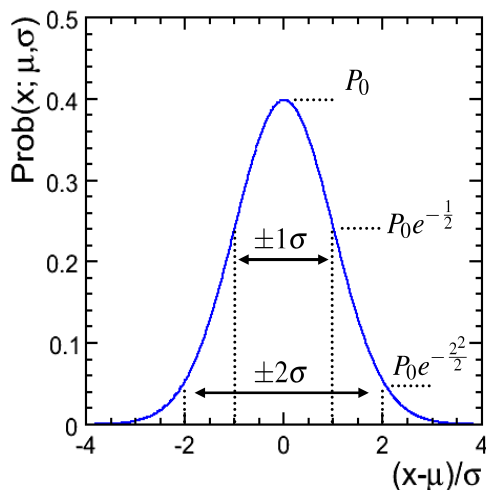


Statistics

Lent Term 2015
Prof. Mark Thomson



Lecture 2 : The Gaussian Limit

Lecture 1: Back to basics

Introduction, Probability distribution functions, Binomial distributions, Poisson distribution

Lecture 2: The Gaussian Limit

The central limit theorem, Gaussian errors, Error propagation, Combination of measurements, Multi-dimensional Gaussian errors, Error Matrix

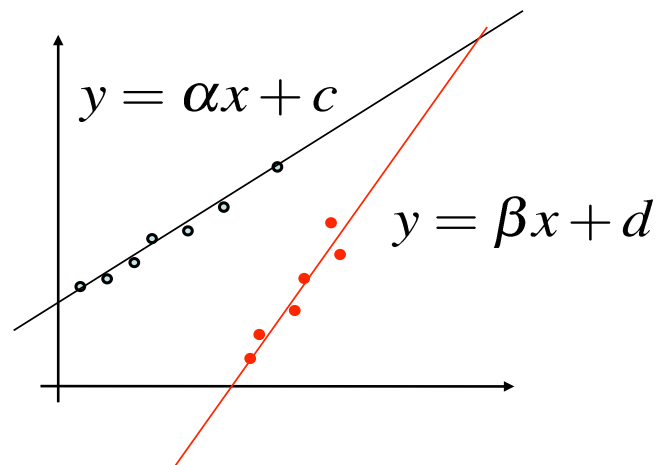
Lecture 3: Fitting and Hypothesis Testing

The χ^2 test, Likelihood functions, Fitting, Binned maximum likelihood, Unbinned maximum likelihood

Lecture 4: Dark Arts

Bayesian statistics, Confidence intervals, systematic errors.

- ★ **Problem:** given the results of two straight line fits with errors, calculate the uncertainty on the intersection



- ★ **Solution:** first learn about
- Gaussian errors
 - Correlations
 - Error propagation

The Central Limit Theorem

- ★ We have already shown that for large μ that a Poisson distribution tends to a Gaussian
- ★ This is one example of a more general theorem, the “**Central Limit Theorem**”*

If n random variables, x_i , each distributed according to **any** PDF, are combined then the sum $y = \sum x_i$ will have a PDF which, for **large** n , tends to a Gaussian

- ★ For this reason the Gaussian distribution plays an important role in statistics

$$G(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(x - \mu)^2}{2\sigma^2} \right\}$$

which by make a suitable coordinate transformation, $x \rightarrow \sigma x + \mu$, gives the **Normal distribution**

$$N(x) = \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{x^2}{2} \right\}$$

Mean = zero
Rms = 1

* The proof of the central limit theorem is non-trivial and isn't reproduced here

A useful integral relationship

- ★ We will often take averages of functions of Gaussian distributed quantities $\langle x^2 \rangle, \langle x^4 \rangle$
- ★ Hence interested in integrals of the form

$$\langle (x - \mu)^n \rangle = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} (x - \mu)^n e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}} \sigma^n \int_{-\infty}^{+\infty} y^n e^{-\frac{y^2}{2}} dy$$

- ★ Define

$$I_n = \int_{-\infty}^{+\infty} x^n e^{-\frac{x^2}{2}} dx$$

$$\text{For } n \text{ odd, } I_n = 0$$

- For even n:

$$\begin{aligned} &= \int_{-\infty}^{+\infty} d(-x^{n-1} e^{-\frac{x^2}{2}}) + (n-1) \int_{-\infty}^{+\infty} x^{n-2} e^{-\frac{x^2}{2}} dx \\ &= \left[-x^{n-1} e^{-\frac{x^2}{2}} \right]_{-\infty}^{+\infty} + (n-1) I_{n-2} \end{aligned}$$

- ★ Hence

$$\frac{I_n}{I_{n-2}} = (n-1) \quad n > 1$$

- ★ By writing

$$\langle (x - \mu)^n \rangle = \frac{\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} (x - \mu)^n e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx}{\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx} \Rightarrow \langle (x - \mu)^n \rangle = \frac{I_n}{I_0} \sigma^n$$

$$\text{e.g. } \langle (x - \mu)^4 \rangle = \frac{I_4}{I_0} \sigma^4 = \frac{I_4}{I_2} \frac{I_2}{I_0} \sigma^4 = (4-1)(2-1) \frac{I_0}{I_0} \sigma^4 = 3\sigma^4$$

Properties of the Gaussian Distribution

- ★ Normalised to unity (it's a PDF)

$$\int_{-\infty}^{+\infty} G(x; \mu, \sigma) dx = 1$$

- Proof:

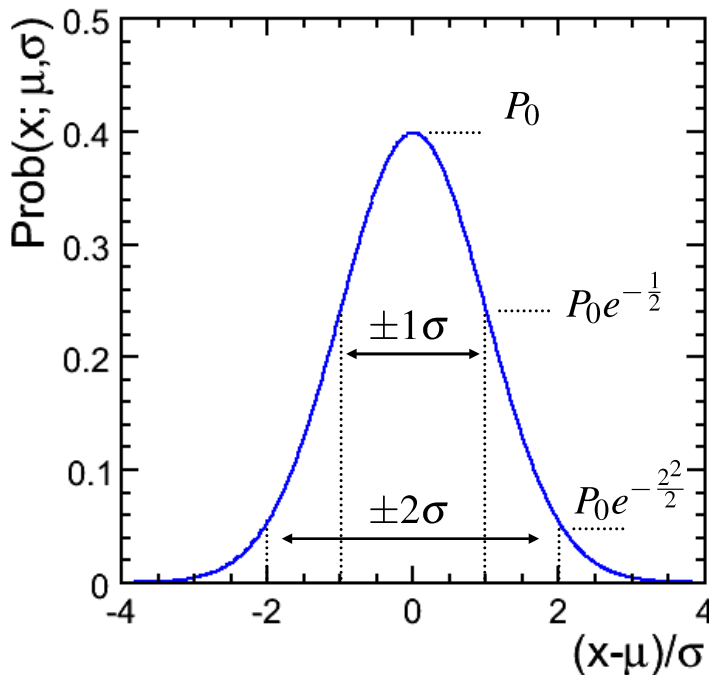
$$\begin{aligned} \int_{-\infty}^{+\infty} G(x; \mu, \sigma) dx &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\ &= \frac{1}{\sqrt{2\pi}\sigma} \cdot \sqrt{2}\sigma \cdot \int_{-\infty}^{+\infty} e^{-y^2} dy \\ &= \frac{1}{\sqrt{2\pi}\sigma} \cdot \sqrt{2}\sigma \cdot \sqrt{\pi} = 1 \end{aligned}$$

- ★ Variance

$$\text{Var}(x) = \langle (x - \mu)^2 \rangle = \sigma^2$$

$$\begin{aligned} \text{Proof: } \text{Var}(x) &= \int_{-\infty}^{+\infty} (x - \mu)^2 G(x; \mu, \sigma) dx = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} (x - \mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\ &= \frac{I_2}{I_0} \sigma^2 \\ &= \sigma^2 \end{aligned}$$

Properties of the 1D Gaussian Distribution, cont.



$$G(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

★ **Natural to introduce** $\chi^2(x)$

$$\chi^2 = \frac{(x-\mu)^2}{\sigma^2}$$

“squared deviation from mean in terms of standard error”

$$G(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\chi^2}{2}\right)$$

★ **Fractions of events**

68.3 % : $ x - \mu < 1\sigma$	$(\chi^2 < 1)$
95.5 % : $ x - \mu < 2\sigma$	$(\chi^2 < 4)$
99.7 % : $ x - \mu < 3\sigma$	$(\chi^2 < 9)$
6×10^{-7} : $ x - \mu > 5\sigma$	$(\chi^2 > 25)$

Averaging Gaussian Measurements

★ Suppose we have two **independent** measurements of a quantity, e.g. the W boson mass:

$$x_1 \pm \sigma_1 \quad \text{and} \quad x_2 \pm \sigma_2$$

there are two questions we can ask:

- Are the measurements compatible? [Hypothesis test – we’ll return to this]
- **What is our best estimate of the parameter x ?** (i.e. how to average)

★ In principle can take any linear combination as an unbiased estimator of x

$$x_{12} = \omega_1 x_1 + \omega_2 x_2 \quad \text{provided} \quad \omega_1 + \omega_2 = 1$$

$$\text{since} \quad \langle x_{12} \rangle = \omega_1 \langle x_1 \rangle + \omega_2 \langle x_2 \rangle = \omega_1 \mu + \omega_2 \mu = \mu$$

★ Clearly want to give the highest weight to the more precise measurements...
e.g. two undergraduate measurements of $g[\text{m s}^{-2}]$

$$10.1 \pm 0.3 \quad 5 \pm 5$$

★ **Method I:** choose the weights to minimise the uncertainty on

$$\sigma_x^2 = \sum_i \omega_i^2 \sigma_i^2$$

$$\text{subject to constraint} \quad f(\omega_1, \omega_2, \dots) = 1 - \sum_i \omega_i = 0$$

$$\begin{aligned}\frac{\partial(\sigma_x^2 + \lambda f)}{\partial \omega_i} &= 0 \\ \Rightarrow 2\omega_i \sigma_i^2 - \lambda &= 0 \\ \omega_i &\propto \frac{1}{\sigma_i^2}\end{aligned}$$

★ Therefore, since the weights sum to unity:

$$\omega_i = \frac{1/\sigma_i^2}{\sum_j 1/\sigma_j^2}$$

★ Hence for two measurements

$$\bar{x} = \frac{\frac{x_1}{\sigma_1^2} + \frac{x_2}{\sigma_2^2}}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}}$$

with

$$\sigma_{\bar{x}}^2 = \frac{1}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}}$$

← Problem: derive this.
(just error propagation as described later)

Averaging Gaussian Measurements II

★ Can obtain the same expression using a natural probability based approach

- We can interpret the first measurement in terms of a probability distribution for the true value of x , i.e. a Gaussian centred on x_1

$$P(x) = P(x; x_1) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x - x_1)^2}{2\sigma_1^2}\right\}$$

- Bayes' theorem then tells us how to modify this in the light of a new measurement

$$P(x; \text{data}) \propto P(\text{data}; x)P(x)$$

$$P(x; \text{data}) \propto \exp\left\{-\frac{(x - x_2)^2}{2\sigma_2^2}\right\} \exp\left\{-\frac{(x - x_1)^2}{2\sigma_1^2}\right\}$$

- So our new expression for the knowledge of x is:

$$P(x) \propto \exp\left\{-\frac{1}{2}\left\{\frac{(x - x_1)^2}{\sigma_1^2} + \frac{(x - x_2)^2}{\sigma_2^2}\right\}\right\}$$

- Completing the square gives plus a little algebra gives

$$P(x) \propto \exp\left\{-\frac{(x - \bar{x})^2}{2\sigma^2}\right\} \quad \text{with} \quad \bar{x} = \frac{\frac{x_1}{\sigma_1^2} + \frac{x_2}{\sigma_2^2}}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}} \quad \text{and} \quad \sigma^2 = \frac{1}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}}$$

- Product of n Gaussians is a Gaussian

Error Propagation I

- ★ Suppose measure a quantity x with a Gaussian uncertainty σ_x ; what is the uncertainty on a derived quantity

$$y = f(x)$$

- Expand $f(x)$ about \bar{x}

$$f(x) = f(\bar{x}) + (x - \bar{x}) \left(\frac{df}{dx} \right)_{\bar{x}} + \dots$$

- Define estimate of y : $\bar{y} = f(\bar{x})$

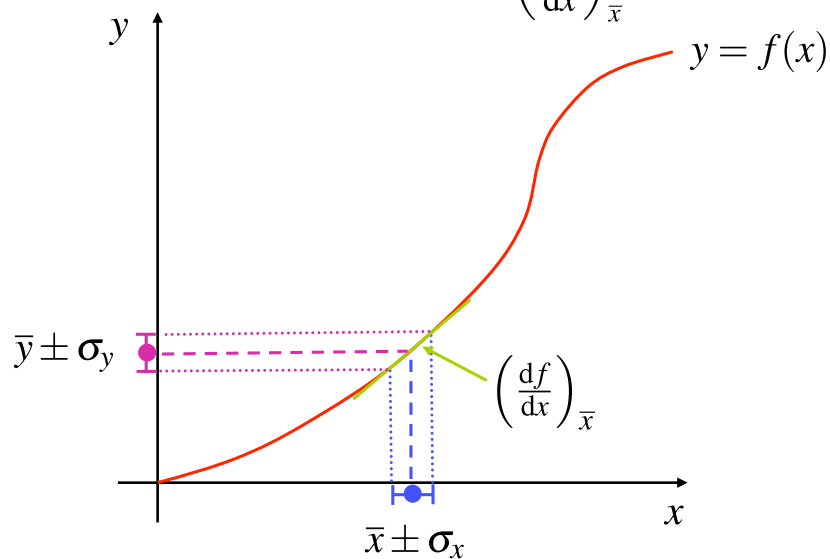
so
$$y - \bar{y} = f(x) - f(\bar{x}) \approx (x - \bar{x}) \left(\frac{df}{dx} \right)_{\bar{x}}$$

$$\langle (y - \bar{y})^2 \rangle = \langle (x - \bar{x})^2 \rangle \left(\frac{df}{dx} \right)_{\bar{x}}^2$$

$$\sigma_y^2 = \left(\frac{df}{dx} \right)_{\bar{x}}^2 \sigma_x^2$$

$$\sigma_y = \left(\frac{df}{dx} \right)_{\bar{x}} \sigma_x$$

- ★ It is easy to understand the origin of $\sigma_y = \left(\frac{df}{dx} \right)_{\bar{x}} \sigma_x$

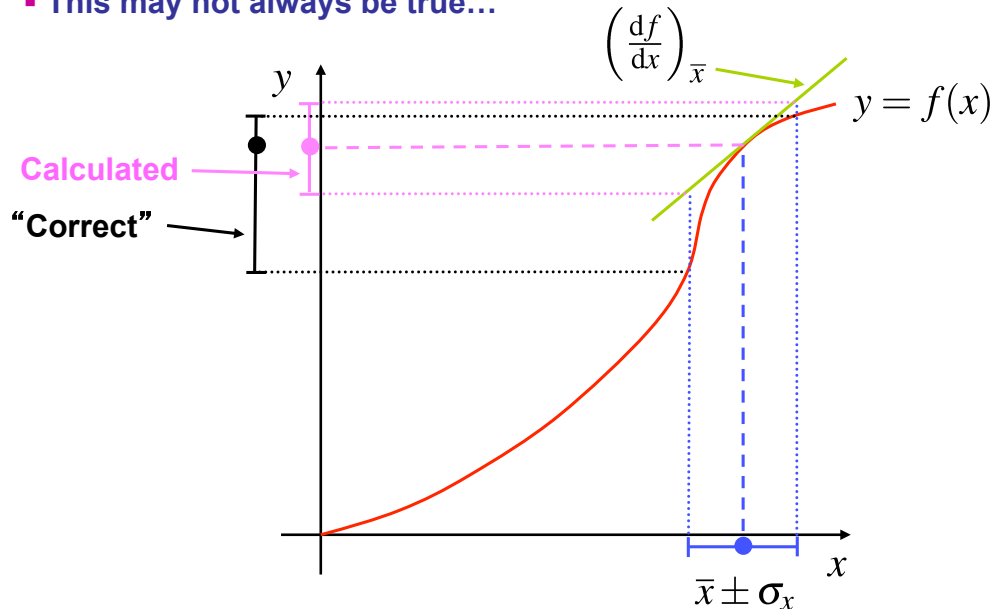


- ★ How does a “small” change in x , i.e. σ_x , propagate to a small change in y , σ_y

$$\frac{\sigma_y}{\sigma_x} \approx \left(\frac{dy}{dx} \right)_{\bar{x}}$$

★ A word of **warning**...

- In deriving the error propagation equation $\sigma_y = \left(\frac{df}{dx} \right)_{\bar{x}} \sigma_x$
- Neglected second order terms in the Taylor expansion
- This is equivalent to saying that the derivative is constant in region of interest
- This may not always be true...



Example

★ Measurement of transverse momentum of a track from a fit

- radius of curvature of track helix, R , given by

$$R = 0.3B(T)p_T(\text{GeV})$$

- track fit returns a Gaussian uncertainty in radius of curvature, and hence, the PDF is Gaussian in $1/p_T$

$$\sigma_{1/p_T}$$

- what is the error in p_T

let $x = 1/p_T$

$$p_T = 1/x$$

$$\frac{dp_T}{dx} = -\frac{1}{x^2} = -p_T^2$$

$$\sigma_{p_T}^2 = \left(\frac{dp_T}{dx} \right)^2 \sigma_x^2$$

$$\sigma_{p_T}^2 = (p_T^2)^2 \sigma_x^2$$

$$\sigma_{p_T} = p_T^2 \sigma_{1/p_T}$$

Error on Error

★ Recall question 2:

Given 5 measurements of a quantity x : 10.2, 5.5, 6.7, 3.4, 3.5

What is the **best estimate** of x and what is the **estimated uncertainty**?

$$\bar{x} = 5.86; \quad s_{n-1} = 2.80; \quad \sigma_{\bar{x}} = \frac{s_{n-1}}{\sqrt{5}} = 1.25$$

So our best estimate of x is: $x = 5.9 \pm 1.3$

★ But how good is our estimate of the error – i.e. what is the “error on the error” ?

▪ It can be shown (but not easy):

$$\text{Var}(s^2) = \frac{1}{n} \left(\langle (x - \mu)^4 \rangle - \frac{n-3}{n-1} \langle (x - \mu)^2 \rangle^2 \right)$$

▪ For a Gaussian distribution $\langle (x - \mu)^4 \rangle = 3\sigma^4$

$$\text{so} \quad \text{Var}(s^2) = \frac{\sigma^4}{n} \left(3 - \frac{n-3}{n-1} \right) = \frac{2\sigma^4}{n-1}$$

▪ Hence (by error propagation – show this) the error on the error estimate is

$$\sigma_s = \frac{\sigma}{\sqrt{2(n-1)}}$$

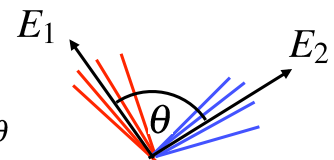
▪ To obtain a 10% estimate of σ ; need rms of 51 measurements !

Combining Gaussian Errors

★ There are many cases where we want to combine measurements to extract a single quantity, e.g. di-jet invariant mass

$$m^2 = E_1 E_2 (1 - \cos \theta)$$

• What is the uncertainty on the mass given $\sigma_{E_1}, \sigma_{E_2}, \sigma_\theta$



★ Start by considering a **simple example**

$$a = x + y$$

• Mean of a is

$$\bar{a} = \bar{x} + \bar{y}$$

• Variance of a is given by:

$$\begin{aligned} \langle (a - \bar{a})^2 \rangle &= \langle (x + y - (\bar{x} + \bar{y}))^2 \rangle \\ \sigma_a^2 &= \langle ([x - \bar{x}] + [y - \bar{y}])^2 \rangle \\ &= \langle (x - \bar{x})^2 \rangle + \langle (y - \bar{y})^2 \rangle + 2\langle (x - \bar{x})(y - \bar{y}) \rangle \\ &= \sigma_x^2 + \sigma_y^2 + 2\langle (x - \bar{x})(y - \bar{y}) \rangle \end{aligned}$$

★ Two important points:

▪ Errors add **in quadrature** (i.e. sum the squares)

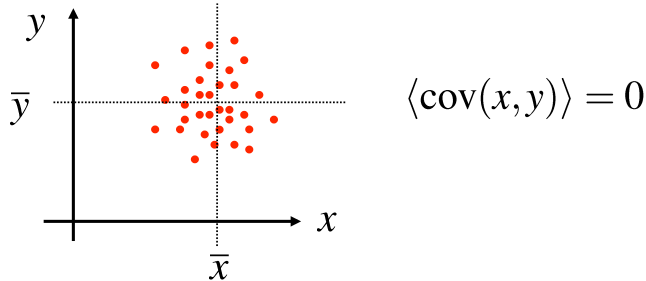
▪ The appearance of a new term, **the covariance** of x and y

$$\text{cov}(x, y) = \langle (x - \bar{x})(y - \bar{y}) \rangle$$

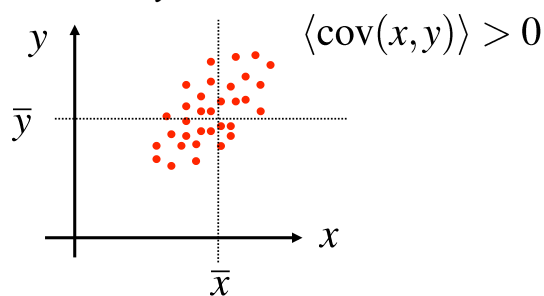
Correlated errors: covariance

★ Consider $\text{cov}(x, y) = \langle (x - \bar{x})(y - \bar{y}) \rangle$

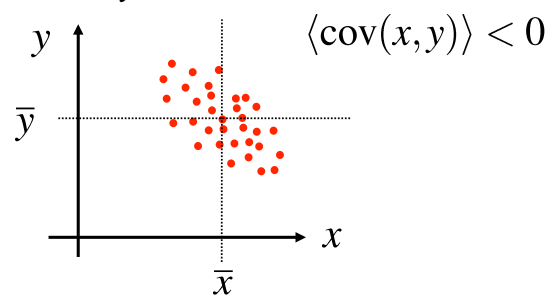
- Suppose in a single experiment measure a value of x and y
- Imagine repeating the measurement multiple times $\Rightarrow \{x_i, y_i\}$
- If the measurements of x and y are uncorrelated, i.e. **INDEPENDENT**



• If x and y are **correlated**



• If x and y are **anti-correlated**



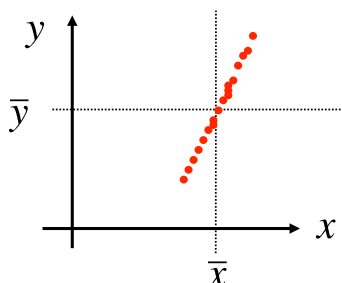
★ Often convenient to express covariance in terms of the **correlation coefficient**

$$\rho = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y}$$

$$\text{cov}(x, y) = \langle (x - \bar{x})(y - \bar{y}) \rangle$$

$$\sigma_x = \langle (x - \bar{x})^2 \rangle^{\frac{1}{2}}$$

• Consider an experiment which returns two values x and y ; where $y - \bar{y} = 2(x - \bar{x})$



$$\begin{aligned} \text{cov}(x, y) &= \langle (x - \bar{x})(2x - 2\bar{x}) \rangle \\ &= 2 \langle (x - \bar{x})^2 \rangle \\ &= 2\sigma_x^2 = \sigma_x \sigma_y \\ \Rightarrow \rho &= +1 \end{aligned}$$

★ Hence (unsurprisingly) the correlation coefficient expresses the degree of correlation with

$$|\rho| \leq 1$$

★ Going back to $a = x + y$

$$\Rightarrow \sigma_a^2 = \sigma_x^2 + \sigma_y^2 + 2\rho\sigma_x\sigma_y$$

Error Propagation II: the general case

★ We can now consider the more general case

$$a = f(x, y)$$

$$a = f(x, y) = f(\bar{x}, \bar{y}) + \frac{\partial f}{\partial x}(x - \bar{x}) + \frac{\partial f}{\partial y}(y - \bar{y}) + \dots$$

$$(a - \bar{a})^2 = (f(x, y) - f(\bar{x}, \bar{y}))^2 \approx \left(\frac{\partial f}{\partial x}\right)^2 (x - \bar{x})^2 + \left(\frac{\partial f}{\partial y}\right)^2 (y - \bar{y})^2 + 2 \frac{\partial f}{\partial x} \frac{\partial f}{\partial y} (x - \bar{x})(y - \bar{y})$$

$$\langle (a - \bar{a})^2 \rangle = \left(\frac{\partial f}{\partial x}\right)^2 \langle (x - \bar{x})^2 \rangle + \left(\frac{\partial f}{\partial y}\right)^2 \langle (y - \bar{y})^2 \rangle + 2 \frac{\partial f}{\partial x} \frac{\partial f}{\partial y} \langle (x - \bar{x})(y - \bar{y}) \rangle$$

$$\sigma_a^2 = \left(\frac{\partial f}{\partial x}\right)^2 \sigma_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 \sigma_y^2 + 2 \frac{\partial f}{\partial x} \frac{\partial f}{\partial y} \text{cov}(x, y)$$

$$\sigma_a^2 = \left(\frac{\partial f}{\partial x}\right)^2 \sigma_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 \sigma_y^2 + 2\rho \frac{\partial f}{\partial x} \frac{\partial f}{\partial y} \sigma_x \sigma_y$$

★ In order to estimate the error on a derived quantity need to know correlations

Example continued

★ Back to the original problem $m = \{E_1 E_2 (1 - \cos \theta)\}^{\frac{1}{2}}$

$$\sigma_m^2 = \left(\frac{\partial m}{\partial E_1}\right)^2 \sigma_{E_1}^2 + \left(\frac{\partial m}{\partial E_2}\right)^2 \sigma_{E_2}^2 + \left(\frac{\partial m}{\partial \theta}\right)^2 \sigma_\theta^2 + 2\rho_{12} \frac{\partial m}{\partial E_1} \frac{\partial m}{\partial E_2} \sigma_{E_1} \sigma_{E_2} + 2\rho_{1\theta} \frac{\partial m}{\partial E_1} \frac{\partial m}{\partial \theta} \sigma_{E_1} \sigma_\theta + 2\rho_{2\theta} \frac{\partial m}{\partial E_2} \frac{\partial m}{\partial \theta} \sigma_{E_2} \sigma_\theta$$

★ First assume independent errors on E_1, E_2, θ and for simplicity neglect σ_θ term

$$\frac{\partial m}{\partial E_1} = \frac{1}{2} \left\{ \frac{E_2}{E_1} (1 - \cos \theta) \right\}^{\frac{1}{2}} \quad \frac{\partial m}{\partial E_2} = \frac{1}{2} \left\{ \frac{E_2}{E_1} (1 - \cos \theta) \right\}^{\frac{1}{2}}$$

giving: $\sigma_m^2 = \frac{1}{4} \left\{ \frac{E_2}{E_1} (1 - \cos \theta) \right\} \sigma_{E_1}^2 + \frac{1}{4} \left\{ \frac{E_1}{E_2} (1 - \cos \theta) \right\} \sigma_{E_2}^2$

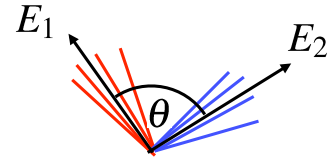
$$\frac{\sigma_m}{m} = \frac{1}{2} \left\{ \frac{\sigma_{E_1}^2}{E_1^2} + \frac{\sigma_{E_2}^2}{E_2^2} \right\}^{\frac{1}{2}}$$

★ EXERCISE: by first considering σ_{m^2} , calculate $\frac{\sigma_m}{m}$, including the σ_θ term

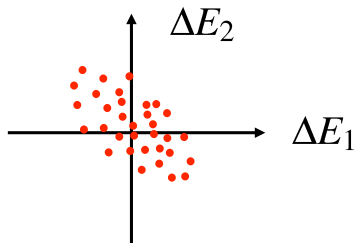
ANS: $\frac{\sigma_m}{m} = \frac{1}{2} \left\{ \frac{\sigma_{E_1}^2}{E_1^2} + \frac{\sigma_{E_2}^2}{E_2^2} + \cot^2 \left(\frac{\theta}{2} \right) \sigma_\theta^2 \right\}^{\frac{1}{2}}$

Estimating the Correlation Coefficient

- ★ Correlations can arise from physical effects, e.g.
 - Would expect E_1 and E_2 to be (slightly) anti-correlated why ?
 - Can always check (in MC) by plotting



$$\Delta E_1 = E_1 - E_1^{\text{MC}} \quad \text{against} \quad \Delta E_2 = E_2 - E_2^{\text{MC}}$$



$$\rho = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y}$$

$$\text{cov}(x, y) = \langle (x - \bar{x})(y - \bar{y}) \rangle$$

$$\sigma_x = \langle (x - \bar{x})^2 \rangle^{\frac{1}{2}}$$

NOTE: uncertainty on correlation coefficient $s_\rho \approx \frac{(1 - \rho^2)}{\sqrt{n - 2}}$

- ★ Correlations also arise when calculating derived quantities from uncorrelated measurements
 - e.g. $x = a + b$ $y = a - b$
 - this type of correlation can be handled mathematically (see later)

Properties of the 2D Gaussian Distribution

- ★ For two **independent** variables (x, y) the joint probability distribution $P(x, y)$ is simply the product of the two distributions

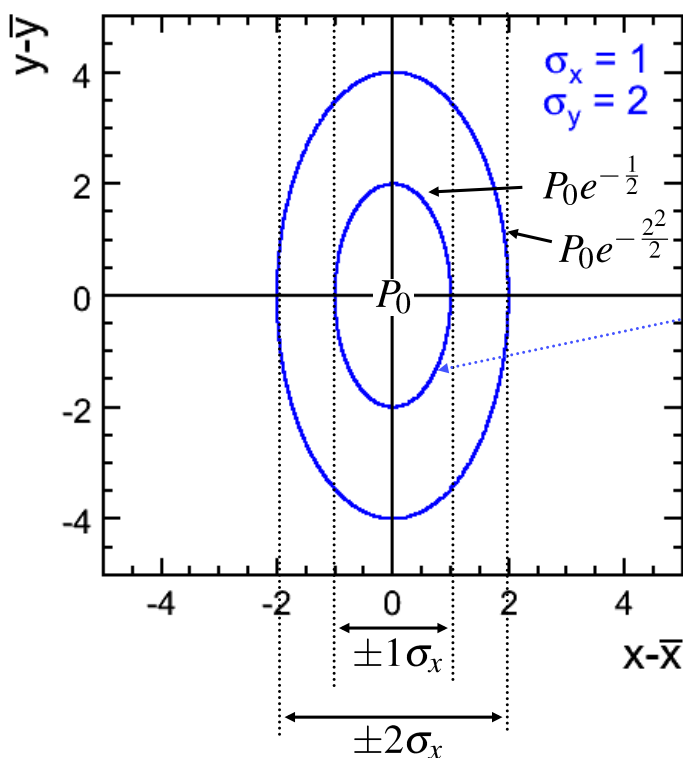
$$P(x, y) = P(x)P(y) = \frac{1}{\sqrt{2\pi}\sigma_x} \exp\left\{-\frac{(x - \bar{x})^2}{2\sigma_x^2}\right\} \frac{1}{\sqrt{2\pi}\sigma_y} \exp\left\{-\frac{(y - \bar{y})^2}{2\sigma_y^2}\right\}$$

$$P(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{-\frac{1}{2} \left[\frac{(x - \bar{x})^2}{\sigma_x^2} + \frac{(y - \bar{y})^2}{\sigma_y^2} \right]\right\}$$

NOTE: $\int_{-\infty}^{+\infty} P(x, y) dy = \frac{1}{\sqrt{2\pi}\sigma_x} \exp\left\{-\frac{(x - \bar{x})^2}{2\sigma_x^2}\right\} = P(x)$

- ★ Can write in terms of χ^2 with

$$P(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{-\frac{\chi^2}{2}\right\} \quad \chi^2 = \chi_x^2 + \chi_y^2 = \frac{(x - \bar{x})^2}{\sigma_x^2} + \frac{(y - \bar{y})^2}{\sigma_y^2}$$



- 68 % of events within $\pm 1 \sigma_x$
- 68 % of events within $\pm 1 \sigma_y$
- Now consider contours of

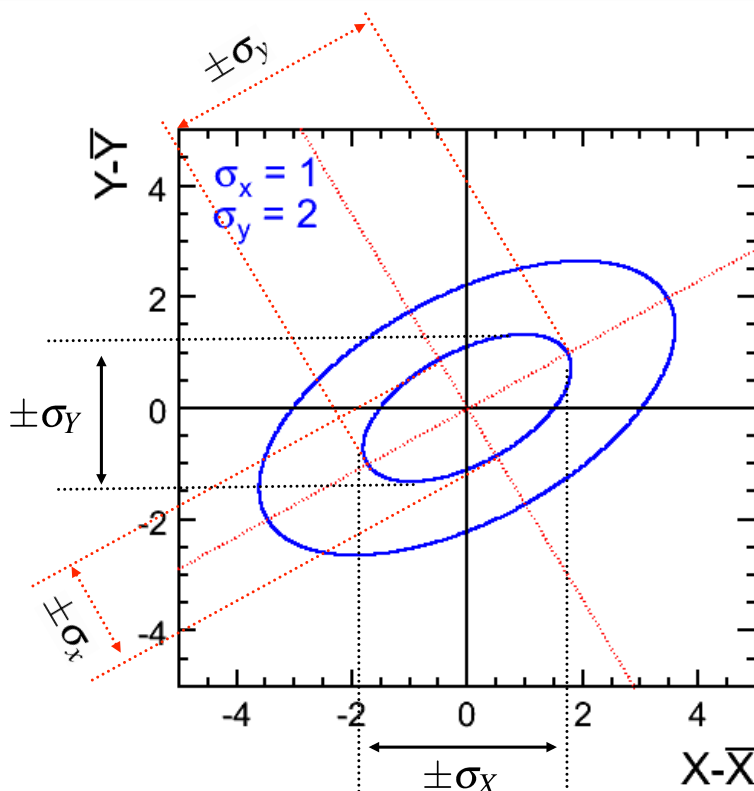
$$\chi^2 = \frac{(x - \bar{x})^2}{\sigma_x^2} + \frac{(y - \bar{y})^2}{\sigma_y^2}$$

- $\chi^2 = 1$ corresponds to contour where PDF falls to $e^{-\frac{1}{2}}$ of peak
- Only 39% of events within $\chi^2 < 1$
- Only 86% of events within $\chi^2 < 4$

Now to introduce correlations...
rotate the ellipse

$$\begin{pmatrix} X - \bar{X} \\ Y - \bar{Y} \end{pmatrix} = \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} x - \bar{x} \\ y - \bar{y} \end{pmatrix}$$

$$s = \sin \theta; \quad c = \cos \theta$$



- Same PDF, but now w.r.t. different axes
- Simple to derive the general error ellipse with correlations...

Let $X = +cx + sy$

$Y = -sx + cy$

To find the equivalent correlation coefficient, evaluate

$$\langle XY \rangle = \langle scy^2 - scx^2 + (c^2 - s^2)xy \rangle = sc(\sigma_y^2 - \sigma_x^2)$$

hence

$$\rho_{XY} \sigma_X \sigma_Y = sc(\sigma_y^2 - \sigma_x^2)$$

To eliminate the rotation angle, write

$$\sigma_X^2 = \langle X^2 \rangle = \langle c^2x^2 + s^2y^2 + 2csxy \rangle = c^2\sigma_x^2 + s^2\sigma_y^2$$

$$\sigma_Y^2 = \langle Y^2 \rangle = \langle c^2y^2 + s^2x^2 - 2csxy \rangle = c^2\sigma_y^2 + s^2\sigma_x^2$$

giving

$$\sigma_X^2 \sigma_Y^2 = s^2 c^2 (\sigma_x^4 + \sigma_y^4) + (c^4 + s^4) \sigma_x^2 \sigma_y^2$$

Compare to:

$$\rho^2 \sigma_X^2 \sigma_Y^2 = s^2 c^2 (\sigma_y^4 + \sigma_x^4 - 2\sigma_x^2 \sigma_y^2)$$

gives

$$\sigma_X^2 \sigma_Y^2 = \rho^2 \sigma_X \sigma_Y + (c^4 + 2s^2 c^2 + s^4) \sigma_x^2 \sigma_y^2$$

hence

$$(1 - \rho^2) \sigma_X^2 \sigma_Y^2 = \sigma_x^2 \sigma_y^2$$

Properties of the 2D Gaussian Distribution

★ Start from uncorrelated 2D Gaussian:

$$P(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp \left\{ -\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2} \right\}$$

★ Make the coordinate transformation

$$x = cX - sY; \quad y = sX + cY \quad P(x,y)dx dy = P(X,Y)dX dY$$

$$\begin{aligned} P(X,Y) &= \frac{1}{2\pi\sigma_x\sigma_y} \exp \left\{ -\frac{(cX - sY)^2}{2\sigma_x^2} - \frac{(cY + sX)^2}{2\sigma_y^2} \right\} \\ &= \frac{1}{2\pi\sigma_x\sigma_y} \exp \left\{ -\frac{X^2}{2} \left[\frac{c^2}{\sigma_x^2} + \frac{s^2}{\sigma_y^2} \right] - \frac{Y^2}{2} \left[\frac{c^2}{\sigma_y^2} + \frac{s^2}{\sigma_x^2} \right] + \frac{2XY}{2} sc \left[\frac{1}{\sigma_x^2} - \frac{1}{\sigma_y^2} \right] \right\} \\ &= \frac{1}{2\pi\sigma_x\sigma_y} \exp \left\{ -\frac{X^2}{2} \left[\frac{c^2\sigma_y^2 + s^2\sigma_x^2}{\sigma_x^2\sigma_y^2} \right] - \frac{Y^2}{2} \left[\frac{c^2\sigma_x^2 + s^2\sigma_y^2}{\sigma_x^2\sigma_y^2} \right] + \frac{2XY}{2} sc \left[\frac{\sigma_y^2 - \sigma_x^2}{\sigma_x^2\sigma_y^2} \right] \right\} \end{aligned}$$

★ From previous page identify

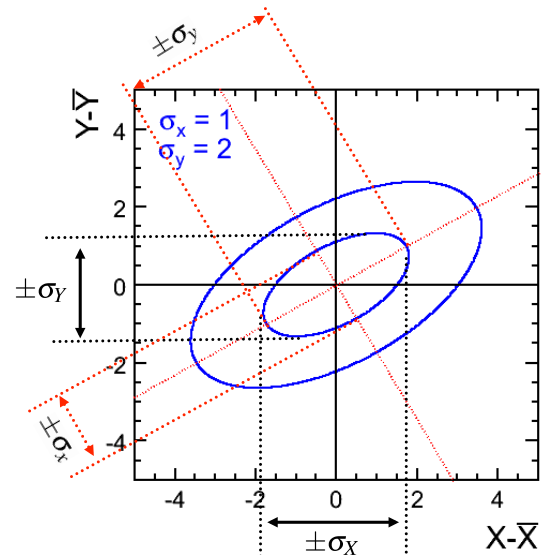
$$\langle X^2 \rangle = \sigma_X^2 = c^2\sigma_x^2 + s^2\sigma_y^2 \quad \langle Y^2 \rangle = \sigma_Y^2 = c^2\sigma_y^2 + s^2\sigma_x^2 \quad (1 - \rho^2) \sigma_X^2 \sigma_Y^2 = \sigma_x^2 \sigma_y^2$$

$$P(X,Y) = \frac{1}{2\pi\sqrt{(1-\rho^2)}\sigma_X\sigma_Y} \exp\left\{-\frac{X^2}{2} \left[\frac{\sigma_Y^2}{(1-\rho^2)\sigma_X^2\sigma_Y^2}\right] - \frac{Y^2}{2} \left[\frac{\sigma_X^2}{(1-\rho^2)\sigma_X^2\sigma_Y^2}\right] + \frac{2\rho XY}{2(1-\rho^2)\sigma_X\sigma_Y}\right\}$$

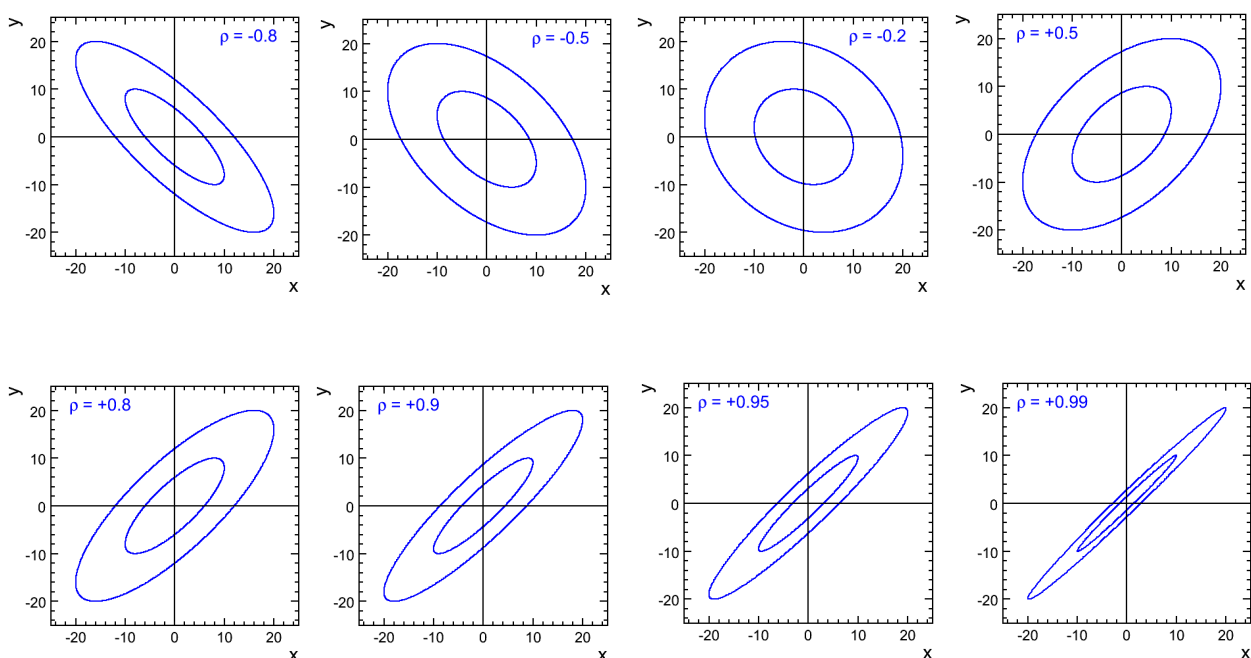
$$= \frac{1}{2\pi\sqrt{(1-\rho^2)}\sigma_X\sigma_Y} \exp\left\{-\frac{1}{2} \frac{1}{1-\rho^2} \left[\frac{X^2}{\sigma_X^2} + \frac{Y^2}{\sigma_Y^2} - \frac{2\rho XY}{\sigma_X\sigma_Y}\right]\right\}$$

★ Note we have now expressed the same ellipse in terms of the new coordinates, where the errors are now correlated.

★ If dealing with correlated errors can always find a linear combination of variables which are uncorrelated



★ Example 2D error ellipses with different correlation coefficients



The Error Ellipse and Error Matrix

- ★ Now we have the general equation for two correlated **Gaussian distributed** quantities

$$P(x,y) = \frac{1}{2\pi\sqrt{(1-\rho^2)}\sigma_x\sigma_y} \exp\left\{-\frac{1}{2}\frac{1}{1-\rho^2}\left[\frac{(x-\bar{x})^2}{\sigma_x^2} + \frac{(y-\bar{y})^2}{\sigma_y^2} - 2\frac{\rho(x-\bar{x})(y-\bar{y})}{\sigma_x\sigma_y}\right]\right\}$$

- ★ Defines the **error ellipse**
- ★ Ultimately want to generalise this to an **N variable hyper-ellipsoid**
- ★ Sounds hard... but is actually rather simple in matrix form
- ★ Define the **ERROR MATRIX**

$$\mathbf{M} = \begin{pmatrix} \langle x^2 \rangle & \langle xy \rangle \\ \langle xy \rangle & \langle y^2 \rangle \end{pmatrix}$$

i.e.

$$\mathbf{M} = \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix}$$

- ★ and define the **DISCREPANCY VECTOR**

$$\mathbf{x} = \begin{pmatrix} x - \bar{x} \\ y - \bar{y} \end{pmatrix}$$

detM

using $\mathbf{M}^{-1} = \frac{1}{1-\rho^2} \begin{pmatrix} \frac{1}{\sigma_x^2} & -\frac{\rho}{\sigma_x\sigma_y} \\ -\frac{\rho}{\sigma_x\sigma_y} & \frac{1}{\sigma_y^2} \end{pmatrix}$ and $|\mathbf{M}| = (1-\rho^2)\sigma_x^2\sigma_y^2$

we can write

$$P(x,y) = \frac{1}{2\pi|\mathbf{M}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}\mathbf{x}^T\mathbf{M}^{-1}\mathbf{x}\right\}$$

- ★ The beauty of this formalism is that it can be extended to any number of correlated Gaussian distributed variables

$$P(x_1, x_2, \dots, x_n) = \frac{1}{(2\pi)^{\frac{n}{2}}|\mathbf{M}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}\mathbf{x}^T\mathbf{M}^{-1}\mathbf{x}\right\}$$

with

$$\mathbf{M} = \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \dots & \rho_{1n}\sigma_1\sigma_n \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \dots & \rho_{2n}\sigma_2\sigma_n \\ \dots & \dots & \dots & \dots \\ \rho_{1n}\sigma_1\sigma_n & \rho_{2n}\sigma_2\sigma_n & \dots & \sigma_n^2 \end{pmatrix}$$

- ★ Can write this in terms of the χ^2 for n-variables (including correlations)

$$P(x_1, x_2, \dots, x_n) = \frac{1}{(2\pi)^{\frac{n}{2}}|\mathbf{M}|^{\frac{1}{2}}} \exp\left\{-\frac{\chi^2}{2}\right\} = P_0 e^{-\frac{\chi^2}{2}}$$

with

$$\chi^2 = \mathbf{x}^T\mathbf{M}^{-1}\mathbf{x}$$

General transformation of Errors

- ★ Suppose we have a set of variables, \mathbf{x}_i , and the error matrix, \mathbf{M} , and now wish to transform to a set of variables, \mathbf{y}_i , defined by
- ★ Taylor expansion about mean:

$$\begin{aligned}
 y_i &= \bar{y}_i + \sum_k \frac{\partial y_i}{\partial x_k} (x_k - \bar{x}_k) + \mathcal{O}(\Delta x^2) \\
 y_i - \bar{y}_i &\approx \sum_k \frac{\partial y_i}{\partial x_k} (x_k - \bar{x}_k) \\
 \langle (y_i - \bar{y}_i)(y_j - \bar{y}_j) \rangle &= \left\langle \sum_k \frac{\partial y_i}{\partial x_k} (x_k - \bar{x}_k) \sum_\ell \frac{\partial y_j}{\partial x_\ell} (x_\ell - \bar{x}_\ell) \right\rangle \\
 &= \sum_{k\ell} \frac{\partial y_i}{\partial x_k} \frac{\partial y_j}{\partial x_\ell} \langle (x_k - \bar{x}_k)(x_\ell - \bar{x}_\ell) \rangle \\
 \mathbf{M}_{\{\mathbf{y}\}}^{ij} &= \sum_{k\ell} \frac{\partial y_i}{\partial x_k} \frac{\partial y_j}{\partial x_\ell} \mathbf{M}_{\{\mathbf{x}\}}^{k\ell} \\
 \mathbf{M}_{\{\mathbf{y}\}} &= \mathbf{T}^T \mathbf{M}_{\{\mathbf{x}\}} \mathbf{T}
 \end{aligned}$$

- ★ **T** is the error transformation matrix

$$\mathbf{T} = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_2} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial y_1}{\partial x_n} & \frac{\partial y_2}{\partial x_n} & \cdots & \frac{\partial y_m}{\partial x_n} \end{pmatrix}$$

For Gaussian errors we can now do anything !

- ★ Can deal with:
 - ♦ correlated errors
 - ♦ arbitrary dimensions
 - ♦ parameter transformations

Examples...

A simple example

- ★ **Measure two uncorrelated variables** $a \pm \sigma_a, b \pm \sigma_b$

Error matrix $\mathbf{M} = \begin{pmatrix} \sigma_a^2 & 0 \\ 0 & \sigma_b^2 \end{pmatrix} \quad \mathbf{M}^{-1} = \begin{pmatrix} 1/\sigma_a^2 & 0 \\ 0 & 1/\sigma_b^2 \end{pmatrix}$

- ★ **Calculate two derived quantities**

$$x = a + b \quad y = a - b$$

- ★ **Transformation matrix**

$$\mathbf{T} = \begin{pmatrix} \frac{\partial x}{\partial a} & \frac{\partial x}{\partial b} \\ \frac{\partial y}{\partial a} & \frac{\partial y}{\partial b} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

- ★ **Giving**

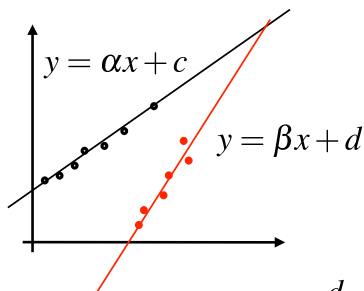
$$\begin{pmatrix} \sigma_x^2 & \rho \sigma_x \sigma_y \\ \rho \sigma_x \sigma_y & \sigma_y^2 \end{pmatrix} = \mathbf{T}^T \mathbf{M} \mathbf{T} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \sigma_a^2 & 0 \\ 0 & \sigma_b^2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

$$= \begin{pmatrix} \sigma_a^2 + \sigma_b^2 & \sigma_a^2 - \sigma_b^2 \\ \sigma_a^2 - \sigma_b^2 & \sigma_a^2 + \sigma_b^2 \end{pmatrix}$$

$$\sigma_x^2 = \sigma_y^2 = \sigma_a^2 + \sigma_b^2; \quad \rho = \frac{\sigma_a^2 - \sigma_b^2}{\sigma_a^2 + \sigma_b^2}$$

A more involved example

- ★ **Given the results of two straight line fits, calculate the uncertainty on the intersection**



- **With the error matrix (note the results of the two fits are uncorrelated)**

$$\mathbf{M} = \begin{pmatrix} \sigma_\alpha^2 & \rho_1 \sigma_\alpha \sigma_c & 0 & 0 \\ \rho_1 \sigma_\alpha \sigma_c & \sigma_c^2 & 0 & 0 \\ 0 & 0 & \sigma_\beta^2 & \rho_2 \sigma_\beta \sigma_d \\ 0 & 0 & \rho_2 \sigma_\beta \sigma_d & \sigma_d^2 \end{pmatrix}$$

- **Lines Intersect at:** $x = \frac{d - c}{\alpha - \beta} \quad y = \frac{\alpha d - \beta c}{\alpha - \beta}$

- **To calculate error on intersection need error transformation matrix, i.e. need the partial derivatives, e.g.** $\frac{\partial x}{\partial \alpha} = \frac{c - d}{(\alpha - \beta)^2}$

- **giving**

$$\mathbf{T} = \begin{pmatrix} \frac{\partial x}{\partial \alpha} & \frac{\partial x}{\partial c} \\ \frac{\partial y}{\partial \alpha} & \frac{\partial y}{\partial c} \\ \frac{\partial x}{\partial \beta} & \frac{\partial x}{\partial d} \\ \frac{\partial y}{\partial \beta} & \frac{\partial y}{\partial d} \end{pmatrix} = \frac{1}{\alpha - \beta} \begin{pmatrix} -\kappa & -\beta \kappa \\ -1 & -\beta \\ \kappa & \alpha \kappa \\ +1 & \alpha \end{pmatrix} \quad \text{with } \kappa = \frac{d - c}{\alpha - \beta}$$

- then its just algebra

$$\begin{pmatrix} \sigma_x^2 & \rho \sigma_x \sigma_y \\ \rho \sigma_x \sigma_y & \sigma_y^2 \end{pmatrix} = \mathbf{T}^T \mathbf{M} \mathbf{T}$$

$$= \begin{pmatrix} -\kappa & -1 & \kappa & +1 \\ -\beta \kappa & -\beta & \alpha \kappa & \alpha \end{pmatrix} \begin{pmatrix} \sigma_\alpha^2 & \rho_1 \sigma_\alpha \sigma_c & 0 & 0 \\ \rho_1 \sigma_\alpha \sigma_c & \sigma_c^2 & 0 & 0 \\ 0 & 0 & \sigma_\beta^2 & \rho_2 \sigma_\beta \sigma_d \\ 0 & 0 & \rho_2 \sigma_\beta \sigma_d & \sigma_d^2 \end{pmatrix} \begin{pmatrix} -\kappa & -\beta \kappa \\ -1 & -\beta \\ \kappa & \alpha \kappa \\ +1 & \alpha \end{pmatrix}$$

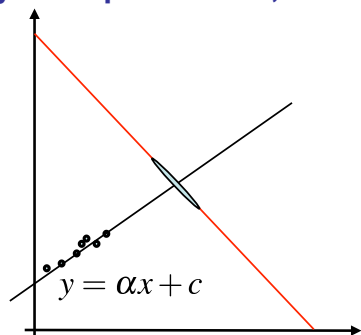
- giving

$$\begin{aligned} \sigma_x^2 &= \frac{1}{(\alpha - \beta)^2} \left[\kappa^2 (\sigma_\alpha^2 + \sigma_\beta^2) + 2\kappa (\rho_1 \sigma_\alpha \sigma_c + \rho_2 \sigma_\beta \sigma_d) + \sigma_c^2 + \sigma_d^2 \right] \\ \rho \sigma_x \sigma_y &= \frac{1}{(\alpha - \beta)^2} \left[\kappa^2 (\beta \sigma_\alpha^2 + \alpha \sigma_\beta^2) + 2\kappa (\rho_1 \beta \sigma_\alpha \sigma_c + \rho_2 \alpha \sigma_\beta \sigma_d) + \beta \sigma_c^2 + \alpha \sigma_d^2 \right] \\ \sigma_y^2 &= \frac{1}{(\alpha - \beta)^2} \left[\kappa^2 (\beta^2 \sigma_\alpha^2 + \alpha^2 \sigma_\beta^2) + 2\kappa (\rho_1 \beta^2 \sigma_\alpha \sigma_c + \rho_2 \alpha^2 \sigma_\beta \sigma_d) + \beta^2 \sigma_c^2 + \alpha^2 \sigma_d^2 \right] \end{aligned}$$

$$\kappa = \frac{d - c}{\alpha - \beta}$$

- ★ OK, it is not pretty, but we now have an analytic expression (i.e. once you have done the calculation, computationally very fast)

- Apply to a special case, intersection with a fixed line $y = 1 - x$



$$\beta = -1; d = +1; \sigma_\beta = 0; \sigma_d = 0$$

$$\begin{aligned} \sigma_x^2 &= \frac{1}{(\alpha - \beta)^2} \left[\kappa^2 \sigma_\alpha^2 + 2\kappa \rho_1 \sigma_\alpha \sigma_c + \sigma_c^2 \right] \\ \rho \sigma_x \sigma_y &= \frac{1}{(\alpha - \beta)^2} \left[-\kappa^2 \sigma_\alpha^2 - 2\kappa \rho_1 \sigma_\alpha \sigma_c - \sigma_c^2 \right] \\ \sigma_y^2 &= \frac{1}{(\alpha - \beta)^2} \left[\kappa^2 \sigma_\alpha^2 + 2\kappa \rho_1 \sigma_\alpha \sigma_c + \sigma_c^2 \right] \end{aligned}$$

Hence $\sigma_x^2 = \sigma_y^2$; $\rho = -1$ which makes perfect sense

- ★ The treatment of Gaussian errors via the **error matrix** is an extremely powerful technique – it is also easy to apply (once you understand the basic ideas)

Summary

★ Should now understand:

- ♦ Properties of the Gaussian distribution
- ♦ How to combine errors
- ♦ Propagation simple of 1D errors
- ♦ How to include correlations
- ♦ How to treat multi-dimensional errors
- ♦ How to use the error matrix

★ Next up, chi-squared, likelihood fits, ...

Appendix: Error on Error - Justification

▪ Assume mean of distribution is zero (can always make this transformation without affecting the variance)

$$\begin{aligned} \text{Var}(s^2) &= \langle (s^2 - \sigma^2)^2 \rangle \\ &= \langle \left(\frac{1}{n} \sum x^2 - \sigma^2 \right)^2 \rangle \\ &= \frac{1}{n^2} \langle \sum_i x_i^2 \sum_j x_j^2 \rangle - 2\sigma^2 \langle \left(\frac{1}{n} \sum x^2 \right) \rangle + \sigma^4 \\ &= \frac{1}{n^2} \left(n \langle x^4 \rangle + n(n-1) \langle x_i^2 x_j^2 \rangle_{i \neq j} \right) - \sigma^4 \\ &\approx \frac{1}{n} \langle x^4 \rangle + \frac{n-1}{n} \sigma^4 - \sigma^4 \quad \left\{ \begin{array}{l} \text{For large } n \\ \langle x_i^2 x_j^2 \rangle_{i \neq j} \approx \sigma^4 \end{array} \right. \\ &= \frac{1}{n} (\langle x^4 \rangle - \langle x^2 \rangle^2) \end{aligned}$$