Practical Statistics for Physicists

Louis Lyons

Imperial College and Oxford

CMS expt at LHC

l.lyons@physics.ox.ac.uk

CERN Latin American School

March 2015

Extra Lecture:
Learning to love the Error Matrix

- Introduction via 2-D Gaussian
- Understanding covariance
- Using the error matrix

  Combining correlated measurements
- Estimating the error matrix

Reminder of 1-D Gaussian or Normal

$$y = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\{-(x-\mu)^2/(2\sigma^2)\}$$

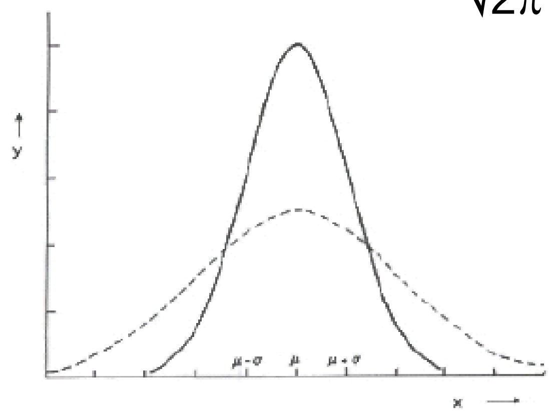


Fig. 1.5. The solid curve is the Gaussian distribution of eqn (1.14). The distribution peaks at the mean μ, and its width is characterised by the parameter σ. The dashed curve is another Gaussian distribution with the same values of μ, but with σ twice as large as the solid curve. Because the normalisation condition (1.15) ensures that the area under the curves is the same, the height of the dashed curve is only half that of the solid curve at their maxima. The scale on the x-axis refers to the solid curve.

Significance of σ

i) RMS of Gaussian $= \sigma$
(Hence factor of 2 in defn of Gaussian)

ii) At $x = \mu \pm \sigma$, $y = y_{max}/\sqrt{e}$
(i.e. $\sigma \sim$ half-width at "half" height)

iii) Fractional area within $\mu \pm \sigma$ is 68%

iv) Height at max $= 1/\sqrt{2\pi}\,\sigma$

Correlations

Basic issue:

For 1 parameter, quote value and error

For 2 (or more) parameters,
(e.g. gradient and intercept of straight line fit)
quote values + errors **+ correlations**

Just as the concept of variance for single variable is more general than Gaussian distribution, so correlation in more variables does not require multi-dim Gaussian

But more simple to introduce concept this way

Gaussian in 2-variables

$$P(x) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_x} e^{-\frac{1}{2}\frac{x^2}{\sigma_x^2}}$$

$$P(y) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_y} e^{-\frac{1}{2} y^2/\sigma_y^2}$$

x + y uncorrelated $\Rightarrow$

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

Down on $P(0,0)$ by $e^{-\frac{1}{2}}$ when

$$\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} = 1$$

Rewrite as

$$(x \quad y)\begin{pmatrix} \frac{1}{\sigma_x^2} & 0 \\ 0 & \frac{1}{\sigma_y^2} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 1$$

Invert

$\Rightarrow$ ERROR MATRIX

$$\begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix}$$

Element E_{ij} - $\langle(x_i - \overline{x}_i)(x_j - \overline{x}_j)\rangle$

Diagonal E_{ij} = variances

Off-diagonal E_{ij} = covariances

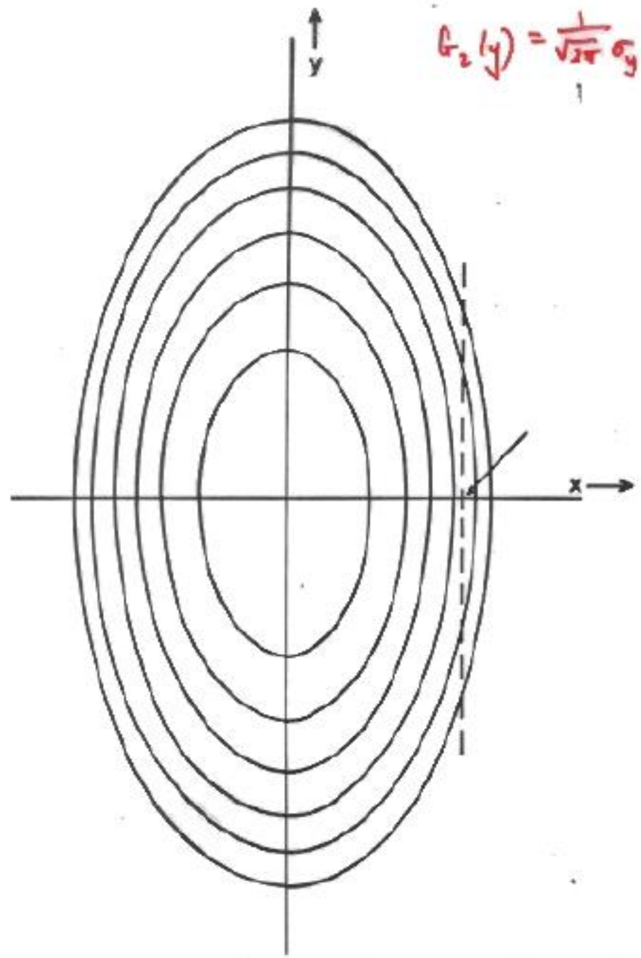
Towards the Error Matrix

x + y indep Gaussians

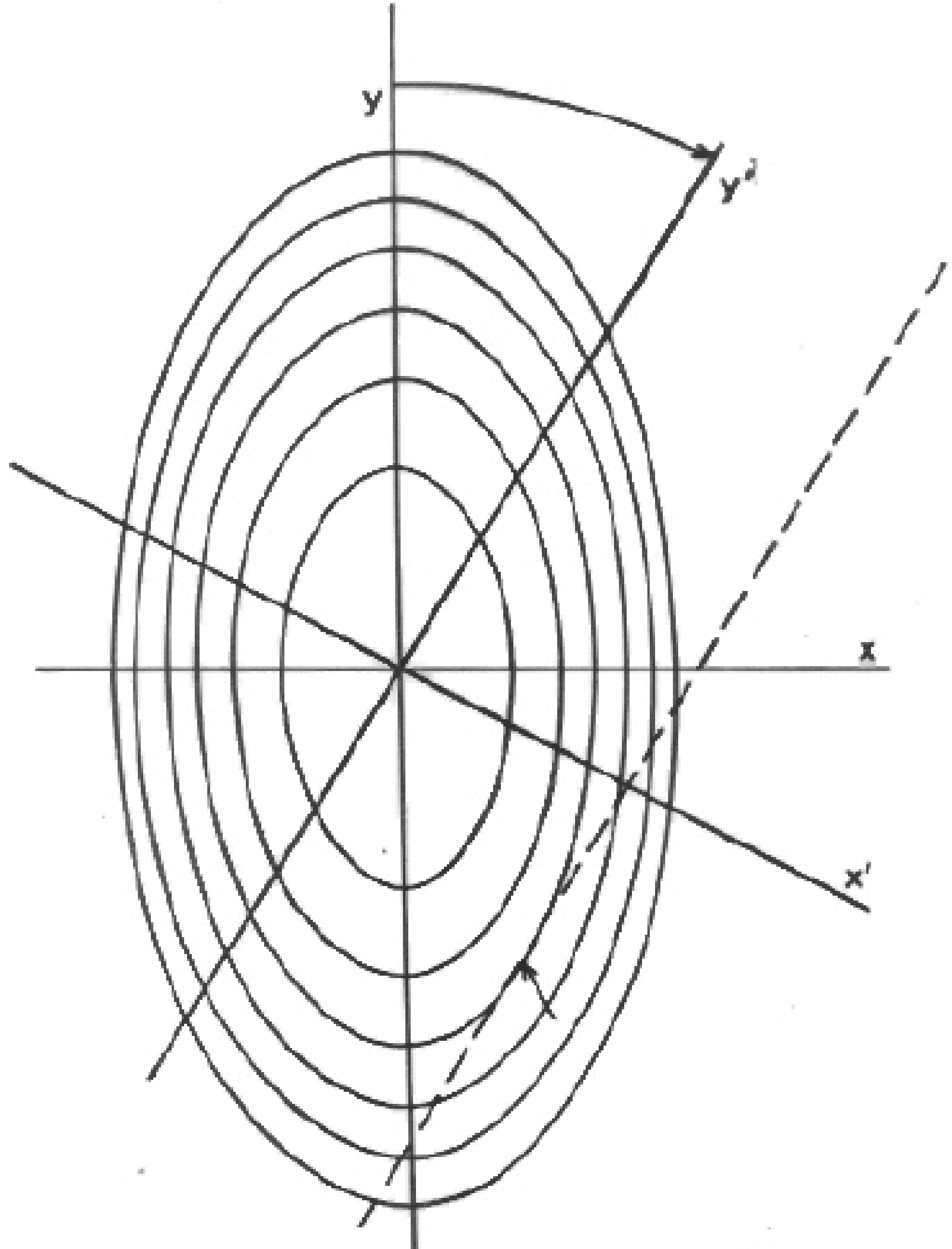
$$P(x,y) = G_1(x)\, G_2(y)$$

$$G_1(x) = \frac{1}{\sqrt{2\pi}\,\sigma_x} \exp\left[-\frac{1}{2}\frac{x^2}{\sigma_x^2}\right]$$

$$G_2(y) = \frac{1}{\sqrt{2\pi}\,\sigma_y} \exp\left[-\frac{1}{2}\frac{y^2}{\sigma_y^2}\right]$$



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


y
y'
x
x'

Specific example

$$\sigma_x = \frac{\sqrt{2}}{4} = .354 \qquad \sigma_y = \frac{\sqrt{2}}{2} = .707$$

Then factor of $e^{-\frac{1}{2}}$ when

$$8x^2 + 2y^2 = 1$$

Now introduce CORRELATIONS by 30° rot^n

$$\frac{1}{2}\left[13x'^2 + 6\sqrt{3}\,x'y' + 7y'^2\right] = 1$$

$$\begin{pmatrix} \frac{13}{2} & \frac{3\sqrt{3}}{2} \\ \frac{3\sqrt{3}}{2} & \frac{7}{2} \end{pmatrix} = \text{Inverse Error Matrix}$$

$$\frac{1}{32} \times \begin{pmatrix} 7 & -3\sqrt{3} \\ -3\sqrt{3} & 13 \end{pmatrix} = \text{Error Matrix}$$

$8x^2 + 2y^2 = 1$

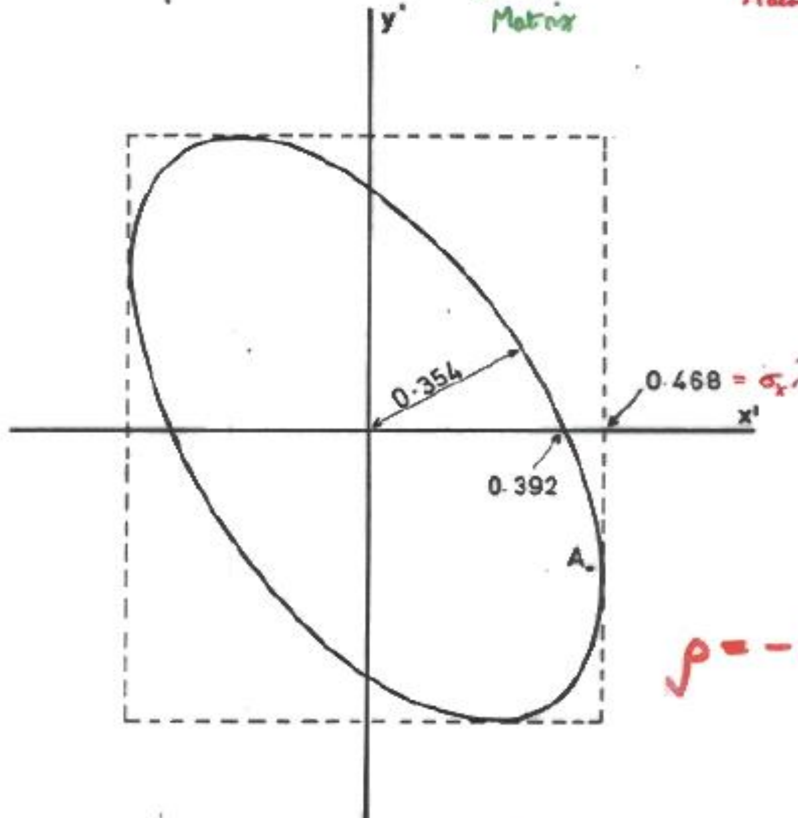
$\frac{1}{2}\left[13x'^2 + 6\sqrt{3}\,x'y' + 7y'^2\right] = 1$

$\begin{pmatrix} \frac{13}{2} & \frac{3\sqrt{3}}{2} \\ \frac{3\sqrt{3}}{2} & \frac{7}{2} \end{pmatrix}$

Inverse Error Matrix

$\frac{1}{32} \times \begin{pmatrix} 7 & -3\sqrt{3} \\ -3\sqrt{3} & 13 \end{pmatrix}$

Error Matrix



$\rho = -0.54$

$(0.468)^2 = \frac{7}{32} = \sigma_{x'}^2$

$(0.392)^2 = 1/6.5$

$\frac{1}{8} = (0.354)^2 =$ Eigenvalue of error matrix $= \sigma_x^2$

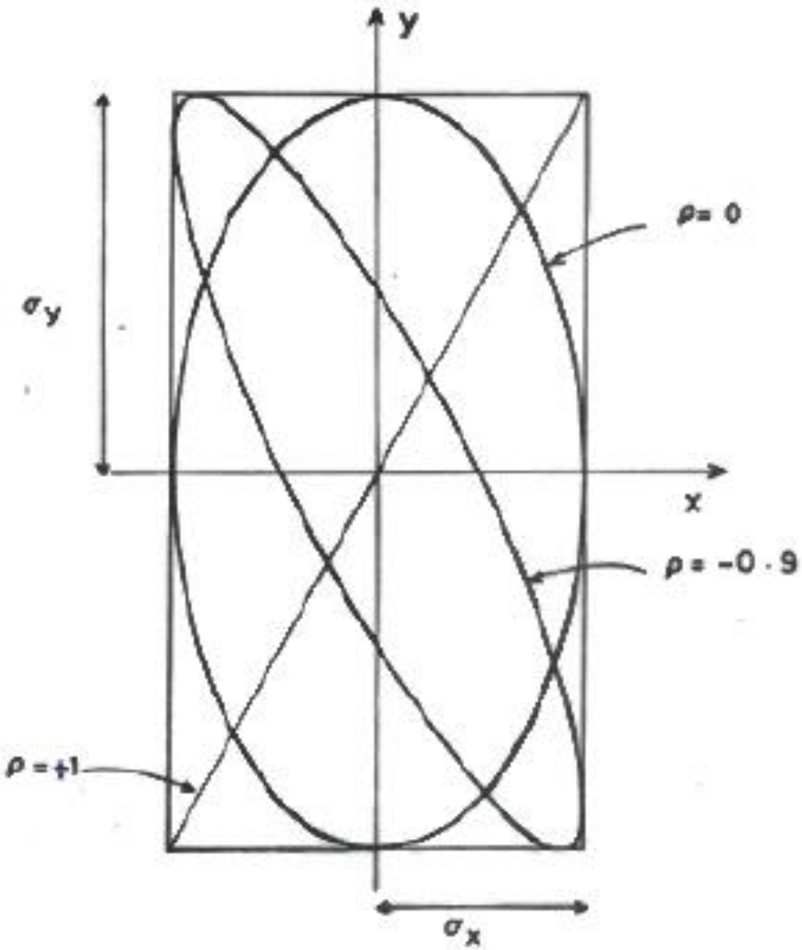
$$\left.\begin{matrix}\sigma_x \\ \sigma_y\end{matrix}\right\}\ \text{constant}$$

ρ varying

Covariance

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

Error Matrix



USING THE ERROR MATRIX

(i) Function of variables

$$y = y(x_a, x_b)$$

Given x_a, x_b error matrix, what is σ_y?

Differentiate, square, average

$$\overline{\delta y^2} = \left(\frac{\partial y}{\partial x_a}\right)^2 \overline{\delta x_a^2} + \left(\frac{\partial y}{\partial x_b}\right)^2 \overline{\delta x_b^2} + 2\frac{\partial y}{\partial x_a}\frac{\partial y}{\partial x_b}\overline{\delta x_a \delta x_b}$$

Zero if x_a, x_b uncorrelated

OR

$$\overline{\delta y^2} = \underbrace{\begin{pmatrix}\frac{\partial y}{\partial x_a} & \frac{\partial y}{\partial x_b}\end{pmatrix}}_{\tilde{D}} \underbrace{\begin{pmatrix}\overline{\delta x_a^2} & \overline{\delta x_a \delta x_b} \\ \overline{\delta x_a \delta x_b} & \overline{\delta x_b^2}\end{pmatrix}}_{\text{Error matrix}} \underbrace{\begin{pmatrix}\frac{\partial y}{\partial x_a} \\ \frac{\partial y}{\partial x_b}\end{pmatrix}}_{\text{Derivative vector } D}$$

$$\sigma_y^2 = \tilde{D} E D$$

(ii) Change of variables $\quad x_a = x_a(p_i, p_j)$
$\quad x_b = x_b(p_i, p_j)$

e.g. Cartesian $\Rightarrow$ polars

or Points in x, y $\Rightarrow$ m, c of straight line fit

Given (p_i, p_j) error matrix $\Rightarrow$ (x_i, x_j) error matrix

Differentiate, $\delta x_a \delta x_b$, average

$$\delta x_a = \frac{\partial x_a}{\partial p_i}\delta p_i + \frac{\partial x_a}{\partial p_j}\delta p_j \qquad (+ \text{ sim for } x_b)$$

Then

$$\overline{\delta x_a^2} = \left(\frac{\partial x_a}{\partial p_i}\right)^2 \overline{\delta p_i^2} + \left(\frac{\partial x_a}{\partial p_j}\right)^2 \overline{\delta p_j^2} + 2\frac{\partial x_a}{\partial p_i}\frac{\partial x_a}{\partial p_j}\overline{\delta p_i \delta p_j}$$

$$\overline{\delta x_a \delta x_b} = \frac{\partial x_a}{\partial p_i}\frac{\partial x_b}{\partial p_i}\overline{\delta p_i^2} + \frac{\partial x_a}{\partial p_j}\frac{\partial x_b}{\partial p_j}\overline{\delta p_j^2} + \left(\frac{\partial x_a}{\partial p_i}\frac{\partial x_b}{\partial p_j} + \frac{\partial x_a}{\partial p_j}\frac{\partial x_b}{\partial p_i}\right) \times \overline{\delta p_i \delta p_j}$$

$+ \overline{\delta x_b^2}$ like $\overline{\delta x_a^2}$

N.B. Change of variables does not have to be $N \to N$

e.g. straight line fit involves $N \to 2$

Then i) & ii) are both examples of $N \to M$ $(M \le N)$
where $M = 1$ in i) $\quad M = N$ in ii)

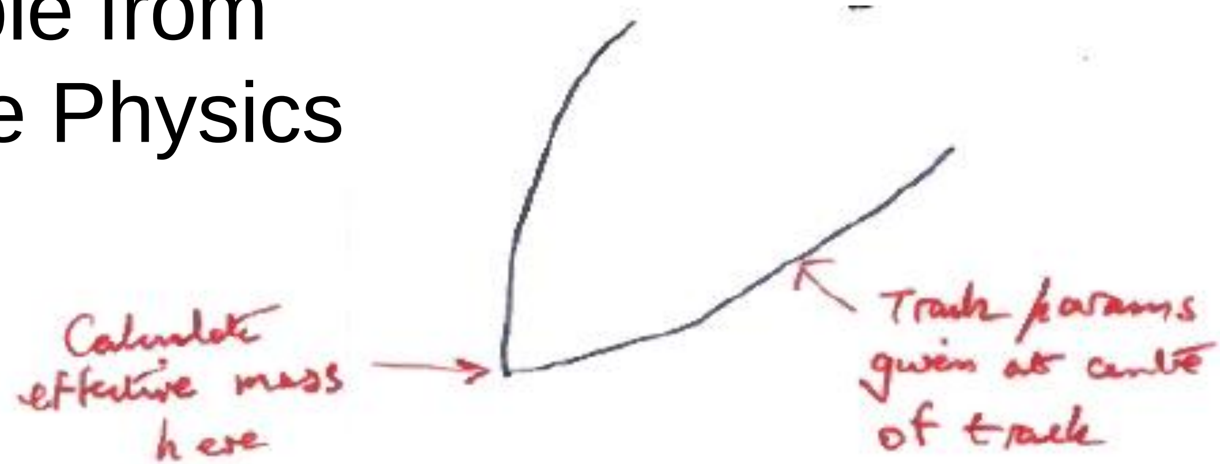
i.e.

$$\underbrace{\begin{pmatrix} \overline{\delta x_a^2} & \overline{\delta x_a \delta x_b} \\ \overline{\delta x_a \delta x_b} & \overline{\delta x_b^2} \end{pmatrix}}_{\text{New error matrix}} = \underbrace{\begin{pmatrix} \frac{\partial x_a}{\partial p_i} & \frac{\partial x_a}{\partial p_j} \\ \frac{\partial x_b}{\partial p_i} & \frac{\partial x_b}{\partial p_j} \end{pmatrix}}_{\tilde{T}} \underbrace{\begin{pmatrix} \overline{\delta p_i^2} & \overline{\delta p_i \delta p_j} \\ \overline{\delta p_i \delta p_j} & \overline{\delta p_j^2} \end{pmatrix}}_{\text{Old error matrix}} \underbrace{\begin{pmatrix} \frac{\partial x_a}{\partial p_i} & \frac{\partial x_b}{\partial p_i} \\ \frac{\partial x_a}{\partial p_j} & \frac{\partial x_b}{\partial p_j} \end{pmatrix}}_{\text{Transform matrix } T}$$

$$E_x = \tilde{T} E_p T$$

BEWARE!

Example from Particle Physics



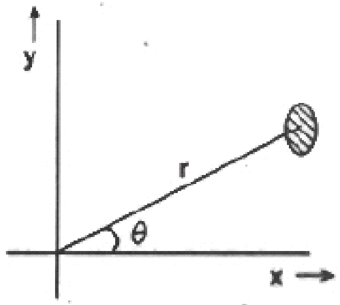
Tracks' error matrix (centre of tracks)

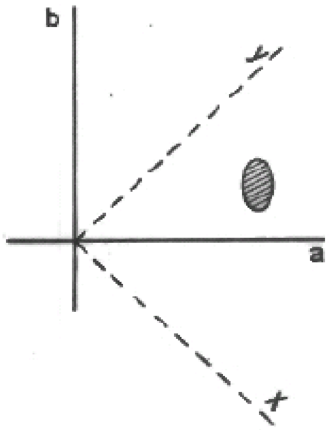
$$\sigma_M^2 = \tilde{D}\,\tilde{T}\,E\,T\,D$$

Deriv vector for mass in terms of track params at vertex

Transformation matrix from centre of tracks to vertex

Examples of correlated variables





USING THE ERROR MATRIX

COMBINING RESULTS

If $a_i \pm \sigma_i$ are independent:

Minimise $S = \sum \left(\frac{a_i - \hat{a}}{\sigma_i}\right)^2$

$$\Rightarrow \quad \hat{a} = \frac{\sum a_i w_i}{\sum w_i} \qquad w_i = 1/\sigma_i^2$$

Now $a_i \pm \sigma_i$ are correlated with error matrix $\underline{\underline{E}}$

$$\underline{\underline{E}} = \begin{pmatrix} \sigma_1^2 & cov(1,2) & cov(1,3) & \cdots \\ cov(1,2) & \sigma_2^2 & cov(2,3) & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

$$S = \sum_{i,j} (a_i - \hat{a}) \, \underline{\underline{E}}^{-1}_{ij} \, (a_j - \hat{a})$$

INVERSE ERROR MATRIX

N.B. $\hat{a}$ CAN LIE OUTSIDE a_i

$\sigma_a \rightarrow 0$ AS $\rho \rightarrow \pm 1$

$$\underline{\underline{E}}^{-1} = \begin{pmatrix} 1/\sigma_1^2 & 0 & 0 & \cdots \\ 0 & 1/\sigma_2^2 & 0 & \\ \vdots & \vdots & \vdots & \end{pmatrix}$$

FOR UNCORRELATED

MORE COMBINING:

SEVERAL PAIRS OF CORRELATED MEAS.

(x_i, y_i) with $\underline{\underline{E}}_i = \begin{pmatrix} \sigma_x^2 & cov \\ cov & \sigma_y^2 \end{pmatrix}_i$

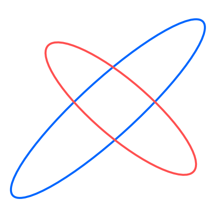
$$S = \sum_i \left\{ (x_i - \hat{x})^2 E^{-1}_{11,i} + (y_i - \hat{y})^2 E^{-1}_{22,i} + 2(x_i - \hat{x})(y_i - \hat{y}) E^{-1}_{12,i} \right\}$$

ice result: —

Inverse error matrix on result $\hat{x}, \hat{y}$

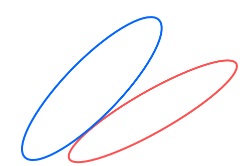
$$= \sum_i \underline{\underline{E}}_i^{-1}$$

cf $\frac{1}{\sigma^2} = \sum \frac{1}{\sigma_i^2}$ for single uncorrelated meas.



Small error

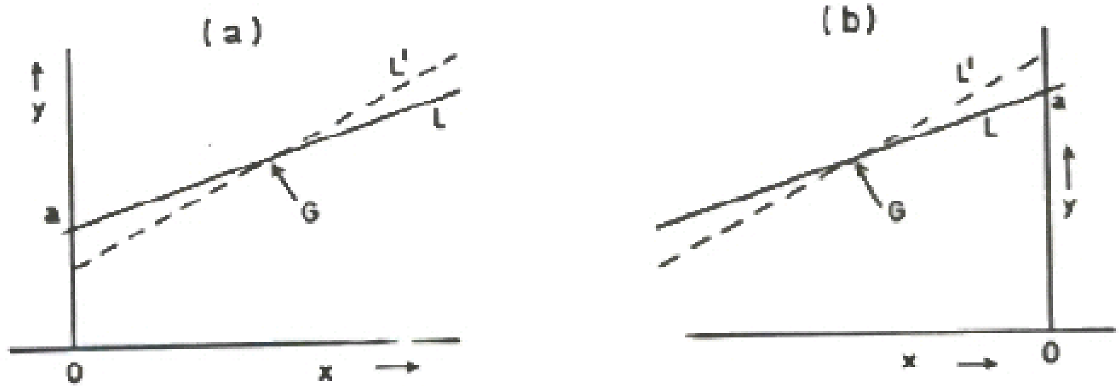
Example: Chi-sq Lecture



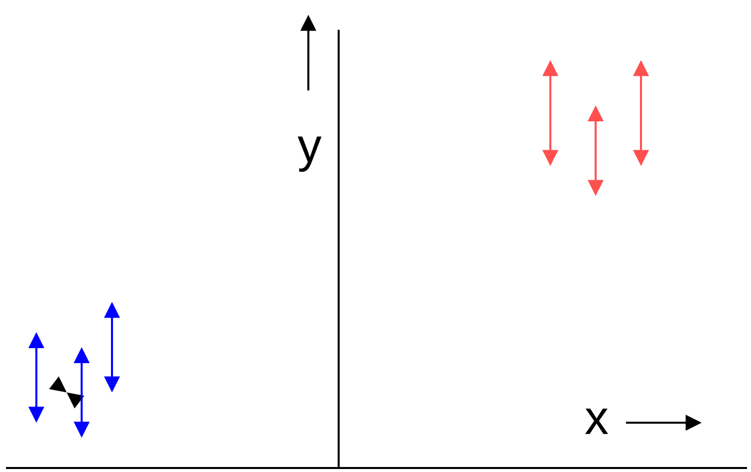
x_{best} outside $x_1 \rightarrow x_2$

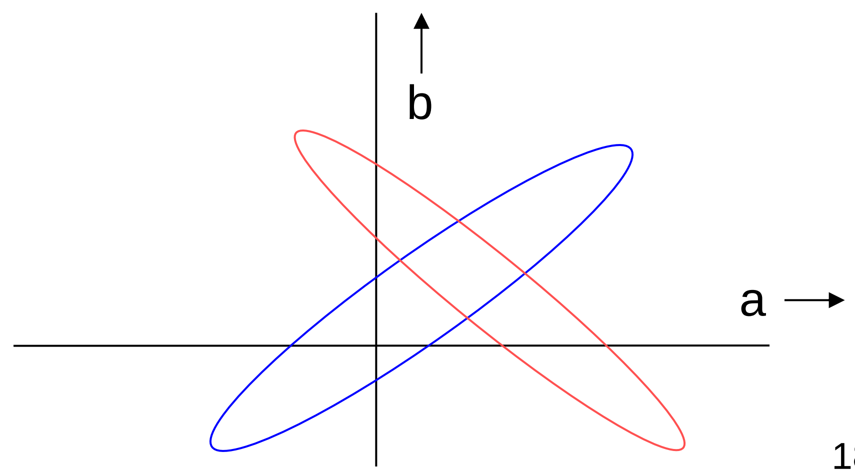
y_{best} outside $y_1 \rightarrow y_2$

COVARIANCE (a, b) $\propto -\langle x \rangle$

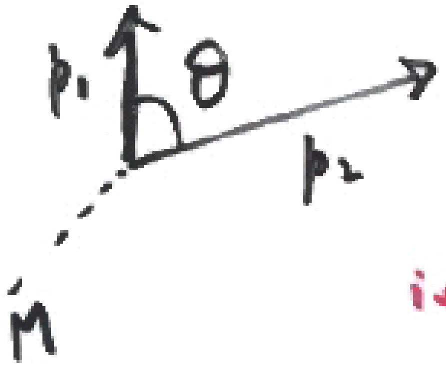


$\langle x \rangle$ pos   Fig. 2.4 $\langle x \rangle$ neg





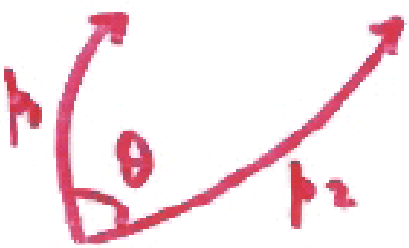
CORRELATIONS + MASS RESOLUTION



$$M^2 = (E_1+E_2)^2 - (\underline{p}_1+\underline{p}_2)^2$$

$$\sim p_1 p_2 \theta \qquad [\, p_i \gg m_i ;\ \theta \ll 1 \,]$$

i.e. $M \uparrow$ as $p_i \uparrow$ + $\theta_i \uparrow$



As $p_i \downarrow$, $\theta \uparrow$

$\therefore$ Smaller σ_M

As $p_i \downarrow$, $\theta \downarrow$

$\therefore$ Larger σ_M

ESTIMATING THE ERROR MATRIX

1) ESTIMATE ERRORS
ESTIMATE CORRELATIONS
(Usually easiest if $\rho = 0$ or ± 1)

2) FOR INDEP SOURCES OF ERRORS,
ADD ERROR MATRICES

e.g. M_W FROM $WW \to 4$ JETS
$WW \to JJ\ell\nu$

$\underline{\underline{E}} = (M_W)_1, (M_W)_2$ ERROR MATRIX

$$\underline{\underline{E}} = \underline{\underline{E}}_{stat} + \underline{\underline{E}}_{B.E.} + \underline{\underline{E}}_{E\ scale} + \underline{\underline{E}}_{FSR} + \underline{\underline{E}}_{colour\ +\ reconn}$$

$$\begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix} \qquad \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2 \\ \sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \qquad \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & 0 \end{pmatrix}$$

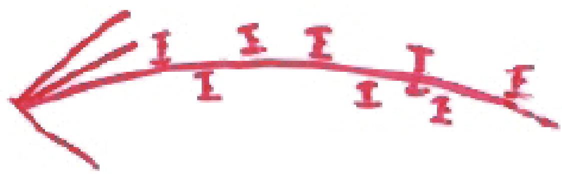
3) TRANSFORMATIONS

e.g. $(x \pm \sigma_x , y \pm \sigma_y)$ with uncorrel. errors

$\Rightarrow$ r, θ with correlations

Indep data points

$\Rightarrow$ correlated a and b



Track fit

4) REPEATED OBSERVATIONS

$(x_i, y_i) \Rightarrow \sigma_x^2 \quad \sigma_y^2$ and

cov(x, y) from $\overline{(x-\bar{x})(y-\bar{y})}$

Conclusion

Error matrix formalism makes life easy when correlations are relevant