$$\chi^2 = (\vec{m} - \vec{y})^T V^{-1} (\vec{m} - \vec{y})$$

V is symmetric and positive definite.

Change of variables:

$$\vec{m} = \vec{f}(\vec{x}) \qquad \mathcal{O}(\vec{m}) = N \quad , \quad \mathcal{O}(\vec{x}) = M$$

Linear approximation: $\vec{m} = \vec{m}_0 + A\vec{x}$

where $A = \dfrac{\partial \vec{f}}{\partial \vec{x}}$

Better notation: $A_{ij} = \dfrac{\partial f_i}{\partial x_j}$

$$i = 1, \ldots, N$$
$$j = 1, \ldots, M$$

$$\chi^2 = (\vec{m}_0 - \vec{y})^T V^{-1} (\vec{m}_0 - \vec{y}) + 2(\vec{m}_0 - \vec{y}) V^{-1} A \vec{x} + \vec{x}^T A^T V^{-1} A \vec{x}$$

$$\mathrm{Cov}(x_i, x_j) = \frac{1}{2}\left(\frac{\partial \chi^2}{\partial x_i \partial x_j}\right)^{-1} = \frac{1}{2}\left(A^T V^{-1} A\right)^{-1}$$

A should be evaluated using $\vec{x}$ which minimizes χ^2.

Example:

$$\vec{m} = \alpha + \beta \vec{z} \qquad \left(m_i = \alpha + \beta z_i \right)$$

Measurements are made at various known z-coordinates. Assume measurements are uncorrelated.

$$V = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \sigma_2^2 & \\ 0 & & \ddots & \\ & & & \sigma_N^2 \end{pmatrix}$$

$$\frac{\partial f_i}{\partial \alpha} = 1 \qquad \frac{\partial f_i}{\partial \beta} = z_i$$

$$A^T V^{-1} A = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & & z_N \end{pmatrix} \begin{pmatrix} 1/\sigma_1^2 & & & \\ & 1/\sigma_2^2 & & \\ & & \ddots & \\ & & & 1/\sigma_N^2 \end{pmatrix} \begin{pmatrix} 1 & z_1 \\ 1 & z_2 \\ \vdots & \vdots \\ 1 & z_N \end{pmatrix}$$

$$= \begin{pmatrix} 1/\sigma_1^2 & 1/\sigma_2^2 & \cdots & 1/\sigma_N^2 \\ z_1/\sigma_1^2 & z_2/\sigma_2^2 & \cdots & z_N/\sigma_N^2 \end{pmatrix} \begin{pmatrix} 1 & z_1 \\ 1 & z_2 \\ \vdots & \vdots \\ 1 & z_N \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}$$

$$a_{11} = \sum_{i=1}^{N} 1/\sigma_i^2$$

$$a_{12} = \sum_{i=1}^{N} z_i/\sigma_i^2$$

$$a_{22} = \sum_{i=1}^{N} z_i^2/\sigma_i^2$$

$$\mathrm{Cov}(\alpha, \beta) = \frac{\begin{pmatrix} a_{22} & -a_{12} \\ -a_{12} & a_{11} \end{pmatrix}}{a_{11} a_{22} - a_{12}^2} = \begin{pmatrix} C_{11} & C_{12} \\ C_{12} & C_{22} \end{pmatrix}$$

What is the expected distribution of y measured at z'?

$$y = \alpha + \beta z' \qquad \text{(Error propagation)}$$

$$\sigma_y^2 = \begin{pmatrix} \frac{\partial y}{\partial \alpha} & \frac{\partial y}{\partial \beta} \end{pmatrix} \left(\mathrm{cov}(\alpha, \beta) \right) \begin{pmatrix} \frac{\partial y}{\partial \alpha} \\ \partial y/\partial \beta \end{pmatrix} = \begin{pmatrix} C_{11} + z' C_{12} & C_{12} + z' C_{22} \end{pmatrix} \begin{pmatrix} 1 \\ z' \end{pmatrix}$$

$$\begin{matrix} 1 & z' \end{matrix}$$

$$= C_{11} + 2 z' C_{12} + z'^2 C_{22} .$$

What if the problem is nonlinear? Then we cant solve for $\hat{m}$ by inverting a matrix. Resort to numerical methods.

Newton's method.

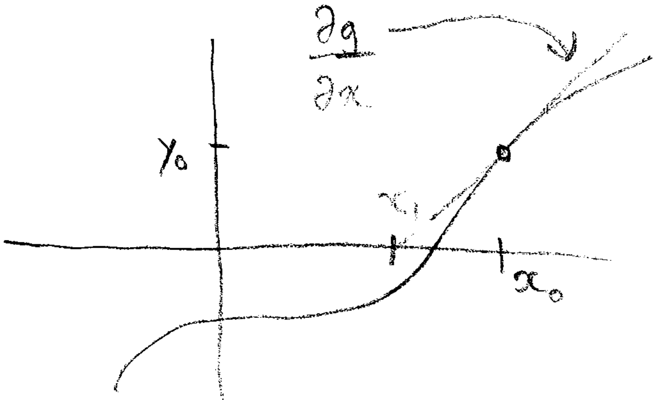
Suppose $y = f(x)$. What value of x will minimize y?

At the minimum, $f'(\hat{x}) = 0$.

$\hat{x}$ is a root of the equation $f'(x) = 0$

Let $g(x) = f'(x)$.

Solve $g(x) = 0$ using Newton's method.



$$x_1 = x_0 - \frac{y_0}{\left.\frac{\partial g}{\partial x}\right|_{x_0}}$$

$$\chi^2 = \vec{u}^T V^{-1} \vec{u} \qquad \vec{u} = \vec{f}(\vec{x}) - \vec{y} \approx \vec{f}_0 + A\Delta\vec{x} - \vec{y}_0$$

$$\frac{\partial \chi^2}{\partial \vec{x}} = 2 \frac{\partial \vec{u}}{\partial \vec{x}} V^{-1} \vec{u} = 2 A^T V^{-1} \vec{u}$$

$$A_{ij} = \frac{\partial u_i}{\partial x_j}$$

$$\frac{\partial \chi^2}{\partial \vec{x}} = 2 A^T V^{-1} A \Delta\vec{x} + 2 A^T V^{-1} (\vec{f}_0 - \vec{y}_0) = 0$$

$$\Delta\vec{x} = -(A^T V^{-1} A)^{-1} A^T V^{-1} (\vec{f}_0 - \vec{y}_0)$$

Progressive fit.

Suppose we have current estimates for parameters α, β with covariance V but add one more measurement y_{N+1} at z_{N+1}

$$\chi^2 = \begin{pmatrix} \hat{\alpha} - \alpha & \hat{\beta} - \beta \end{pmatrix} V^{-1} \begin{pmatrix} \hat{\alpha} - \alpha \\ \hat{\beta} - \hat{\beta} \end{pmatrix} + \left(\frac{\hat{\alpha} + \hat{\beta} z_{N+1}}{\sigma_{N+1}^2} \right)^2$$

$$\frac{\partial \chi^2}{\partial \hat{\alpha}} = 2 V^{-1} \begin{pmatrix} \hat{\alpha} - \alpha \\ \hat{\beta} - \beta \end{pmatrix} + \frac{2(\hat{\alpha} + \hat{\beta} z_{N+1})}{\sigma_{N+1}^2}$$

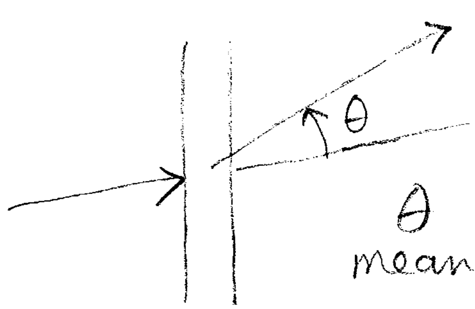
$$\frac{\partial \chi^2}{\partial \hat{\beta}} = 2 V^{-1} \begin{pmatrix} \hat{\alpha} - \beta \\ \hat{\beta} - \beta \end{pmatrix} + \frac{2 z_{N+1} (\hat{\alpha} + \hat{\beta} z_{N+1})}{\sigma_{N+1}^2}$$

$$\left(V^{-1} + \frac{1}{\sigma_{N+1}^2} \begin{pmatrix} 1 & z_{N+1} \\ z_{N+1} & z_{N+1}^2 \end{pmatrix} \right) \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = V^{-1} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \left[V^{-1} + \frac{1}{\sigma_{N+1}^2} \begin{pmatrix} 1 & z_{N+1} \\ z_{N+1} & z_{N+1}^2 \end{pmatrix} \right]^{-1} V^{-1} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

$$\hat{V} = \left[V^{-1} + \frac{1}{\sigma_{N+1}^2} \begin{pmatrix} 1 & z_{N+1} \\ z_{N+1} & z_{N+1}^2 \end{pmatrix} \right]^{-1}$$

Treatment of multiple scattering from thin sheets of material.



θ is a random variable with mean 0 and RMS θ_0.

$\rightarrow x \leftarrow$

$$\theta_0 = \frac{13.6\,\text{MeV}}{\beta c p} z \sqrt{\frac{x}{X_0}} \left[1 + 0.038 \log\left(\frac{x z^2}{X_0 \beta^2} \right) \right]$$

(PDG - Passage of particles through matter)

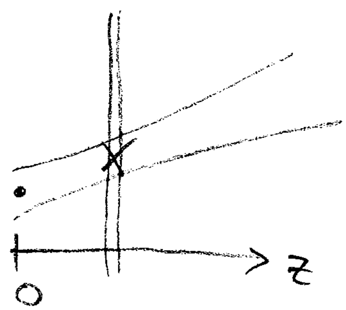
Scattering from a thin sheet does not change $\hat{\alpha}$, only $\hat{\beta}$..

$$\langle \hat{\alpha}' \rangle = \hat{\alpha}$$

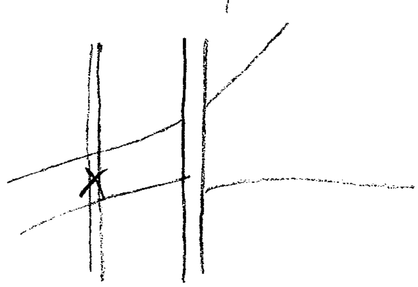
$$\langle \hat{\beta}' \rangle = \hat{\beta}$$

$$\sigma^2_{\hat{\beta}'} = \sigma^2_{\hat{\beta}} + \theta_0^2$$

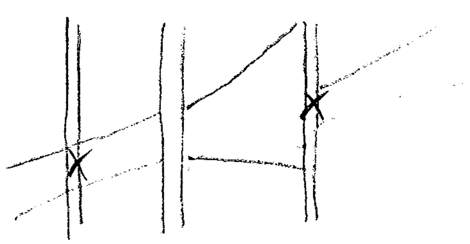
Initial measurement:



Multiple scattering



Next measurement:



Often we want the best estimates for $\hat{\alpha}$ and $\hat{\beta}$ at the origin (eg, before any multiple scattering).

Constrained fits.

Example: Minimize $f(x,y) = x$ subject to the constraint

$$g(x,y) = (x-x_0)^2 + (y-y_0)^2 - R^2 = 0$$

Method of Lagrange Multipliers:

Find stationary points of the function

$$\mathcal{L} = f(x,y) - \lambda g(x,y) = 0$$

with respect to x, y, λ.

$$\frac{\partial \mathcal{L}}{\partial x} = \frac{\partial \mathcal{L}}{\partial y} = \frac{\partial \mathcal{L}}{\partial \lambda} = 0 .$$

$$\frac{\partial \mathcal{L}}{\partial x} = 1 - 2\lambda(x - x_0) = 0$$

$$\frac{\partial \mathcal{L}}{\partial y} = -2\lambda(y-y_0) = 0$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = R^2 - (x-x_0)^2 - (y-y_0)^2$$

Solution: $y = y_0$

$$x - x_0 = \frac{1}{2\lambda}$$

$$R^2 - \frac{1}{4\lambda^2} = 0 \qquad \lambda = \pm\frac{1}{2R}$$

$$x = x_0 \pm R .$$

as expected.

In general suppose we have measurements $\hat{y}$ with covariance matrix V

$$\chi^2 = (\hat{y} - \vec{y})^T V^{-1} (\hat{y} - \vec{y})$$

is minimized when $\hat{y} = \vec{y}$. What if we impose constraints

$$\vec{g}(\vec{y}, \vec{p}) = 0$$

$\vec{p}$ are parameters that can be expressed as functions of $\vec{y}$.

Example: $\vec{y} = (p_x, p_y, p_z)$

$$\vec{g} = \begin{pmatrix} p_x^2 + p_y^2 + p_z^2 - P^2 \\ E^2 - P^2 - M^2 \end{pmatrix} = 0$$

This introduces two constraints and two parameters (P and E).

$$L = (\hat{y} - \vec{y})^T V^{-1} (\hat{y} - \vec{y}) + 2\vec{\lambda}^T \vec{g}(\vec{y}, \vec{p})$$

$\vec{\lambda}$ is a vector of Lagrange multipliers.

$\vec{g}$ can be linearized about a point $\vec{y}_0, \vec{p}_0$:

$$\vec{g}(\vec{y}, \vec{p}) = \vec{g}(\vec{y}_0, \vec{p}_0) + A(\vec{y} - \vec{y}_0) + B(\vec{p} - \vec{p}_0)$$

$$A = \left.\frac{\partial \vec{g}}{\partial \vec{y}}\right|_{\vec{y}_0, \vec{p}_0} \qquad B = \left.\frac{\partial \vec{g}}{\partial \vec{p}}\right|_{\vec{y}_0, \vec{p}_0}$$

Then,
$$\begin{aligned} L &= \vec{e}^T V^{-1} \vec{e} + 2\vec{\lambda}^T (\vec{g}_0 + A(\vec{y} - \vec{y}_0) + B(\vec{p} - \vec{p}_0)) \\ &= \vec{e}^T V^{-1} \vec{e} + 2\vec{\lambda}^T (\vec{g}_0 - A(\hat{y} - \vec{y}) + A(\hat{y} - \vec{y}_0) + B(\vec{p} - \vec{p}_0)) \\ &= \vec{e}^T V^{-1} \vec{e} + 2\vec{\lambda}^T (\vec{g}_0 + A(\hat{y} - \vec{y}_0) - A\vec{e} + B\vec{d}) \end{aligned}$$

$$\vec{e} = (\hat{y} - \vec{y}) \qquad \vec{d} = \vec{p} - \vec{p}_0$$

$$\frac{\partial L}{\partial \vec{e}} = 2\vec{e}^T V^{-1} - 2\vec{\lambda}^T A = 0$$

$$\frac{\partial L}{\partial \vec{d}} = 2\vec{\lambda}^T B = 0$$

Thus, $\vec{e} = VA^T\vec{\lambda}$

$B^T\vec{\lambda} = 0$

Solve for $\vec{\lambda}$:

$$g(\vec{y},\vec{p}) = \vec{g}_0 + A(\hat{y} - \vec{y}_0) - A\vec{e} + B\vec{d} = 0$$

$$= \vec{g}_0 + A(\hat{y} - \vec{y}_0) - AVA^T\vec{\lambda} + B\vec{d} = 0$$

$$\vec{\lambda} = (AVA^T)^{-1}\left(\vec{g}_0 + A(\hat{y} - \vec{y}_0) + B\vec{d}\right)$$

$$\equiv W\left(\vec{g}_0 + A(\hat{y} - \vec{y}_0) + B\vec{d}\right)$$

But $B^T\vec{\lambda} = 0$ so solve for $\vec{d}$:

$$B^T W\left(\vec{g}_0 + A(\hat{y} - \vec{y}_0) + B\vec{d}\right) = 0$$

$$\vec{d} = -(B^T W B)^{-1} B^T W(\vec{g}_0 + A(\hat{y} - \vec{y}_0))$$

Then,

$$\vec{e} = VA^T W\left(1 - B(B^T W B)^{-1} B^T W\right)\left(\vec{g}_0 + A(\hat{y} - \vec{y}_0)\right)$$

$$\vec{y}_1 = \hat{y} - \vec{e}$$

$$\vec{p}_1 = \vec{p}_0 + \vec{d}$$

$$\mathrm{Cov}(\vec{p},\vec{p}) = (B^T W B)^{-1}$$

$$\mathrm{Cov}(\vec{y},\vec{p}) = -VA^T W B(B^T W B)^{-1}$$

$$\mathrm{Cov}(\vec{y},\vec{y}) = V - VA^T W A V + VA^T W B(B^T W B)^{-1} B^T W A V$$