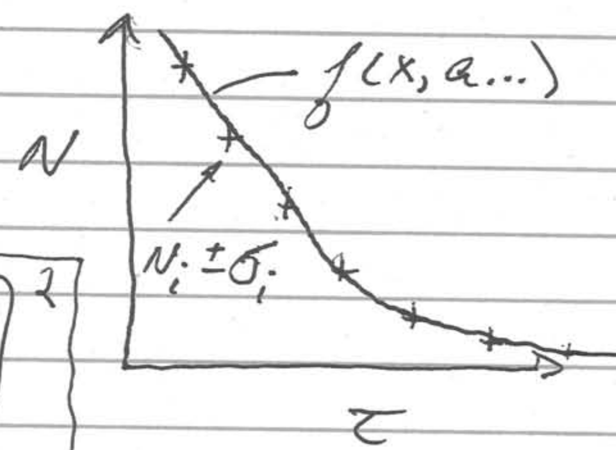
Least-Squares Method

Imagine have a set of n measurements, y_i, at points x_i

→ each measurement has error σ_i

We want to fit to some function
$f(x; a_1, a_2 \ldots a_n)$
↳ fit parameters - unknown



Calculate figure of merit

$$\boxed{S = \sum_{i=1}^{n} \left[\frac{y_i - f(x_i; a_j)}{\sigma_i} \right]^2}$$

↳ would be a χ^2 if errors Gaussian (not always the case)

Seek a_j for which S is a minimum

Analytically, determining a_j would require solving system of equations

$$\frac{\partial S}{\partial a_j} = 0$$

→ not always an analytic solution

∴ need to probe ~~analytically~~ computationally for minimum

Errors

Obtained from covariance matrix

$$(V^{-1})_{ij} = \frac{1}{2} \frac{\partial^2 S}{\partial a_i \partial a_j} \rightarrow \text{evaluate at minimum}$$

- change of a_i near minimum ⇒ variance of a_i

$$V = \begin{bmatrix} \sigma_1^2 & cov(1,2) & cov(1,3) & \cdots \\ \vdots & \sigma_2^2 & - & - \\ \vdots & \vdots & \sigma_3^2 & - \end{bmatrix}$$

Straight Line Fit

Consider points fitted to a simple linear function $y = ax + b$

To get 'a' and 'b':

$$\frac{\partial S}{\partial a} = 0 \quad ; \quad \frac{\partial S}{\partial b} = 0$$

→ assigning

$$A = \sum \frac{x_i}{\sigma_i^2} \qquad B = \sum \frac{1}{\sigma_i^2}$$

$$C = \sum \frac{y_i}{\sigma_i^2} \qquad D = \sum \frac{x_i^2}{\sigma_i^2}$$

$$E = \sum \frac{x_i y_i}{\sigma_i^2}$$

Gives

$$a = \frac{EB - CA}{DB - A^2}$$

$$b = \frac{DC - EA}{DB - A^2}$$

Covariance matrix yields

$$\sigma_a^2 = B/(BD - A^2) \; ; \; \sigma_b^2 = D/(BD - A^2)$$

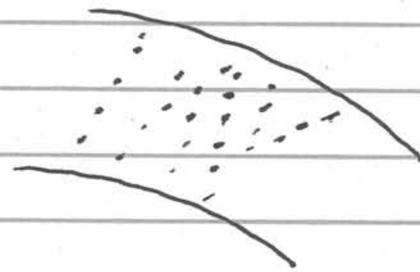
$$\mathrm{cov}(a, b) = -A/(BD - A^2)$$

Charged Particle Trajectories

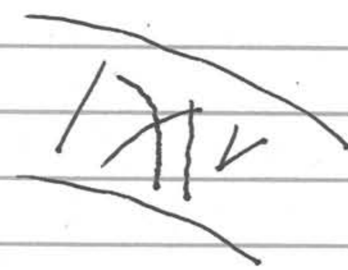
Pattern recognition

- individual hits in tracking detectors

How go from this? $\Rightarrow$



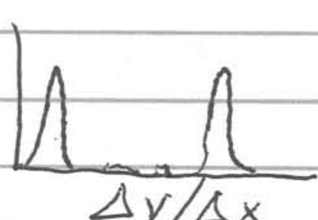
To this? $\Rightarrow$



It is impractical to
try all combinations of hits to find the best associations for tracks
- but have advantage of considering full range of configurations

Can also analyze hits by several calculable parameters

eg. $\frac{(Y_i - Y_j)}{(X_i - X_j)}$ for straight tracks



<u>Road Method</u>: Tracking

If have a track with measured positions relative to that calculated

$$\eta_i \equiv y_i - \epsilon_i = \vec{y} - \vec{\epsilon}$$

$$= x_i \cdot a_1 + 1 \cdot a_2$$

linear relation

$$\vec{\eta} = X \cdot \vec{a}$$

To obtain $\vec{a}$, we need the covariance matrix of $\vec{y}$

$$C_y = \begin{pmatrix} \sigma_1^2 & & & 0 \\ & \sigma_2^2 & & \\ & & \sigma_3^2 & \\ 0 & & & \sigma_4^2 \end{pmatrix} = G_y^{-1}$$

diagonal because each measurement is independent

Using least-squares, one minimizes

$$\chi^2 = \vec{\epsilon}^T G_y \vec{\epsilon}$$

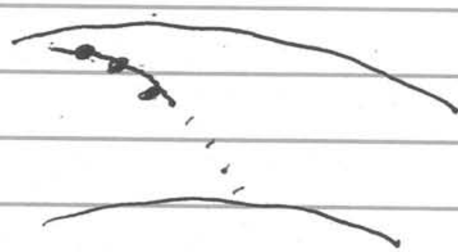
Get

$$\vec{a} = (X^T G_y X)^{-1} X^T G_y \vec{y} \quad \text{(values)}$$

$$C_a = (X^T G_y X)^{-1} \quad \text{(errors)}$$

For large trackers with many hits

- may start at outer edge of detector
- consider 3 hit triplets
    - fit to a parabola

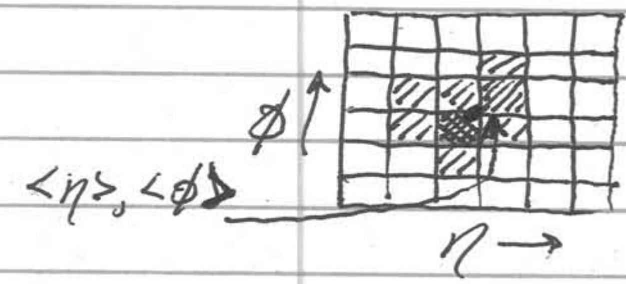


    - ~~[illegible]~~ extrapolate inward

- forming track segments this way
- a series of segments is combined to form a final track
    - drop points with large 'residuals' (distance from track of good χ^2)
- refit

Clustering

- a series of "hits" or "cells" in a 2D plane
  - often a 3rd coordinate involved



Dark = more energy deposited

- a position measurement is determined for centroid of distribution

$$\text{e.g.} \quad \langle \eta \rangle = \frac{\sum_{i=1}^{\#\text{cells}} \eta_i E_i}{\sum_{i=1}^{\#\text{cells}} E_i}$$

- accuracy of location due to
  - granularity of detector
  - characteristic size + fluctuations of "pulseheight" measurements
    → "Energy" in calorimeters

A Simple Cone Algorithm

Need to *find* the cluster of high energy hits/cells in a calorimeter

- one method is to draw a cone in detector with vertex at interaction point

$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$

- Look for cells with large energy deposition
- draw cone ΔR in radius around this cell
- calculate a new η', ϕ' centroid

An iterative process is often needed

Quark Jets:

Aftermath of parton fragmentation: jets

- a spray of particles arrayed roughly around direction of an initiating colored object: a quark or gluon

- there is not a perfect mapping between the kinematics of this 'jet' and the partons

- at detector level

  - final state particles ~~from~~ including prompt decays of ~~hadronic~~ fragmentation products
  - collision + shower in calorimeter
    - many overlapping showers

Several reconstruction schemes: two most common

ITERATIVE FIXED-CONE
JADE (energy or momentum space algorithm)

JADE Algorithm

We will iteratively consider the closeness of cells in "momentum space"
- as opposed to using η, ϕ more directly

Use

$$y_{ij} = \frac{2 E_i E_j (1 - \cos\theta_{ij})}{E_{vis}^2} \qquad M_{ij}^2$$

Cells within a threshold on y_{ij} are combined:
yields a new "pseudojet" with its own E_i and direction

The specifics of many algorithms can vary by experiment

Example 4.7 Find the best straight line through the following measured points

x	0	1	2	3	4	5
y	0.92	4.15	9.78	14.46	17.26	21.90
σ	0.5	1.0	0.75	1.25	1.0	1.5

Applying (4.75) to (4.82), we find

$a = 4.227 \qquad b = 0.878$

$\sigma(a) = 0.044 \qquad \sigma(b) = 0.203$ and

$\mathrm{cov}(a, b) = -0.0629.$