Least Squares in Matrix Form

- Need to estimate motion
- Idea is to given observation, make a model to fit that data
  - Can be multivariate

$$y_i = \beta_1 + \beta_2 x_{2i} \cdots + \beta_k x_{ki} + \varepsilon_i \qquad (i = 1, 2, \ldots n)$$

In matrix form

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \quad X = \begin{pmatrix} 1 & x_{21} & \cdots & x_{k1} \\ \vdots & & & \\ 1 & x_{2n} & \cdots & x_{kn} \end{pmatrix} \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

Unknown param ← → Unobserved disturbances

$$y = X\beta + \varepsilon$$

We prefer overdetermined structure so that it can accomodate noise

Residual form:

$$y = X\beta + e \quad \to \text{Residuals}$$

$$e = y - X\beta$$

Sum of squares of residuals as function of b:

$$S(b) = \sum e_i^2 = e'e = (y - Xb)'(y - Xb)$$

$$S(b) = yy' - y'Xb - b'X'y + b'X'Xb$$

Then take derivative to get least squares estimator

$$\frac{\delta S}{\delta b} = -2X'y + 2X'Xb$$

To minimise

$$X'Xb = X'y \rightarrow \text{Normal equations}$$

$$\Rightarrow b = (X'X)^{-1}X'y$$

↳ X should atleast have rank k

$n \geq k \rightarrow$ Considerably smaller

No. of parameters is smaller than or equal to number of observations

Can be proved by taking Hessian matrix

$$\frac{\partial^2 S}{\partial b \partial b'} = 2X'X$$

Can be extended to non-linear estimates as well.

$\boxed{\text{Read the geometric proj part}}$

For non-linear parameters:

$P = RQ + t$

9 elements

$$A_{3\times 12} = \begin{bmatrix} R_{9\times 1} \\ t_{3\times 1} \end{bmatrix} = b_{12\times 1}$$

Try to solve for b in a least squares problem in terms of

$$[r_{11}, r_{12} \dots r_{33}] \ \& \ [t_x \ t_y \ t_z]^T$$

There is a problem though as there are constraints

$$RR^T = R^TR = I$$

$$\|R_1\| = \|R_2\| = \|R_3\| = 1$$

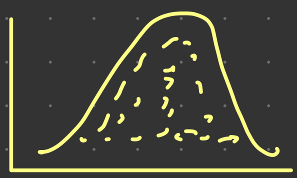
$$R_i R_j = 0$$

Makes problem non convex

$$y_i = a e^{\frac{x-\mu}{2\sigma^2}}$$

Given set of $50 = m$ observation and estimate for a, μ, σ

Linearise y about $[a_0, \mu_0, \sigma_0]^T = \beta_0$



$$y_0 = f(\beta)\big|_{\beta=\beta_0} = f(a_0, \mu_0, \beta_0)$$

using taylor series

$$y = y_0 + f'(\beta)\,\delta(\beta)_{\beta=\beta_0}$$

for i^{th} observation,

$$y_i = y_0 + J_i f(\beta_i)$$

↳ Jacobian due to i^{th} observation

$$[J_{i1} \quad J_{i2} \quad J_{i3}] = \left[\frac{\partial y_i}{\partial a}, \frac{\partial y_i}{\partial \mu} \quad \frac{\partial y_i}{\partial \sigma}\right] = \left[e^{-\frac{(x-\mu_0)^2}{2\sigma_0^2}} \quad \frac{a_0 e(\,)(x-\mu_0)}{2\sigma_0^2}\right.$$

$$\frac{a_0 e(\,)(x-\mu_0)}{\sigma_0^3}$$

$$J_i \delta\beta_i = y_i - y_0$$

As $\delta\beta_1 = \delta\beta_2 \cdots \delta\beta_m = \delta\beta$

$$J_i \delta\beta = y_i - f(x_i, \beta_0)$$

So

$$\begin{bmatrix} J_1 \\ J_2 \\ \vdots \\ \vdots \\ J_m \end{bmatrix}_{m\times3} \delta\beta_{3\times1} = \begin{bmatrix} y_1 - f(x_1 - \beta_0) \\ y_2 - f(x_2, \beta_0) \\ \vdots \\ \vdots \\ y_m - f(x_m, \beta_0) \end{bmatrix}_{m\times1}$$

$\Longrightarrow$ Of the form

$$A_{m\times3}\,\delta\beta_{3\times1} = Y_{m\times1}$$

$$J_{m\times3}\,\delta\beta_{3\times1} = Y_{m\times1}$$

So, $\delta\beta = [J^T J]^{-1} J^T Y$

↳ End up with similar

$\beta_{k+1} = \beta_k \mp \delta\beta$ (or) $\beta(n+1) = \beta(n) \mp \delta\beta$

Linearise about $\beta(n+1)$ to solve for new $\delta\beta$ and keep doing the process time $\beta(n+1) - \beta(n) < \xi$ (or) $n >$ max iterations

LM algorithm:

$$[J^T J + \lambda I]\delta\beta = J^T [Y]$$

(or)

$$\delta\beta = [J^T J + \lambda I]^{-1} J^T Y$$

↳ Damping factor

↓

Check it out on paper / wikipedia