LECTURE III

# LEAST SQUARES

18/11/2019



Consider the differential equation

$$\frac{d}{dt} y = f(y, t, \mu)$$

giving the state $y \in \mathbb{R}^p$ of the system at time $t$

(e.g. $y \in \mathbb{R}^6$, orbital elements)

$\mu \in \mathbb{R}^{p'}$ dynamical parameters

Denote by $\Phi_{t_0}^{t}(y_0, \mu)$ the integral flow.

DEF Observation function $R = (R_1, \dots, R_K)$

$$R_j = R_j(y, t, \nu)$$

state at time $t$

$\nu \in \mathbb{R}^{p''}$ Kinematical parameters

DEF Prediction function

$$\tilde{r}(t) = R\left( \Phi_{t_0}^{t}(y_0, \mu), t, \nu \right)$$

We can group the multidimensional data and predictions into 2 vectors, with components $r_i$, $\tilde{r}_i(t_j)$. ← $r_i$ refers to time $t_j$

DEF Residuals $\xi = (\xi_1, \dots, \xi_m)$ $\quad \xi_i = r_i - \tilde{r}_i(t_j) \quad i = 1, \dots, m.$

observed computed

## Least squares method

principle of least squares: the solution of the OD problem makes the value of the target function

$$Q(\xi) = \frac{1}{m}\xi^T\xi$$

achieving its minimum value.

Note that $\quad \xi_i = \xi_i(y_0, \mu, \nu)$

Select some components of the vector $(y_0, \mu, \nu)$ to form the vector $\quad x \in \mathbb{R}^N \quad$ fit parameters

to be determined by the least squares fit.

Define the following quantities

- consider parameters : the remaining components of $(y_0, \mu, \nu)$

REQUIREMENT : $\quad m \geq N$

number of observations

number of parameters to be determined

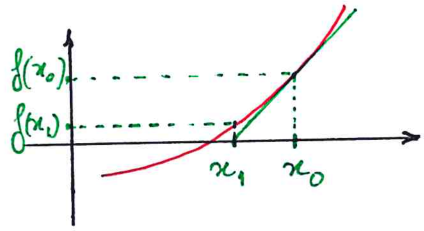
- design matrix $\quad B = \frac{\partial \xi}{\partial x}(x)$

We search for the minimum of $\; Q(x) = Q(\xi(x))\;$ by looking for its stationary points.

Therefore, we consider the equation

$$\boxed{\frac{\partial Q}{\partial x} = \frac{2}{m}\,\xi^T B = 0} \qquad (*)$$

Use Newton's method to search for solutions of $(*)$.




solve $f(x) = 0$

iterative method $\quad x_{k+1} = x_k - \dfrac{f(x_k)}{f'(x_k)}$

HINT Newton's method needs a starting guess (preliminary orbit)

In this case we have to solve $(*)$, so that we need

$$\frac{\partial^2 Q}{\partial x^2} = \frac{2}{m}\left(B^T B + \xi^T H\right)$$

where $H = \dfrac{\partial^2 \xi(x)}{\partial x^2}$ is a tensor with 3 indexes.

Set $\quad C_{New} = B^T B + \underline{\xi^T H}$

NxN matrix

the components of this matrix are $\sum_i \xi_i \dfrac{\partial^2 \xi_i}{\partial x_j \partial x_k}$

$C_{New}$ is a NON-NEGATIVE matrix in a neighborhood of a local minimum.

Given the residuals $\xi(x_k)$ obtained from $x_k$ at iteration $k$, from Newton (-Raphson) method we have

$$\boxed{x_{k+1} = x_k - C_{New}^{-1} B^T \xi}$$

Differential corrections : a variant of Newton's method

$$\boxed{x_{k+1} = x_k - (B^T B)^{-1} B^T \xi} \qquad (**)$$

where $B = B(x_k)$.

DEF $C = B^T B$ normal matrix

HINT In $(**)$ we used $C$ in place of $C_{New}$ of Newton's method.

Equation $(**)$, written as $C(x_{k+1} - x_k) = -B^T \xi$

is called normal equation.

Denote by $x_*$ the value of $x$ at convergence of the iterations.

DEF $\Gamma = C^{-1}$ covariance matrix

The value of $\Gamma$ in $x_*$ can be used to estimate the uncertainty of the solution of the differential correction algorithm:

$$Q(x) \simeq \underset{\substack{\uparrow \\ \text{minimal value}}}{Q(x_*)} + \frac{2}{m} \underbrace{\xi^T B}_{\substack{\uparrow \\ \text{is zero at } x_*}} (x - x_*) + \frac{1}{m} (x - x_*)^T \underset{\substack{\uparrow \\ C \text{ at } x_*}}{C} (x - x_*)$$

The eigenvalues of $\Gamma$ are proportional to the length of the axes of the confidence ellipsoids $\boxed{\frac{1}{m} (x - x_*)^T C (x - x_*) \leq \sigma^2}$

where $\sigma$ can be selected within a probabilistic interpretation of the observational errors.

## EXAMPLE (linear case)

data : $(\alpha_i, \delta_i)$ $\quad i = 1, .., m \quad m \geq 3$

right ascension — declination

Linear least squares fit to compute $\alpha, \dot{\alpha}, \ddot{\alpha}, \delta, \dot{\delta}, \ddot{\delta}$

at time $\bar{t} = \frac{1}{m} \sum_{i=1}^{m} t_i$

linear model: $\quad \alpha(t) = \alpha + \dot{\alpha}(t-\bar{t}) + \frac{1}{2}\ddot{\alpha}(t-\bar{t})^2$

residuals: $\quad \xi_i = \alpha_i - \alpha(t_i) \qquad \xi = (\xi_1, .., \xi_m)$

$$Q(\alpha, \dot{\alpha}, \ddot{\alpha}) = \frac{1}{m}\sum_{i=1}^{m} \xi_i^2 = \frac{1}{m}\sum_{i=1}^{m} \left| \alpha_i - \alpha - \dot{\alpha}(t_i - \bar{t}) - \frac{\ddot{\alpha}}{2}(t_i - \bar{t})^2 \right|^2$$

stationarity condition $\quad \frac{\partial Q}{\partial(\alpha, \dot{\alpha}, \ddot{\alpha})} = 0$

$$\begin{cases} \frac{\partial Q}{\partial \alpha} = -\frac{\cancel{2}}{\cancel{m}} \sum_i \left[ \alpha_i - \alpha - \dot{\alpha}(t_i - \bar{t}) - \frac{\ddot{\alpha}}{2}(t_i - \bar{t})^2 \right] = 0 \\ \frac{\partial Q}{\partial \dot{\alpha}} = -\frac{\cancel{2}}{\cancel{m}} \sum_i \left[ \quad '' \quad \right](t_i - \bar{t}) = 0 \\ \frac{\partial Q}{\partial \ddot{\alpha}} = -\frac{2}{m} \sum_i \left[ \quad '' \quad \right] \frac{(t_i - \bar{t})^2}{2} = 0 \end{cases}$$

We obtain the system

$$\underbrace{\begin{bmatrix} m & \sum_i (t_i - \bar{t}) & \frac{1}{2}\sum_i (t_i - \bar{t})^2 \\ \sum_i (t_i - \bar{t}) & \sum_i (t_i - \bar{t})^2 & \frac{1}{2}\sum_i (t_i - \bar{t})^3 \\ \frac{1}{2}\sum_i (t_i - \bar{t})^2 & \frac{1}{2}\sum_i (t_i - \bar{t})^3 & \frac{1}{4}\sum_i (t_i - \bar{t})^4 \end{bmatrix}}_{B^T B} \begin{pmatrix} \alpha \\ \dot{\alpha} \\ \ddot{\alpha} \end{pmatrix} = \underbrace{\begin{pmatrix} \sum_i \alpha_i \\ \sum_i \alpha_i (t_i - \bar{t}) \\ \frac{1}{2}\sum_i \alpha_i (t_i - \bar{t})^2 \end{pmatrix}}_{-B^T \xi}$$