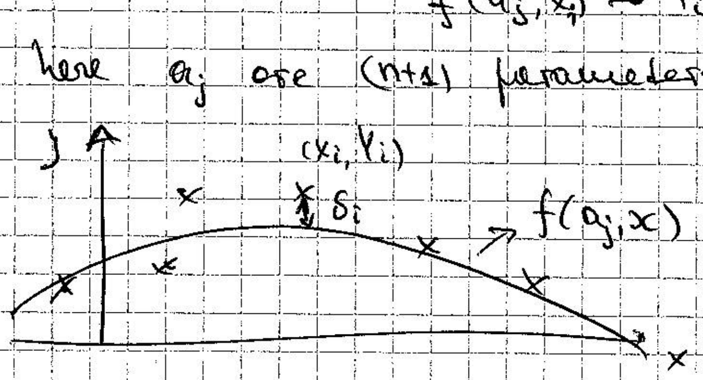
Lecture. „Function approximation and data fitting."

(1)

Problem: One has N data points (x_i, Y_i), which are to be represented by an approximation function $f(a_j, x)$, so that

$$f(a_j, x_i) \approx Y_i \quad \text{and} \quad \|f - Y_i\| = \min \|f(a_j, x) - Y_i\|$$

here a_j are $(n+1)$ parameters.



Any set of a_j gives some approx. of data set. One needs the „best" fit.

We will consider linear app. problem

$$f(a_j, x) = \sum_{j=0}^{n} a_j \varphi_j(x)$$, where $\varphi_i(x)$ are linear independent functions

$$= \sum_{j=0}^{n} \varphi_j(a_j, x)$$ → nonlinear appr. problem.

$\varepsilon_j = Y_j - f(a_j, x_i)$ – deviation.

!! Then approximation in the „root-mean-square" (as least square) sense, one takes the norm $\|\varepsilon_i\|^2 = \sum_{i=1}^{N} \varepsilon_i^2 =$

$$= \sum_{j=1}^{N} \left(Y_i - f(a_j, x_i)\right)^2 \Rightarrow \text{minimum condition is}$$

$$\frac{\partial \sum_{i=1}^{N} \varepsilon_i^2}{\partial a_j} = 2 \sum_{i=1}^{N} \left(Y_i - f(a_j, x_i)\right) \frac{\partial f(a_j, x_i)}{\partial a_j} = 0, \; j = 0 \ldots n.$$

$$= 2 \sum_{i=1}^{N} \left(Y_i - f(a_j, x_i)\right) \varphi_j(x_i) = 0 \Rightarrow$$

normal equation.

$$\sum_{i=1}^{N} Y_i \varphi_j(x_i) = \sum_{i=1}^{N} \sum_{k=0}^{n} a_k \varphi_k(x_i)\varphi_j(x_i)$$

normal equation for linear least square problem.

① $\sum_{k=0}^{n} a_k A_{kj} = C_j \iff \boxed{\vec{a}\cdot A = \vec{C}}$ → linear system of equation. can be solve by appropriate method

can be singular

$\vec{C} = \left(\sum_{i=1}^{N} Y_i \varphi_0(x_i), \ldots \sum_{i=1}^{N} Y_i \varphi_{n+1}(x_i)\right)$

$\vec{a} = (a_0, \ldots a_n)$

$$A = \begin{pmatrix} \sum_{i=1}^{N} \varphi_0(x_i)\varphi_0(x_i) & \cdots & \sum_{i=1}^{N} \varphi_0(x_i)\varphi_n(x_i) \\ & \ddots & \\ & & \sum_{i=1}^{N} \varphi_n(x_i)\varphi_n(x_i) \end{pmatrix}$$

symmetric $(n+1)\times(n+1)$ matrix.

② if one defines: $\Phi = \begin{pmatrix} \varphi_0(x_1) & \cdots & \varphi_n(x_1) \\ \vdots & & \vdots \\ \varphi_0(x_N) & \cdots & \varphi_n(x_N) \end{pmatrix}$ $(n+1)\times N$

$$A = \Phi^T \Phi \;;\quad \vec{C} = \Phi^T \vec{Y} \Rightarrow$$

$$A\vec{a} = \vec{C} \Rightarrow \Phi^T \Phi \vec{a} = \Phi^T \vec{Y}$$

$\boxed{\Phi \vec{a} = \vec{Y}}$ → over determined linear system.

__Remark 1__

if Y_i is known with different error, data point can be included in approximation function with different weight w_i →

$$f(\vec{a}; x_i) = \sum_{j=0}^{n} w_i a_j \varphi_j(x_i) \quad \Big\| \quad \frac{\partial f}{\partial a_j} = w_i \varphi_j(x)$$

⇓

in normal equation ⇒ $W = w^2$

$$\sum_{i=1}^{N} \left(Y_i - \sum_{k=0}^{n} w_i a_k \varphi_k(x)\right) w_i \varphi_j(x) = 0$$

⇓

$$\sum_{i=1}^{N} \sum_{k=0}^{n} a_k w_i^2 \varphi_k(x_i)\varphi_j(x_i) = \sum_{i=1}^{N} Y_i w_i^2$$

Remark 2 Why "mean-root-squares"?

if "measurement" errors of Y_i are independently random and have a normal (Gauss) distribution around the "true" function with standard deviation σ equal for of points (x_i, Y_i) $\Rightarrow$ the probability of the all data set is

$$P \sim \prod_{i=1}^{N} \left\{ \exp\left(-\frac{1}{2}\left(\frac{Y_i - f(x_i)}{\sigma}\right)^2\right)\right.$$

maximizing $\Leftrightarrow$ ~~minimizing~~ maximizing its log $\Leftrightarrow$

to minimizing its $-\log$ $\Rightarrow$

$$\boxed{\min\left(\sum_{i=1}^{N}\frac{1}{2\sigma^2}\left(Y_i - f(x_i)\right)^2\right)}\ , \text{i.e.}$$

the least-square fitting is a maximum likehood estimation of the fitted parameters, if the ~~—~~ errors are independent and normaly distributed with constant standard deviation

Legendre Approximation (Orthogonal functions)

Let set of functions $\varphi_j(x)$ are ~~orthogonal~~ orthonormal $\Rightarrow$

$$\int W(x)\varphi_k(x)\varphi_j(x)\,dx = \delta_{kj} \Rightarrow$$

$$\Downarrow\quad A_{kj} = \sum_{i=1}^{N} W_i\,\varphi_k(x_i)\,\varphi_j(x_i) \xrightarrow[N\to\infty]{} N\delta_{kj} \text{ is diagonal matrix}$$

$$\sum_{k=0}^{n} a_k \underbrace{\sum_{i=1}^{N} W_i\,\varphi_k(x_i)\,\varphi_j(x_i)}_{N\delta_{kj}} = \sum_{i=1}^{N} W_i Y_i\,\varphi_j(x_i)$$

$$N\to\infty \Rightarrow\quad N\begin{pmatrix} a_0 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix} \cong \begin{pmatrix} \sum_{i=1}^{N} W_i Y_i\,\varphi_0(x_i) \\ \vdots \\ \vdots \end{pmatrix} =$$

(*) $\boxed{a_j \cong \frac{1}{N}\sum_{i=1}^{N} W_i \varphi_j(x_i) Y_i}$ for orthogonal functions one should

normalizes (*) =>

$$a_j \cong \left(\sum_{i=1}^{N} w_i \varphi_j(x_i) Y_i\right)\left(\sum_{i=1}^{N} w_i \varphi_j^2(x_i)\right)^{-1}.$$

Remark: Fourier Series as least squares approximation

(**) $[-1,1]$ => strictly

$\Downarrow$ One can chose any set of orthonormal functions.

For example: trigonometric functions (valid also on the equally spaced discrete set of points):

$$\begin{cases} \sum_{i=0}^{N} \sin(k\pi x_i)\sin(j\pi x_i) = \sum_{i=0}^{N} \cos(k\pi x_i)\cos(j\pi x_i) = N\delta_{kj} \\ \text{also if} \quad x_i = \frac{2i-N}{N};\ i=0,\dots N \end{cases}$$

$\Downarrow$ in the sence of LSA:

$$\begin{cases} a_j \overset{(**)}{=} \frac{1}{N+1}\sum_{i=1}^{N} \underbrace{f(x_i)}_{Y_i}\cos(j\pi x_i) \\ b_j = \frac{1}{N+1}\sum_{i=1}^{N} f(x_i)\sin(j\pi x_i) \end{cases}$$

in continuous case $\sum \to \int$ =>

$$a_j[b_j] = \int_{-1}^{+1} f(x)\cos(j\pi x)[\sin(\pi x)]\,dx$$

which are coeff. of Fourier series:

$$f(x) = \frac{1}{2}a_0 + \sum_{k=1}^{\infty} a_k\cos(k\pi x) + b_k\sin(k\pi x)$$

Error Analysis for Least Square Coeff-s. (5)

1) Assume that Y_i is known with error $\Rightarrow$

2) Let a_j^0 coeffs. obtained assuming E_i one has corrected fit for this error $\Rightarrow$

$$\sum_i w_i^2 \varphi_k(x_i) \Rightarrow \begin{cases} E_i = Y_i - \sum_{j=0}^{M} a_j^0 \varphi_j(x_i) \\ \varepsilon_i = Y_i - \sum_{j=0}^{M} a_j \varphi_j(x_i) \end{cases}$$

residual error

not corrected coeff.

in LSA $\Rightarrow$ $\sum_{i=1}^{N} (w_i \varepsilon_i)^2 = \min$, we are interested in

$$\delta a_j \equiv a_j - a_j^0$$

error in $\vec{a}$ due to E_i

$$\sum_{i=0}^{n} w_i^2 \varphi_k(x_i) E_i = \sum_{i=0}^{n} \left(Y_i - \sum_{j=0}^{n} a_j^0 \varphi_j(x_i) \right) w_i^2 \varphi_k(x_i)$$

$$\sum_{j=0}^{N} a_j^0 \sum_{i=0}^{n} w_i^2 \varphi_j(x_i) \varphi_k(x_i) = \sum_{i=1}^{N} w_i^2 Y_i \varphi_k(x_i) - \sum_{i=1} w_i^2 \varphi_k(x_i) E_i$$

$k = 0 \ldots n$

standart normal equation is

$$(*) \sum_{j=0}^{N} a_j \sum_{i=0}^{N} w_i^2 \varphi_j(x_i) \varphi_k(x_i) = \sum_{i=1}^{N} w_i^2 Y_i \varphi_k(x_i)$$

substruct for $(*)$

$$\sum_{j=0}^{M} \delta a_j \underbrace{\sum_{i=0}^{N} w_i^2 \varphi_j(x_i) \varphi_k(x_i)}_{A_{jk}} = \sum_{j=0}^{N} \delta a_j A_{jk} = \sum_{i=1} w_i^2 \varphi_k(x_i) E_i$$

$A_{jk} \Rightarrow$

$$\delta a = A^{-1} \begin{pmatrix} \sum_{i=1}^{N} w_i^2 \varphi_0 E_i \\ \vdots \\ \sum_{i=1}^{N} w_i^2 \varphi_n E_i \end{pmatrix}$$

$C \equiv A^{-1}$

$\delta a_i = \sum_{k=0}^{n}$

$$(\delta a_j)^2 = \left[\sum_{k=0}^{n} C_{jk} \sum_{i=1}^{N} w_i \varphi_k(x_i) E_i\right]\left[\sum_{p=0}^{n} C_{jp} \sum_{q=1}^{N} w_q \varphi_p(x_q) E_q\right] =$$

$$= \sum_{k=0}^{n}\sum_{p=0}^{n} C_{jk} C_{jp} \sum_{i=1}^{N}\sum_{q=1}^{N} w_i w_q \varphi_k(x_i)\varphi_p(x_q) E_i E_q$$

Assumptions 1) $w_i E_i =$ const., one is chosing weight w_i in this way

2) Error distribution is anti-symmetric about zero.

$\Downarrow$ if say $E_i = -E_q$ $\qquad \overbrace{-x_q = x_i}$ $\quad q \quad 0 \quad i$

$$\sum_{\substack{i=1\\ i\neq q}}^{N} \sum_{q=1}^{N} w_i w_q \underbrace{\varphi_k(x_i)\varphi_p(x_q)}_{\text{symmetric}} E_i E_q = 0$$

$|w_i E_i| = |w_q E_q| \Longrightarrow$

$$(\delta a_j)^2 = \underset{\text{mean value}}{\overline{(wE)^2}} \sum_{k=0}^{n} C_{jk} \sum_{p=0}^{n} C_{jp} \underbrace{\sum_{i=1} w_i \varphi_k(x_i)\varphi_p(x_i)}_{A_{kp}} =$$

$$= \overline{(wE)^2} \sum_{k=0}^{n} C_{jk} \underbrace{\left(\sum_{p=0}^{n} C_{jp} A_{pk}\right)}_{\delta_{jk}} =$$

diagonal elements of A^{-1}

$$= \overline{(wE)^2} \sum_{k=0}^{n} C_{jk}\delta_{jk} = \overline{(wE)^2}\, C_{jj} = (\delta a_j)^2$$

$\overline{(wE)^2} - ?$ $\qquad$ A - normal eq matrix

~~$E_i - \varepsilon_i = \sum_{j=0}^{n} \delta a_j \varphi_j(x_i) = \sum_{j=0}^{n} \varphi_j(x_i)$~~

$$E_i - \varepsilon_i = \sum_{j=0}^{n} \delta a_j \varphi_j(x_i) = \sum_{j=0}^{n} \varphi_j(x_i) \sum_{k=0}^{n} C_{jk} \sum_{q=1}^{N} w_q \varphi_k(x_q) E_q$$

$* \sum_i w_i \varepsilon_i$

7

$$\sum_{i=1}^{N} W_i \varepsilon_i E_i - \sum_{i=1}^{N_i} \varepsilon_i^2 = \sum_{i=1}^{N} W_i \varepsilon_i \left(\sum_{j=0}^{n} \varphi_j(x_i) \sum_{k=0}^{n} C_{jk} \sum_{q=1}^{N} w_q \varphi_k(x_q) E_q \right) =$$

$$= \sum_{j=0}^{n} \sum_{k=0}^{n} \sum_{q=1}^{N} C_{jk} w_q \varphi_k(x_q) E_q \left(\sum_{i=1}^{N} W_i \varepsilon_0 \varphi_j(x_i) \right)$$

$$\frac{\partial \sum_{i=1}^{N} w_i \varepsilon_i^2}{\partial a_j} = 2 \sum_{i=1}^{N} w_i \varepsilon_i \frac{\partial \varepsilon_i}{\partial a_j} = 2 \sum_{i=1}^{N} \varepsilon_i w_i \varphi_j(x_i) = 0$$

def. LSA.

$$\Rightarrow \sum_{i=1}^{N} w_i \varepsilon_i E_i = \sum_{i=1}^{N} W_i \varepsilon_i^2$$

$$(E_i - \varepsilon_i) =$$

$$* \sum_i W_i E_i \Rightarrow \sum_{i=1}^{N} w_i E_i^2 - \sum_{i=1}^{N} W_i \varepsilon_i E_i =$$

$$= \sum_{i=1}^{N} W_i E_i \sum_{j=0}^{n} \delta a_j \varphi_j(x_i) = \sum_{j=0}^{n} \sum_{k=0}^{n} \sum_{q=1}^{N} \sum_{i=1}^{N} C_{jk} w_q \varphi_j(x_i) \varphi_k(x_q) E_q E_i =$$

$$= \sum_{j=0}^{n} \sum_{k=0}^{n} C_{jk} \sum_{i=1}^{N} W_i^2 E_i^2 \varphi_j(x_i) \varphi_k(x_i)$$

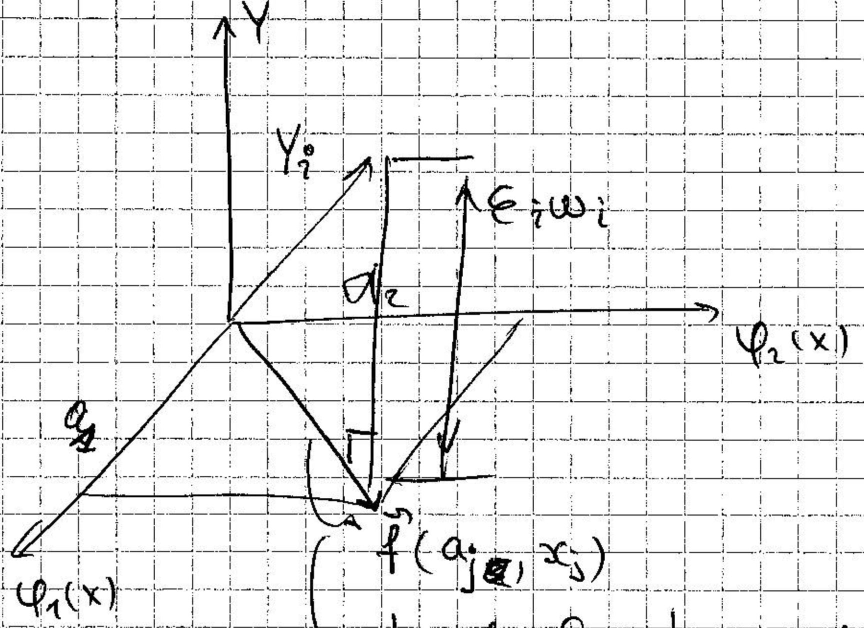
$$N\overline{(W_i E_i)^2} - \sum_{i=1}^{N} W_i^2 E_i \varepsilon_i = \overline{(WE)^2} \underbrace{\sum_{j=0}^{n} \sum_{k=0}^{n} C_{jk} A_{jk}}_{n}$$

$$N\overline{(WE)^2} - \sum_{i=1}^{N} W_i^2 \varepsilon_i^2 = n\overline{(WE)^2}$$

$$\overline{(WE)^2} = \frac{1}{N-n} \sum_{i=1}^{N} W_i^2 \varepsilon_i^2 \Rightarrow$$

$$\boxed{(\delta a_j)^2 = \frac{C_{jj}}{N-n} \sum_{i=1}^{N} W_i^2 \varepsilon_i^2}$$

one should find residual error ε_i^2



should be perpendicular to minimise $\varepsilon_i \Rightarrow$

$$\sum_{i=1}^{N} (\omega_i \varepsilon_i)^2 = \sum_{i=1}^{N} \omega_i^2 Y_i^2 - \sum_{i=1}^{N} \omega_i^2 \vec{f}^2(a_j, x_i)$$

Pythagorus theorem

$$\sum_{i=1}^{N} (\omega_i \varepsilon_i)^2 = \sum_{i=1}^{N} \omega_i^2 \left(Y_i - f(a_j, x_i)\right)^2 =$$

$$= \underline{\sum_{i=1}^{N} \omega_i^2 Y_i^2} - 2\sum_{i=1}^{N} \omega_i^2 Y_i f(a_j, x_i) + \underline{\sum_{i=1}^{N} \omega_i^2 f^2(a_j, x_i)}$$

$$\parallel$$

$$\sum_{i=1}^{N} \omega_i^2 Y_i^2 - \sum_{i=1}^{N} (\omega_i \varepsilon_i)^2$$

$$\cancel{2}\sum_{i=1}^{N} (\omega_i \varepsilon_i)^2 = \cancel{2}\sum_{i=1}^{N} (\omega_i Y_i)^2 - \cancel{2}\sum_{i=1}^{N} \omega_i^2 Y_i \underbrace{f(a_j, x_i)}_{\sum_{k=0}^{n} a_k \varphi_k(x_i) \omega_i^2}$$

$$\Downarrow$$

$$(\delta a_j)^2 = \frac{C_{jj}}{N-n} \left(\underbrace{\sum_{i=1}^{N} \omega_i^2 Y_i^2}_{\text{extra}} - \sum_{k=0}^{n} a_k \sum_{i=1}^{N} \omega_i^4 Y_i \varphi_k(x_i) \right)$$