Curve fitting

\#4

(x_i, y_i)

$$y_i = f(x_i) + \varepsilon_i$$

$$f(x) = a_1 f_1(x) + a_2 f_2(x) + \dots a_m f_m(x)$$

$$a = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} \qquad F(x) = F_x = [f_1(x) \dots f_m(x)]$$

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \qquad x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \qquad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_m \end{bmatrix}.$$

$$y = F_x a + \varepsilon$$

m - fixed

n - big enough.

$$Q(a) = \sum_{i=1}^{n} (y_i - f_a(x_i))^2 \sim \min_a$$

$$Q(a) = \|y - Ba\|^2 =$$

$$= \|B(a - (B^T B)^{-1} B^T y\|^2$$

$$+ \|y\|^2 - \|B(B^T B)^{-1} B^T y\|^2$$

$$B = \begin{bmatrix} F_{x_1} \\ F_{x_2} \\ \vdots \\ F_{x_n} \end{bmatrix} = \begin{bmatrix} f_1(x_1) & \cdots & f_m(x_1) \\ \vdots & & \\ f_1(x_n) & \cdots & f_m(x_n) \end{bmatrix} \quad \underline{\underline{n \times m}}$$

if $\hat{a} = (B^T B)^{-1} B^T y \Rightarrow Q(\hat{a}) \sim \min$

<u>Columns of B</u> are independent.

$n \geq m$, $f_1 \ldots f_m$ - lin indep

$x_1 \ldots x_n$ "random enough"

<u>Col. of B</u> indep $\Rightarrow$ $Bu = 0 \Rightarrow u = 0$

$n \square_m$

$B^T B = m\square_m$

rank $B = m$

$\underline{B^T B u = 0 \overset{?}{\Rightarrow} u = 0}$

$$0 = \langle B^T B u, u \rangle = \langle Bu, Bu \rangle = \|Bu\|^2$$

$$\Rightarrow \|Bu\| = 0 \Rightarrow Bu = 0 \Rightarrow u = 0$$

$\Rightarrow$ rank $B^T B = m \Rightarrow B^T B$ is invertible.

$$\exists !\ \hat{a} : Q(\hat{a}) \sim \min.$$

$$\hat{a} = \underbrace{(B^T B)^{-1}}\ \underbrace{B^T y}$$

$$\left.\begin{matrix}(x_1, y_1)\\ \vdots \\ (x_n\ y_n)\end{matrix}\right] \Rightarrow y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix},\quad B = \begin{bmatrix} f_1(x_1) & \cdots & f_m(x_1) \\ & & \\ f_1(x_n) & \cdots & f_m(x_n) \end{bmatrix}$$

$$(x_{n+1}, y_{n+1}) \qquad y_{n+1} \qquad f_1(x_{n+1}) \cdots f_m(x_{n+1})$$

$$B^T y = \begin{bmatrix} F_1 \\ \vdots \\ F_n \end{bmatrix}^T y \qquad \Big|\quad F_i = F_{x_i} = F(x_i)$$

$$= \left[F_1^T \cdots F_n^T\right]\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = y_1 F_1^T + \cdots + y_n F_n^T$$

$$= \sum_{i=1}^n y_i F_i^T = \sum_{i=1}^n y_i \underbrace{\begin{bmatrix} f_1(x_i) \\ \vdots \\ f_m(x_i) \end{bmatrix}}_{v_i} = \sum_{i=1}^n v_i = v$$

$$(x_i, y_i) \mapsto v_i = y_i \begin{bmatrix} f_1(x_i) \\ \vdots \\ f_m(x_i) \end{bmatrix}$$

v_i

$1m$-vector

$$B^T B = \begin{bmatrix} F_1^T & \cdots & F_n^T \end{bmatrix} \begin{bmatrix} F_1 \\ \vdots \\ F_n \end{bmatrix} \quad - \quad m \times m$$

$m \quad n \qquad n \quad m$

$$= F_1^T F_1 + \cdots + F_n^T F_n =$$

$$= \sum_{i=1}^{n} \underbrace{F_i^T F_i}_{= T_i} = \sum_{i=1}^{n} T_i = T$$

$$T_i = F_i^T F_i = \begin{bmatrix} f_1(x_i) \\ \vdots \\ f_m(x_i) \end{bmatrix} \begin{bmatrix} f_1(x_i) & \cdots & f_m(x_i) \end{bmatrix} =$$

$$= \begin{bmatrix} f_1^2(x_i) & f_1(x_i) f_2(x_i) & \cdots & f_1(x_i) f_m(x_i) \\ \vdots & & & \vdots \\ f_m(x_i) f_1(x_i) & & \cdots & f_m^2(x_i) \end{bmatrix}$$

$m \times m$

(T, z) $\qquad \hat{a} = T^{-1} z$

$m \quad m \qquad m$

$$
\begin{tikzcd}
\text{Set}_1: \left.\begin{matrix}(x_1, y_1)\\ \vdots \\ (x_n, y_n)\end{matrix}\right] \arrow[r, mapsto] & (T, z) \arrow[d] & \\
 & \oplus \arrow[r] & (T+T', z+z') \\
\text{Set}_2 \left.\begin{matrix}(x_1', y_1')\\ \vdots \\ (x_n', y_n')\end{matrix}\right] \arrow[r] & (T', z') \arrow[u] &
\end{tikzcd}
$$

Accumulate:

$$(x_i, y_i) \mapsto (T_i, z_i) \Rightarrow (T, z)$$

$$(T, z) \mapsto \hat{a} = T^{-1} z.$$

$\hat{a}$ - inbiased estimate of a:

$$\hat{a} = \underbrace{(B^T B)^{-1} B^T}_{= R} y = R y$$

$$y = Ba + \varepsilon$$

$$\hat{a} - a = (B^T B)^{-1} B^T y - a$$

$$= (B^T B)^{-1} B^T (Ba + \varepsilon) - a$$

$$= \Big(\underbrace{(B^T B)^{-1} B^T B}_{= I} - I\Big) a + R\varepsilon$$

$$= R\varepsilon$$

$$\underline{E(\hat{a} - a)} = E R \varepsilon = R E \varepsilon = R E \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix} =$$

$$= R \begin{bmatrix} E\varepsilon_1 \\ \vdots \\ E\varepsilon_n \end{bmatrix} = R O = \underline{O}$$

$$\varepsilon_i : \quad \underline{E\varepsilon_i = 0} \qquad Var\,\varepsilon_i = E\varepsilon_i^2 = \sigma^2.$$

$$E(\hat{a} - a) = 0 \implies \hat{a} - \text{unbiased}$$

Accuracy of $\hat{a}$.

$$Var(\hat{a}) = E(\hat{a} - E\hat{a})(\hat{a} - E\hat{a})^T =$$

(variance matr. or var.covar matr)

$$= E(\hat{a} - a)(\hat{a} - a)^T = E R \varepsilon \varepsilon^T R^T$$

$$= R E \varepsilon \varepsilon^T R^T$$

$$E\varepsilon\varepsilon^T = E \begin{bmatrix} \varepsilon_1\varepsilon_1 & \varepsilon_2\varepsilon_2 & \cdots & \varepsilon_1\varepsilon_n \\ \vdots & & & \\ \varepsilon_n\varepsilon_1 & \cdots & \cdots & \varepsilon_n^2 \end{bmatrix}$$

$$E\varepsilon_i^2 = \sigma^2 = \mathrm{Var}\,\varepsilon_i$$

$\{\varepsilon_i\}$ - indep. $E\varepsilon_i\varepsilon_j = 0$ if $i \neq j$ $\Rightarrow$

$$E\varepsilon\varepsilon^T = \mathrm{Var}\,\varepsilon = \sigma^2 I$$

$$\mathrm{Var}(\hat{a}) = R\cdot\sigma^2 I\cdot R^T = \sigma^2 R R^T$$

$$= \sigma^2 \underbrace{(B^TB)^{-1}B^T}_{R} B (B^TB)^{-1}$$

$$= \sigma^2 (\underbrace{B^TB}_{=T})^{-1} = \sigma^2 T^{-1}$$

$$\mathrm{Var}(\hat{a}_j) = E(\hat{a}_j - a_j)^2 = \underline{\sigma^2 (T^{-1})_{jj}}$$

$$\mathrm{Cov}(\hat{a}_j, \hat{a}_k) = E(\hat{a}_j - a_j)(\hat{a}_k - a_k) = \sigma^2 (T^{-1})_{jk}$$

Estimation of $f(x)$

$$\widehat{f(x)} = \hat{a}_1 f_1(x) + \dots + \hat{a}_m f_m(x)$$

$$= F_x \hat{a} \qquad F_x = [f_1(x) \cdots f_m(x)]$$

$$E(\widehat{f(x)} - f(x)) = E(F_x\hat{a} - F_x a) =$$

$$= E F_x(\hat{a} - a) = F_x \underbrace{(E\hat{a} - a)}_{=0} = 0$$

$\widehat{f(x)}$ - unbiased est. of $f(x)$

$$Var(\widehat{f(x)}) = E(\widehat{f(x)} - f(x))^2 =$$

$$= E(F_x(\underbrace{\hat{a} - a}_{R\varepsilon}))^2 = E(F_x R\varepsilon (F_x R\varepsilon)^T)$$

$$\overline{\ m\ } \quad m\square_n \quad |n$$

$$= E(F_x R \varepsilon \varepsilon^T R^T F_x^T)$$

$$= F_x R \cdot \underbrace{E(\varepsilon \varepsilon^T)}_{= \sigma^2 I} \cdot R^T F_x^T = \sigma^2 F_x \underbrace{R R^T}_{= T^{-1}} F_x^T$$

$$= \sigma^2 F_x T^{-1} F_x^T$$

$$\overline{\ m\ } \quad m\square_m \quad |m$$

$$\begin{array}{c} (T, b) \\ \downarrow \end{array}$$

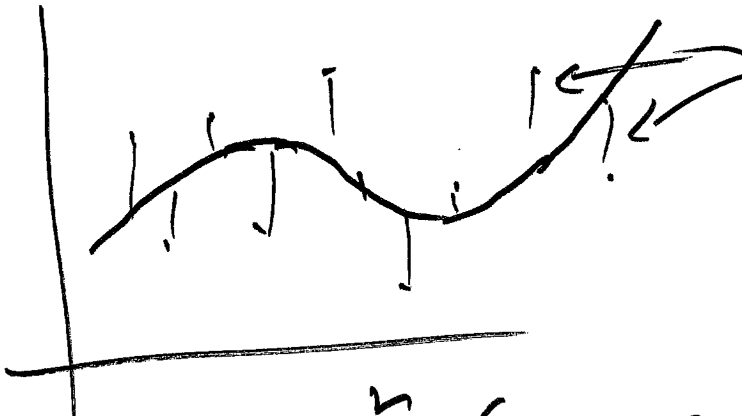
$$x \longrightarrow \circledast \xleftarrow{\text{apply}} \qquad \longmapsto \widehat{f(x)},\ Var(\widehat{f(x)})$$

$$\widehat{f(x)} = F_x T^{-1} b$$

$$Var(\widehat{f(x)}) = \sigma^2 F_x T^{-1} F_x^T$$

Assumed that σ^2 is given.

If σ^2 is not given.
need to estimate it.



$$Q(\hat{a}) = \sum_{i=1}^{n} \left(y_i - f_a(x_i)\right)^2$$

$$E\, Q(\hat{a}) = (n-m)\,\sigma^2 \quad \text{(Later)}$$

$$\hat{\sigma}^2 = \frac{Q(\hat{a})}{n-m}$$ – unbiased est of σ^2

$$Q(\hat{a}) = \|y - B\hat{a}\|^2 = \|y\|^2 - \|B\underbrace{(B^TB)^{-1}B^Ty}_{\hat{a}}\|^2$$

$$= \|y\|^2 - \|B\hat{a}\|^2$$

$$= \underbrace{\sum_{i=1}^{n} y_i^2}_{=V} - \underbrace{(B\hat{a})^T}_{n} \cdot B\hat{a}\Big|_n$$

$$= V - \hat{a}^T \underbrace{B^T B}_{=T} \hat{a}$$

$$Q(\hat{a}) = V - \hat{a}^T T \hat{a} \qquad \hat{a} = T^{-1} z$$

$$= V - z^T \underbrace{T^{-1} T T^{-1}}_{= T^{-1}} z$$

$$= V - z^T T^{-1} z$$

$$V = \sum_{i=1}^{n} \underbrace{y_i^2}_{= V_i} = \sum_{i=1}^{n} V_i$$

$$\widehat{\sigma^2} = \frac{V - z^T T^{-1} z}{n - m}$$

need to add V and n to can. info.

Can info: $\left(T, z, V, n \right)$

$$m \underset{m}{\square} \quad |m \quad \bullet \quad \bullet$$

$$\left.\begin{matrix}(x_1, y_1)\\ \vdots \\ (x_n, y_n)\end{matrix}\right] \mapsto (T, z, v, n)$$

$$\hat{a} = T^{-1} z$$

$$\widehat{Var}(\hat{a}) = \hat{\sigma}^2 T^{-1}$$

$$\hat{\sigma}^2 = \frac{v - z^T T^{-1} z}{n - m}$$

apply

$$x \longrightarrow \circledast \Longrightarrow$$

$$\widehat{Var}(f(x)) = \hat{\sigma}^2 F_x T^{-1} F_x^T$$

$$\widehat{f(x)} = F_x T^{-1} z$$

$$(T, z, n, v) \oplus (T', z', n', v') =$$

$$= (T + T', z + z', n + n', v + v')$$

Simple lin. regression as a particular case.

$$f(x) = a + bx = F_x\, a$$

$$= \underbrace{[1 \;.\; x]}_{F} \begin{bmatrix} a \\ b \end{bmatrix}$$

σ^2 is known -(for simplicity.)

$$(x_i\; y_i)_{i=1,n.} \longmapsto (n, X, Y, Z, U)$$

$$X = \sum_i x_i \quad Y = \sum_i y_i \quad Z = \sum_i x_i y_i$$

$$U = \sum_i x_i^2$$

In gen. case can. Ino.: (T, γ)

$$T = \sum_{i=1}^{n} F_i^T F_i \;, \qquad F_i = [1 \;\; x_i]$$

$$= \sum_{i=1}^{n} \begin{bmatrix} 1 \\ x_i \end{bmatrix} [1 \;\; x_i] = \sum_{i=1}^{n} \begin{bmatrix} 1 & x_i \\ x_i & x_i^2 \end{bmatrix}$$

$$= \begin{bmatrix} \sum_i 1 & \sum_i x_i \\ \sum_i x_i & \sum_i x_i^2 \end{bmatrix} = \begin{bmatrix} n & X \\ X & U \end{bmatrix}$$

$$\gamma = \sum_i y_i F_i^T = \sum_{i=1}^{n} y_i \begin{bmatrix} 1 \\ x_i \end{bmatrix} = \begin{bmatrix} \sum_i y_i \\ \sum y_i x_i \end{bmatrix} = \begin{bmatrix} Y \\ Z \end{bmatrix}$$

$$\widehat{\begin{bmatrix} a \\ b \end{bmatrix}} = T^{-1} z = \begin{bmatrix} n & X \\ X & U \end{bmatrix}^{-1} \begin{bmatrix} Y \\ Z \end{bmatrix}$$

$$= \frac{1}{nU - X^2} \begin{bmatrix} U & -X \\ -X & n \end{bmatrix} \begin{bmatrix} Y \\ Z \end{bmatrix}$$

$$= \frac{1}{nU - X^2} \begin{bmatrix} UY - XZ \\ -XY + nZ \end{bmatrix}$$

$$\hat{a} = \frac{UY - XZ}{nU - X^2}$$

$$\hat{b} = \frac{nZ - XY}{nU - X^2} = \frac{Z - \frac{XY}{n}}{U - \frac{X^2}{n}}$$

$$\mathrm{Var}\left(\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix}\right) = \sigma^2 T^{-1} = \frac{\sigma^2}{nU - X^2} \begin{bmatrix} U & -X \\ -X & n \end{bmatrix}$$

$$\mathrm{Var}\, \hat{a} = \frac{\sigma^2 U}{nU - X^2}$$

$$\mathrm{Var}\, \hat{b} = \frac{\sigma^2 n}{nU - X^2}$$

$\hat{f}(x) = \hat{a} + \hat{b}x$ – unb est of $f(x)$

$$\mathrm{Var}\left(\hat{f}(x)\right) = \sigma^2 F_x T^{-1} F_x^T$$

$$= \sigma^2 [1 \quad x] \frac{1}{nU - X^2} \begin{bmatrix} U & -X \\ -X & n \end{bmatrix} \begin{bmatrix} 1 \\ x \end{bmatrix}$$

$$= \frac{\sigma^2}{nU - X^2} \underbrace{[1 \quad x] \begin{bmatrix} U & -X \\ -X & n \end{bmatrix}}_{[U - xX \quad -X + nx]} \begin{bmatrix} 1 \\ x \end{bmatrix}$$

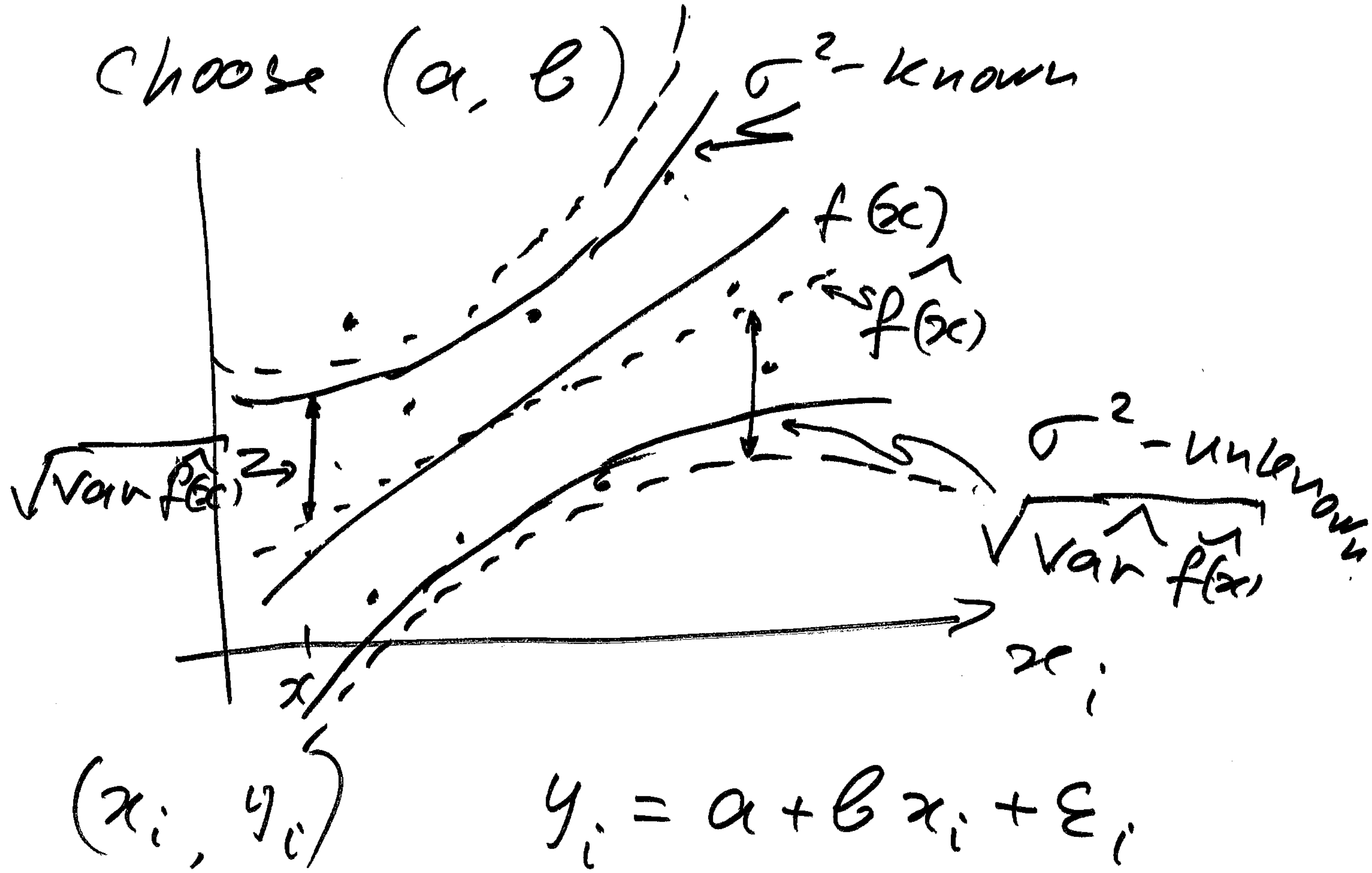
$$= \frac{\sigma^2}{nU - X^2} \left(U - xX - Xx + nx^2 \right)$$

$$-\frac{X^2}{n} + \frac{X^2}{n}$$

$$= \frac{\sigma^2}{nU - X^2} \left(U - \frac{X^2}{n} + \frac{X^2}{n} - 2Xx + nx^2 \right)$$

$$= \sigma^2 \left(\frac{1}{n} + \frac{\left(x - \frac{X}{n}\right)^2}{U - \frac{X^2}{n}} \right)$$

Comments on HW 2, Prob 2.



$y_i = a + b x_i + \varepsilon_i$

con. info. $\Rightarrow \hat{f}(x)$

$\quad\quad\quad\quad\quad\;\; \mathrm{var}\,\hat{f}(x)$

Estimation problems.

$y_1 = x + \varepsilon_1$ $\quad$ x - unknown

$y_2 = x + \varepsilon_2$ $\quad$ ε_i - i.i.d.

$E\varepsilon_i = 0 \quad E\varepsilon_i^2 = \sigma^2$

$\hat{x} = \frac{y_1 + y_2}{2}$ - unbiased est of x.

$\hat{x} = \frac{1}{3} y_1 + \frac{2}{3} y_2$

$\hat{x} = \alpha y_1 + \beta y_2$ $\quad$ if $\alpha + \beta = 1$

$\Rightarrow$ unbiased

$$E\hat{x} = E(\alpha(x + \varepsilon_1) + \beta(x + \varepsilon_2))$$

$$= \underbrace{(\alpha + \beta)}_{=1} x + \alpha \underbrace{E\varepsilon_1}_{=0} + \beta \underbrace{E\varepsilon_2}_{=} = x$$

$$y_i = x + \varepsilon_i \qquad i = 1, \dots n$$

(y_1, σ_1^2)

(y_2, σ_2^2)

$\vdots$

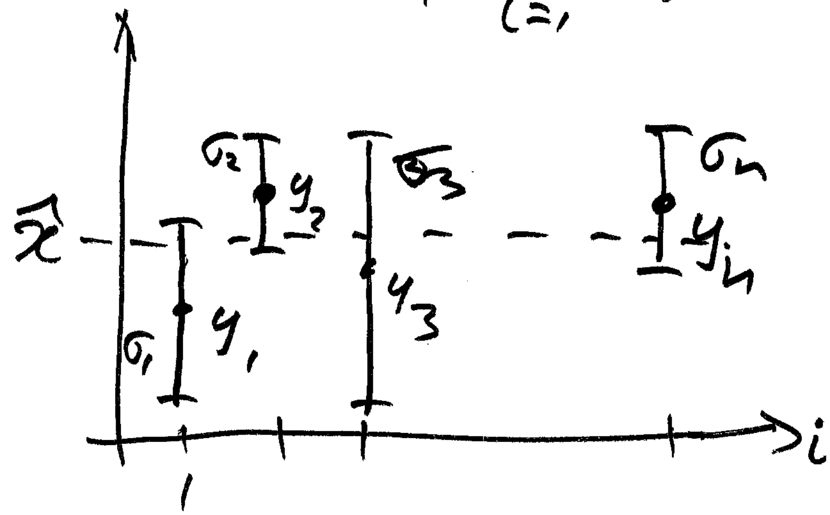
(y_n, σ_n^2)

$$E\,\varepsilon_i = 0 \quad E\,\varepsilon_i^2 = \sigma_i^2$$

ε_i independent, but not identically distributed

$$\hat{x} = \frac{1}{n}\sum_{i=1}^{n} y_i \quad ?$$

Is it good?



$$\hat{x} = \sum_{i=1}^{n} \alpha_i y_i$$

want $\hat{x}$ - unbiased est of x

$$\Rightarrow \sum_{i=1}^{n} \alpha_i = 1$$

$$E\,\hat{x} = E\sum_{i=1}^{n} \alpha_i (x + \varepsilon_i) = \underbrace{\sum_i \alpha_i}_{=1} x + \sum_i \alpha_i \underbrace{E\,\varepsilon_i}_{=0}$$

$$= x$$

Choose $\alpha_i \sim \dfrac{1}{\sigma_i^2} \Rightarrow \alpha_i = \dfrac{c}{\sigma_i^2}$

Unb $\Rightarrow 1 = \sum_i \alpha_i = \sum_i \dfrac{c}{\sigma_i^2} = c \cdot \sum \dfrac{1}{\sigma_i^2}$

$$\Rightarrow c = \frac{1}{\sum \frac{1}{\sigma_i^2}} \qquad \alpha_i = \frac{1}{\sigma_i^2 \sum \frac{1}{\sigma_i^2}}$$