Least-Squares Applications

Gabriel Hugh Elkaim

Least-Squares Applications

- Least-squares data fitting
- Growing sets of regressors
- System identification
- Growing sets of measurements and recursive least-squares

Least-Squares Data Fitting (1.3)

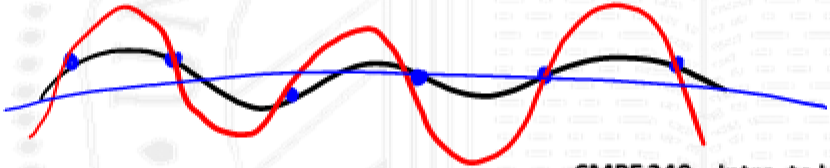
given a set of functions $\{f_1 \dots f_n\} : S \longrightarrow \mathbb{R}$ regressors basis functions

data measurements (s_i, g_i) $i = 1..m$ $g_i \in \mathbb{R}$

$s_i \in S$

Find the coefficients $x_1 \dots x_n \in \mathbb{R}$

$$x_1 f_1(s_i) + x_2 f_2(s_i) + \dots + x_n f_n(s_i) \simeq g_i \quad i = 1 \dots m.$$



Least-Squares Data Fitting (2.3)

choose x to min the mean square fitting error

$$\frac{1}{m}\sum_{i=1}^{m}\left(x_1 f_1(s_i) + \dots + x_n f_n(s_i) - g_i\right)^2$$

$$\min_x \|Ax - g\|^2 \qquad A_{ij} = f_j(s_i)$$

function "j" evaluated at point "s_i"

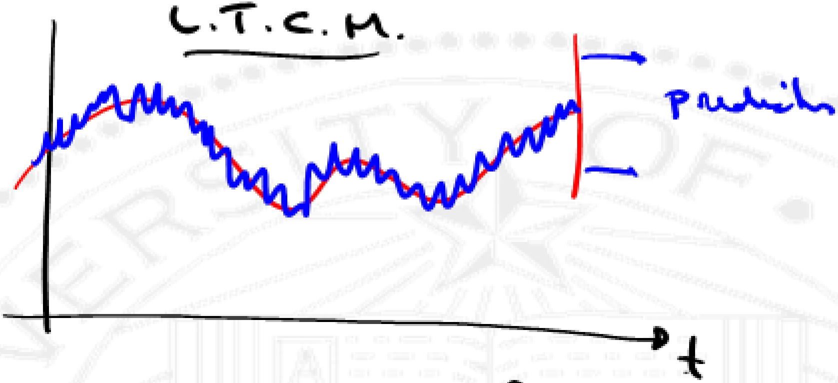
$$x = (A^T A)^{-1} A^T g$$

approximation of $f_{ls}(s) = x_1 f_1(s) + \dots + x_n f_n(s)$.

Least-Squares Data Fitting (3.3)

$f_1 \ldots f_n$ smooth

- Interpolation
- Extrapolation
- Data smoothing
- Develop a simplified approximate model for my data.



Celestial Orbits

LS Polynomial Fitting (1.2)

Polynomial of order n will hit n points exactly.

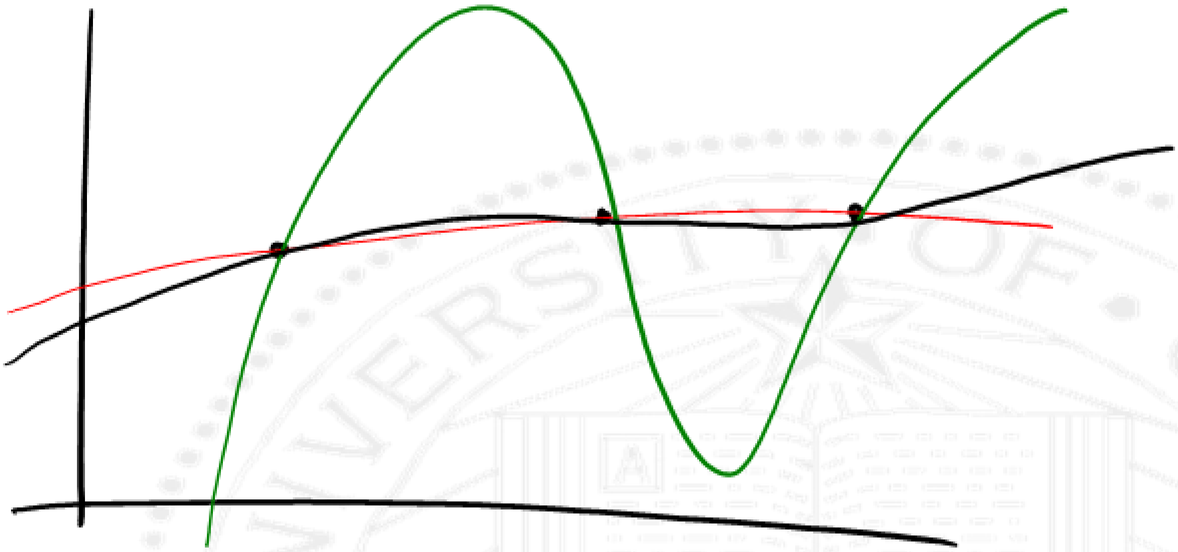
fit a polynomial of order < n.

$$p(t) = a_0 + a_1 t + a_2 t^2 + \dots + a_{n-1} t^{n-1}$$

fit this to data (t_i, y_i) $i = 1 \dots n$.

Basis function $f_j(t) = t^{j-1}$ $j = 1 \dots n$.

$$A_{ij} = t_i^{j-1}$$



$\hat{a} = (A^TA)^{-1}A^Ty$.

LS Polynomial Fitting (2.2)

$$A = \underbrace{\begin{bmatrix} 1 & t_1 & t_1^2 & t_1^3 & \cdots & t_1^{n-1} \\ \vdots & & & & & \\ 1 & t_m & t_m^2 & t_m^3 & \cdots & t_m^{n-1} \end{bmatrix}}_{\text{Vandermonde Matrix}}$$

$t_k \neq t_l \quad k \neq l.$

$m \gg n$

p has no more than $n-1$ zeros

p is identically 0 only when $a=0$

$$\boxed{p(t) = a_0 + a_1 t + \cdots + a_{n-1} t^{n-1}}$$

$Aa = 0$

$g(t) = \frac{4t}{(1-10t^2)}$

$m = 100$

$t = 0.01$

LS Polynomial Example

$a_0 + a_1 t$

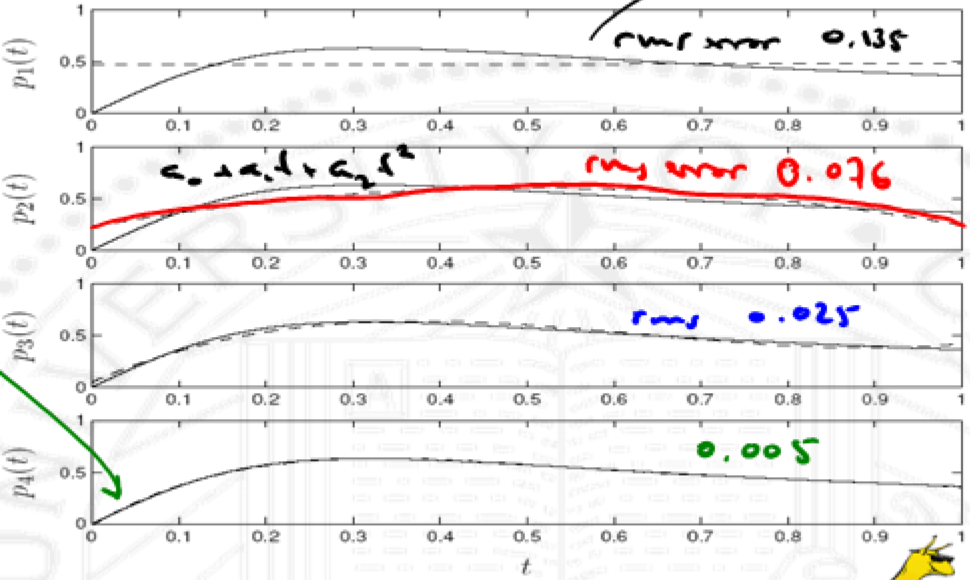
rms error 0.135

$a_0 + a_1 t + a_2 t^2$

rms error 0.076

rms 0.025

0.005



Growing Sets of Regressors (1.2)

consider a family of LS problems

$$\min \left\| \sum_{i=1}^{p} x_i a_i - y \right\|$$

$p = 1 \ldots n$

$a_1 \ldots a_p$ "regressors"

projecting y onto span of $\{a_1 \cdots a_p\}$

approximating y by a linear combination $\{a_1 \ldots a_p\}$

regressing y onto $\{a_1 \ldots a_p\}$

Growing Sets of Regressors (2.2)

As p increases, optimal residual decreases

$$p \uparrow \quad \|x_i a_i - y_i\| \downarrow$$

$A = QR$

$$\hat{a}_p = A(:, 1:p) \backslash y \;\leftarrow \text{MATLAB}.$$

for all $p \le n$

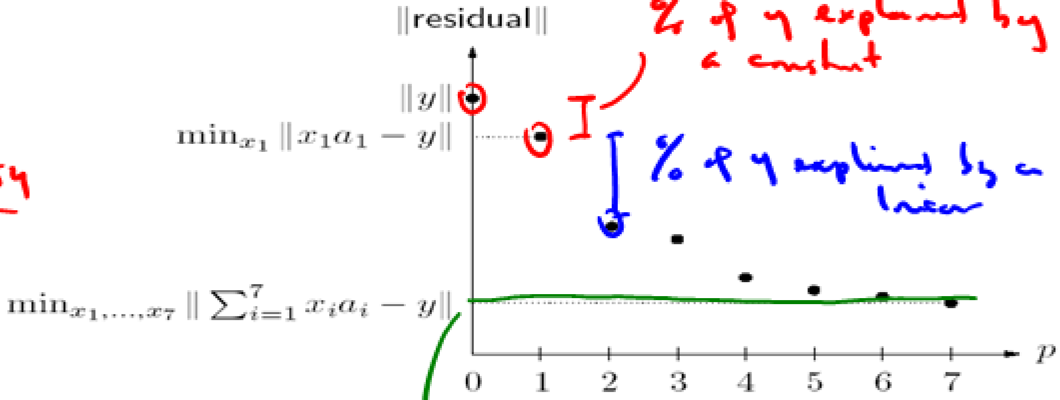
$$x_{ls}^{(p)} = (A_p^T A_p)^{-1} A_p^T y = R_p^{-1} Q_p^T y.$$

R_p is the leading $p \times p$ submatrix of R

Q_p is $[q_1 \cdots q_p]$ first p columns of Q.

Norm of Optimal Residual vs. p

$\|y - A^{(p)} x_{ls}^{(p)}\|$



min floor.

Sensor Integrity Monitor

A B such that $By = \phi$ if y is consistent

$y = Ax$.

$B \in N(A)$

Least squares via QR factorization

$A = QR$

$A x_{ls} = \underbrace{QR}_{A} \underbrace{R^{-1} Q^T}_{A^{+}} y = QQ^T y$ ← NOT IDENTITY

Projection onto $\mathcal{R}(A)$

$\|y - A x_{ls}\| = \|(I - QQ^T) y\|$

$= \sqrt{\|y\|^2 - \sum_{i=1}^{n} (q_i^T y)^2}$

$QQ^T y = \sum_{i=1}^{n} q_i \underbrace{(q_i^T y)}$ → scalar, projection of y onto q_i

reconstructs y.

$$\|(I - QQ^T)y\|^2 = \|y\|^2 - 2y^TQQ^Ty + \|QQ^Ty\|^2$$

$$y^TQ\underbrace{Q^TQ}Q^Ty$$

$$y^TQQ^Ty$$

$$\|Q^Ty\|^2$$

Full QR Factorization

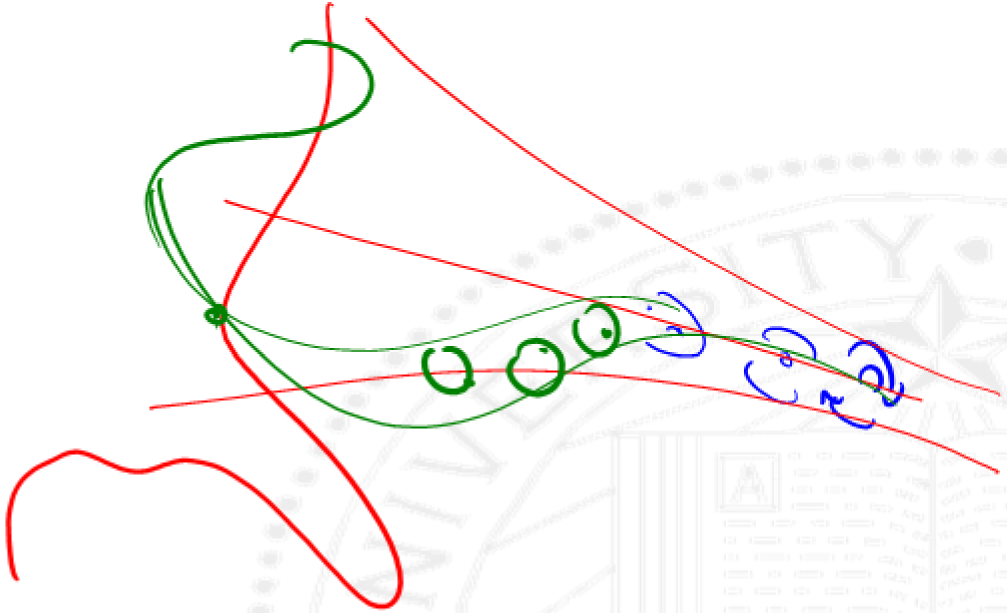
$$[A \,\vdots\, I] = QR = [Q_1 \,\vdots\, Q_2]\begin{bmatrix} R_1 \\ 0 \end{bmatrix}$$

$$\|y - Ax_{ls}\|^2 = \|Q_2Q_2^Ty\|^2 = \|Q_2^Ty\|^2$$

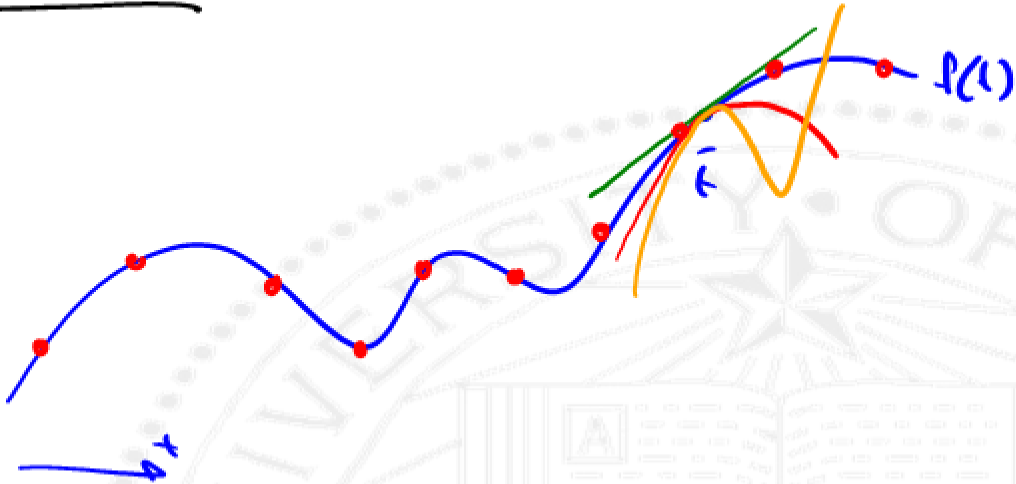
$$\|Q_2^Ty\| = \sqrt{\sum_{i=1}^{n-m}(\tilde{q}_i^Ty)^2} \qquad Q_2 = [\tilde{q}_1, \ldots \tilde{q}_{n-m}]$$

$Q_1 Q_1^T y$ — part of y that I hit with Ax_{ls}.

$Q_2 Q_2^T y$ — part of y that is orthogonal to $R(A)$.

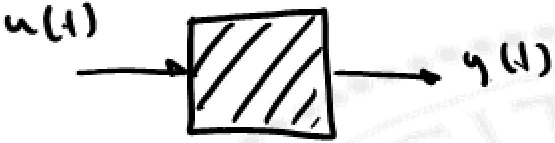


PARTICLE FILTER



$$f(x) \simeq f(\bar{x}) + \left.\frac{\partial f}{\partial x}\right|_{\bar{x}} (x-\bar{x}) + \left.\frac{\partial^2 f}{\partial x^2}\right|_{\bar{x}} (x-\bar{x})^2 \cdots$$

LS System Identification (1.3)



given (u, y) find a model based on the I/O data

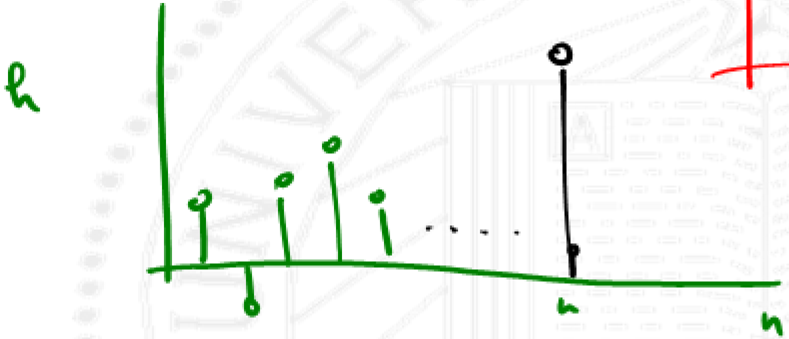
FIR filter — moving average w/ delays

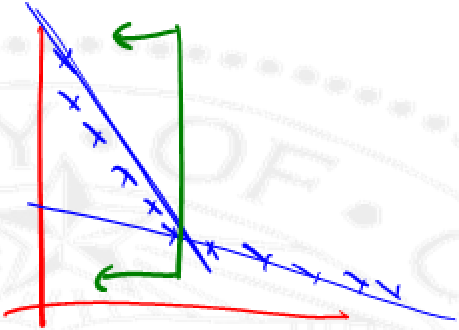
$$\hat{y}(t) = h_0 u(t) + h_1 u(t-1) + \dots + h_n u(t-n)$$

$$\{h_0 \dots h_n\} \in \mathbb{R}$$

LS System Identification (2.3)

$$\hat{h} = (U^T U)^{-1} U^T y.$$





LS System Identification (3.3)

$$\begin{bmatrix}\hat{y}_n\\ \vdots \\ \hat{y}_N\end{bmatrix} = \underbrace{\overbrace{\begin{bmatrix} u_n & u_{n-1} & \cdots & u_0\\ u_{n+1} & u_n & u_{n-1} & \cdots \\ & \ddots & \ddots & \ddots \\ u_N & & & \end{bmatrix}}^{U}}_{\text{TOEPLITZ MATRIX}} \overbrace{\begin{bmatrix}h_0\\ \vdots\\ h_n\end{bmatrix}}^{h}$$

$$e = \begin{bmatrix} y_n - \hat{y}_n \\ \vdots \\ y_N - \hat{y}_N \end{bmatrix}$$

$$\min_{h} \|e\|$$

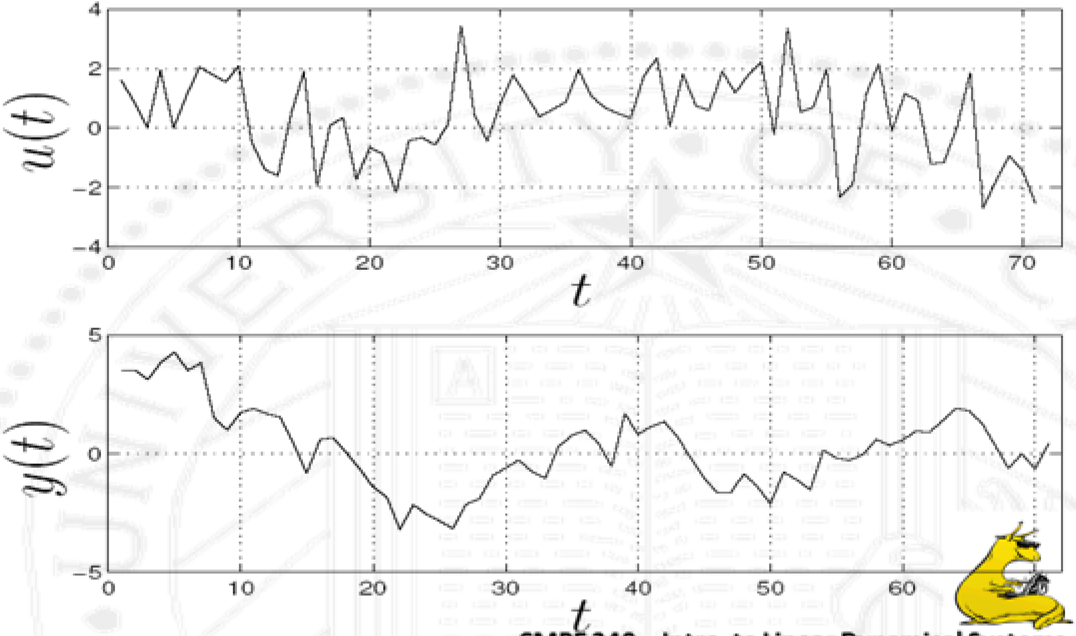
LS SysId Example (1.2)

NOT GENERATED
BY FIR PROCESS

$$\frac{\|e\|}{\|y\|} \approx 0.37$$

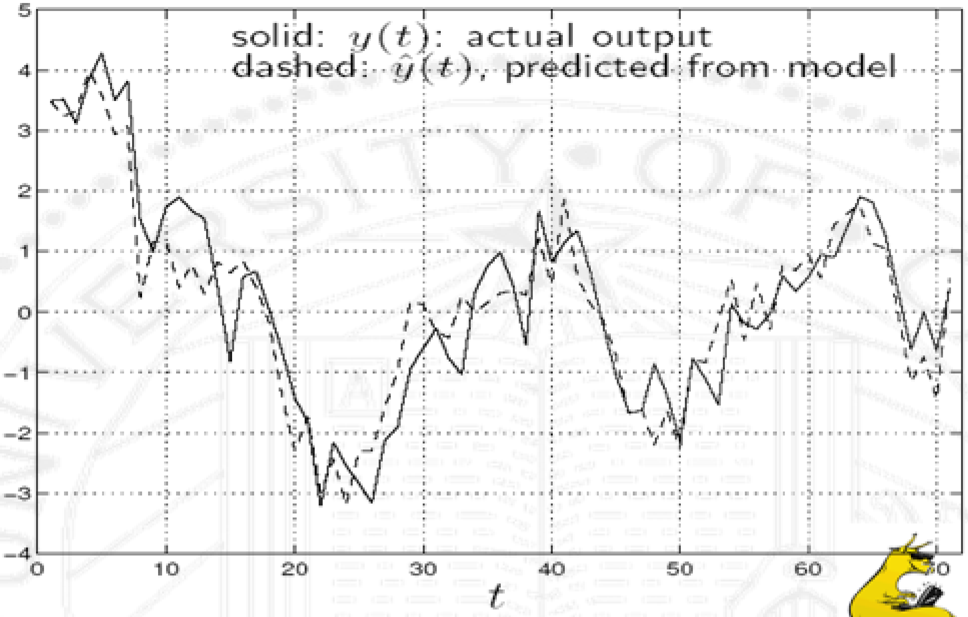
FIR 7 taps

$n=7$.

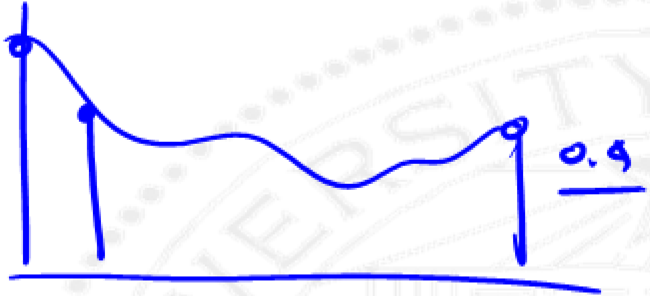


$$h = \begin{bmatrix} 0.029 \\ 0.2519 \\ 0.4326 \\ \vdots \\ .4635 \\ .2084 \\ .4412 \end{bmatrix}$$

LS SysId Example (2.2)

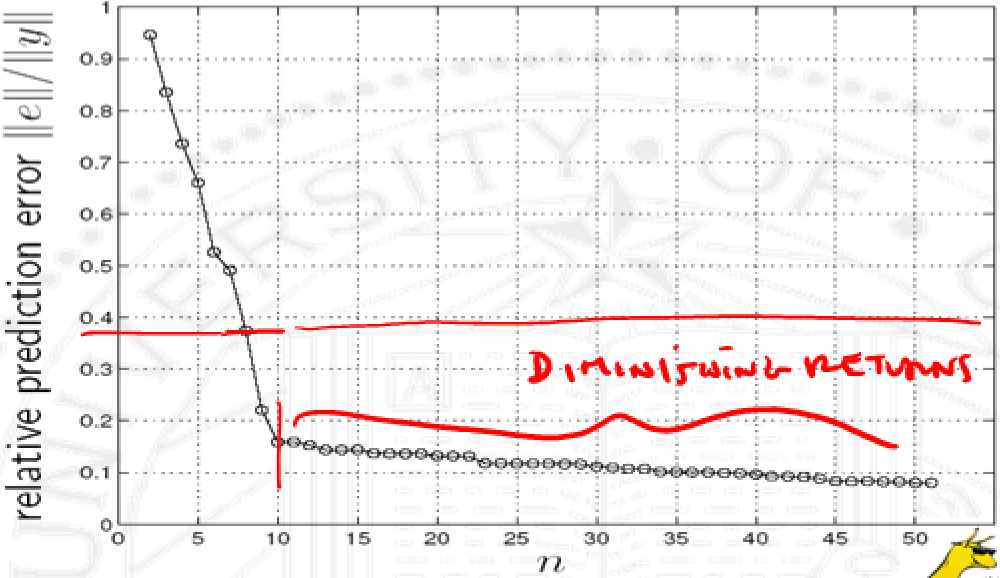


Model Order Selection (1.2)



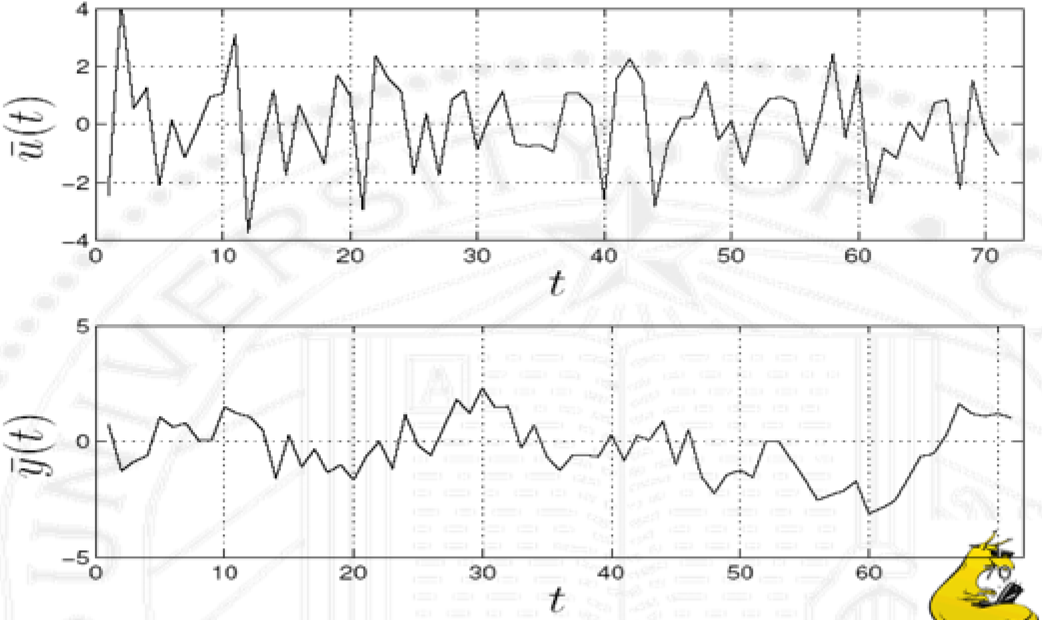
Bigger n better on data to form the model.

Model Order Selection (2.2)

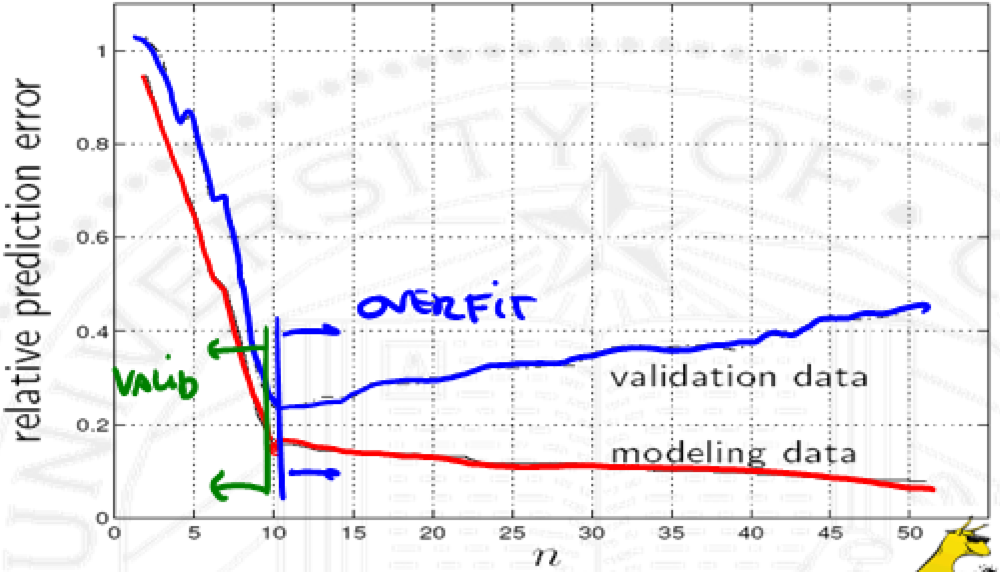


$\|e\| = \|\bar{y} - Ax^{(p)}_{l,s}\|$

Cross-Validation (1.3)



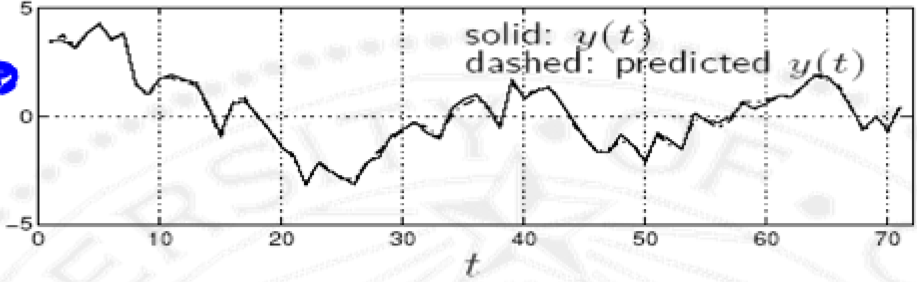
Cross-Validation (2.3)



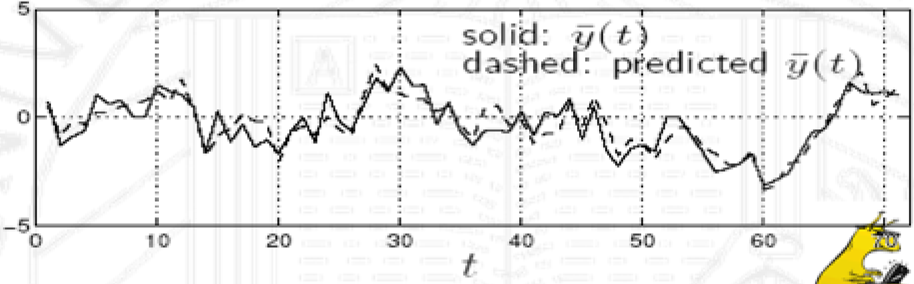
Cross-Validation (3.3)

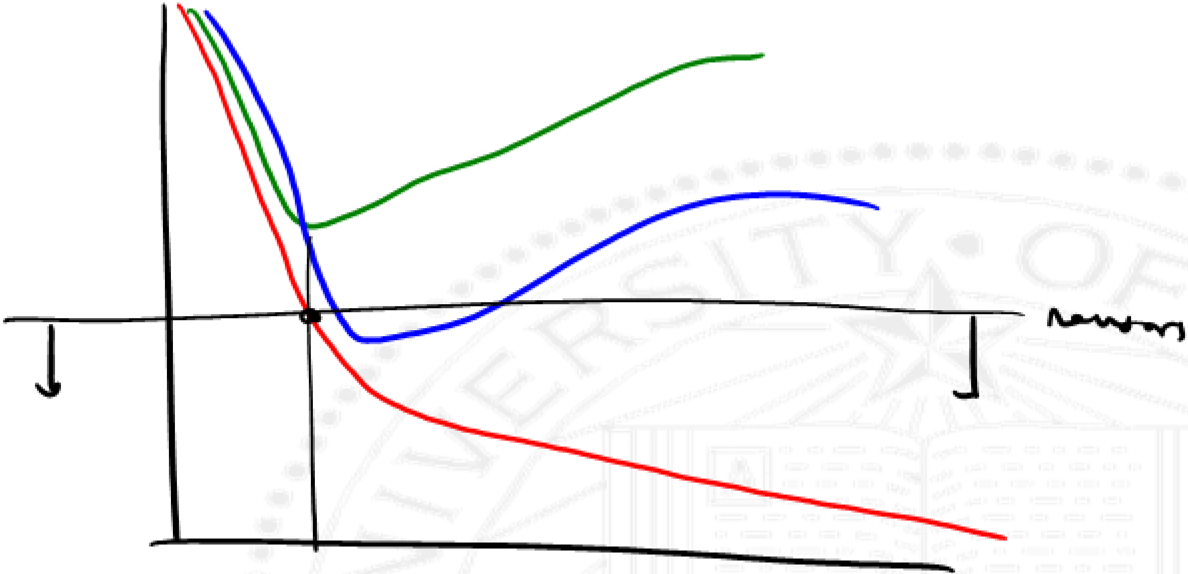
u=10

TRAINING SET



VALIDATION SET





Growing Sets of Measurements

$$\min_x \|Ax - y\|^2 = \sum_{i=1}^{m} (a_i^T x - y_i)^2$$

a_i^T are the rows of A $(a_i \in \mathbb{R}^n)$

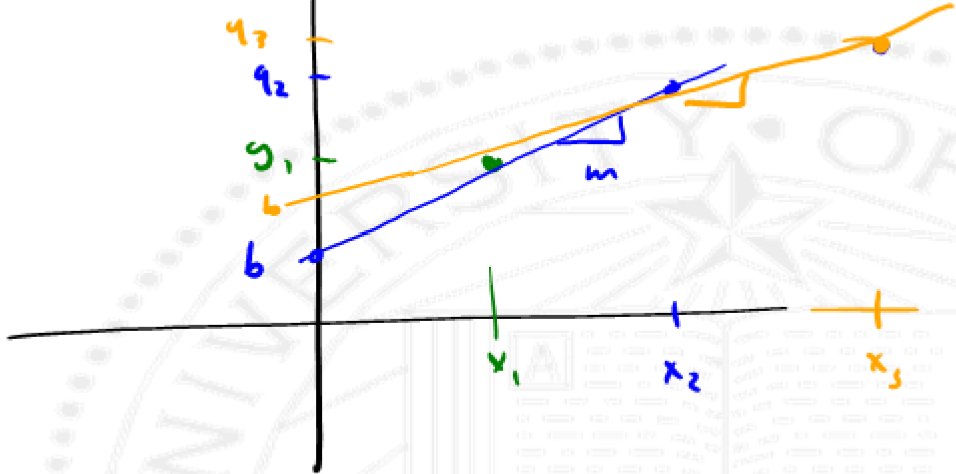
$x \in \mathbb{R}^n$ is a vector to be estimated.

each pair (a_i, y_i) corresponds to ONE additional measurement

Solution $$x_{ls} = \left(\sum_{i=1}^{m} a_i a_i^T\right)^{-1} \sum_{i=1}^{m} y_i a_i^T$$

$$y = mx + b \rightarrow y = \begin{bmatrix} x & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix}$$

$\hat{m}, \hat{b}$



Recursive Least-Squares

Compute $x_{ls}(m) = \left(\sum_{i=1}^{m} a_i a_i^T \right)^{-1} \sum_{i=1}^{m} y_i a_i^T$

$P(m)^{-1} \sim O(n^3)$.

$P(0) = 0 \in \mathbb{R}^{n \times n}$

$q(0) = 0 \in \mathbb{R}^n$

for $m = 0, 1, \ldots$

$P(m+1) = P(m) + a_{m+1} a_{m+1}^T$

$q(m+1) = q(m) + y_{m+1} a_{m+1}^T$

if $P(m)^{-1}$ exists $\longrightarrow$ $x_{ls}(m) = P^{-1}(m)\, q(m)$

$\{a_1 \ldots a_m\}$ spans $\mathbb{R}^n$.

Fast Update for *RLS* $\sim o(n^2)$

$$P(m+1)^{-1} = \left(P(m) + a_{m+1} a_{m+1}^T\right)^{-1}$$

$$(P + aa^T)^{-1} = P^{-1} - \frac{1}{1 + a^T P^{-1} a}(P^{-1}a)(P^{-1}a)^T$$

$P = P^T$ $\quad P \leftarrow (P + aa^T)$ on mult

Verification of Rank One Update

$$(P + aa^T)\left(P^{-1} - \frac{1}{1 + a^T P^{-1} a}(P^{-1}a)(P^{-1}a)^T\right)$$

$$I + aa^T P^{-1} - \frac{1}{1 + a^T P^{-1} a} P(P^{-1}a)(P^{-1}a)^T - \frac{aa^T (P^{-1}a)(P^{-1}a)^T}{1 + a^T P^{-1} a}$$

$$I + aa^T P^{-1} - \frac{aa^T P^{-1}}{1 + a^T P^{-1} a} - \frac{a^T P^{-1} a}{1 + a^T P^{-1} a} \cdot aa^T P^{-1}$$

$$I + aa^T P^{-1} - \underbrace{\left[\frac{1 + a^T P^{-1} a}{1 + a^T P^{-1} a}\right]}_{1} aa^T P^{-1} = I.$$

Questions?

Rank one down date