Least-Squares Applications

Gabriel Hugh Elkaim

Winter 2016

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)

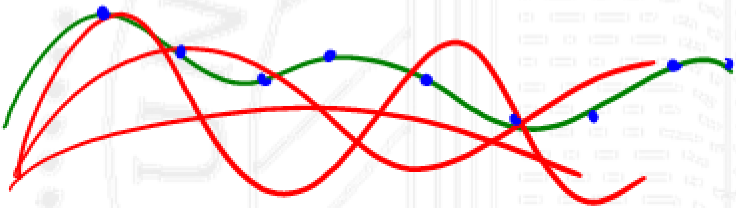
given a set of functions $\{f_1 \dots f_n\} : S \rightarrow \mathbb{R}$ regressors / basis functions

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

$s_i \in S$ (usually) $m \gg n$

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

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



Least-Squares Data Fitting (2.3)

choose x to min mean square fitting error

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

$$\min_x \|Ax-g\|^2$$

$A_{ij} = f_j(s_i)$ function "j" evaluated at point s_i

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

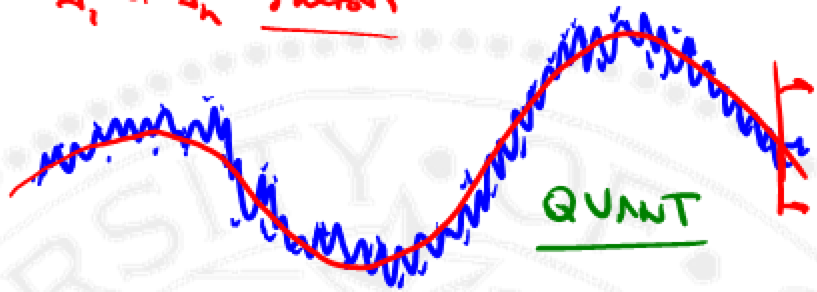
approximating f $f_{ls_{fit}}(s) = x_1 f_1(s)+\dots+x_n f_n(s)$

LTCM

Least-Squares Data Fitting (3.3)

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

$f_1 \cdots f_n$ smooth



Celestial Orbits ← Gauss least squares fitting

LS Polynomial Fitting (1.2)

polynomial of order n will hit n points exactly

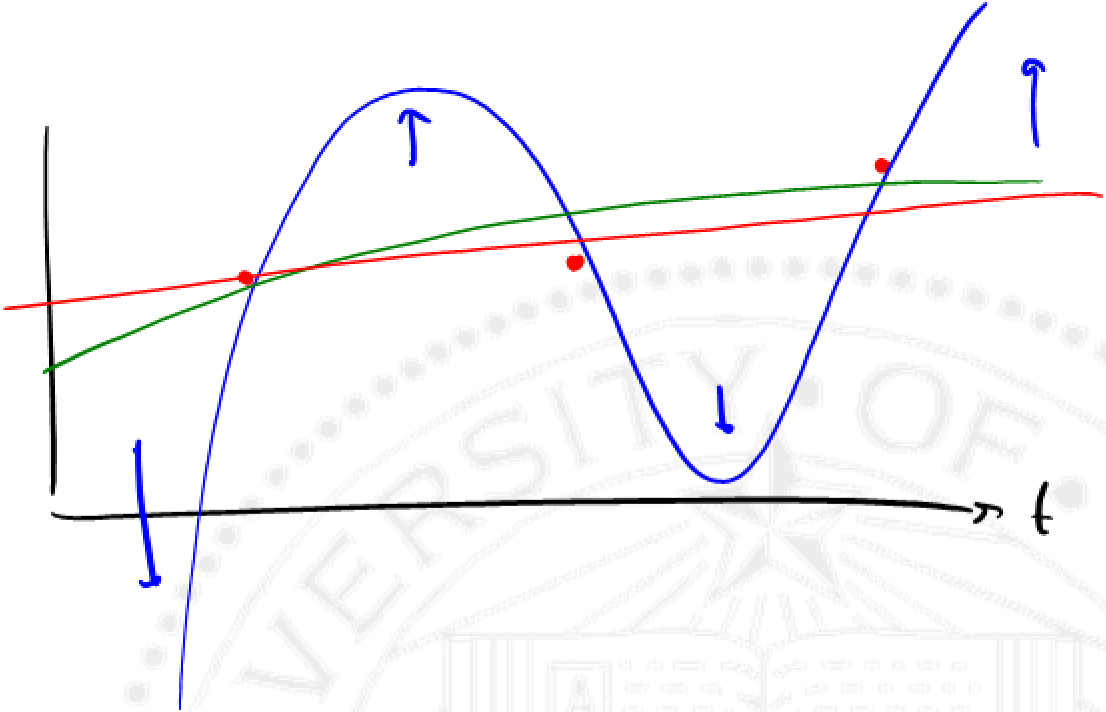
fit polynomial of order $< n$

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

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

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

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


t

LS Polynomial Fitting (2.2)

$$A = \underbrace{\begin{bmatrix} 1 & t_1 & t_1^2 & \cdots & t_1^{n-1} \\ \vdots & t_2 & t_2^2 & \cdots & t_2^{n-1} \\ \vdots & \vdots & & & \\ 1 & t_m & t_m^2 & \cdots & t_m^{n-1} \end{bmatrix}}_{\text{Vandermode Matrix}}$$

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

$m >> n$

$Aa = 0 \quad p(t) = a_0 + \dots + a_{n-1} t^{n-1}$

vanishes at $t_1 \dots t_m$.

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

p no more than $n-1$ zeros

p is identically 0 only when $a = 0$

$g(t) = 4t/(1+10t^2)$

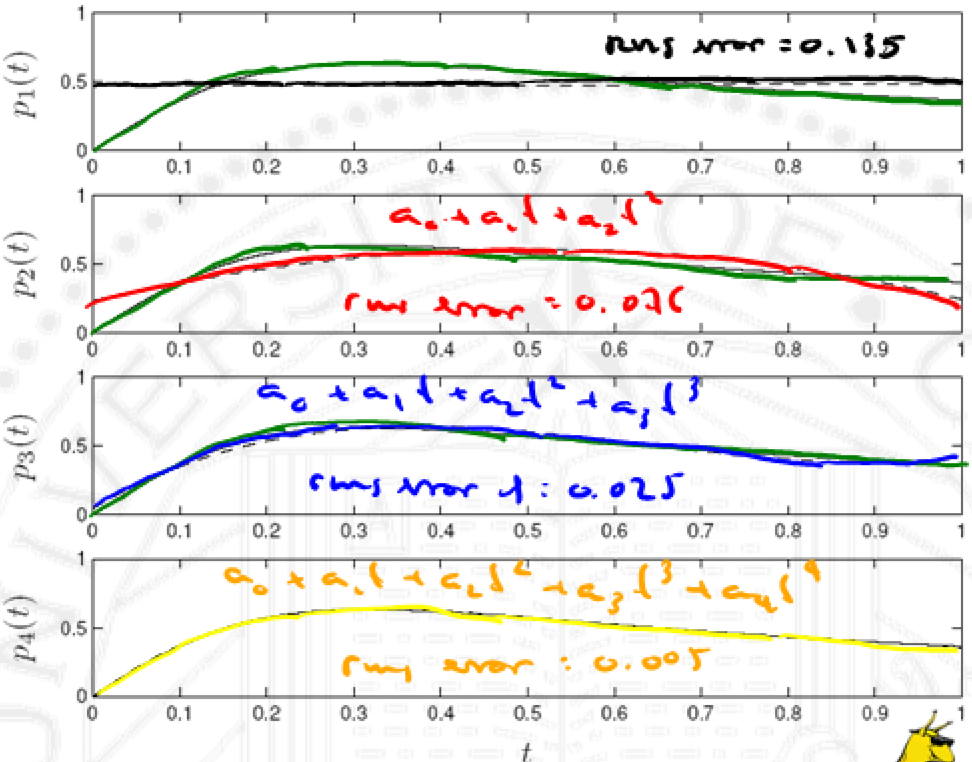
LS Polynomial Example

$a_0 + a_1 t$

$m = 100$

$t = 0.01$

p_1



Growing Sets of Regressors (1.2)

consider a family of least squares problems.

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

$a_1 \ldots a_p$ regressors.

project y onto $\operatorname{span}\{a_1 \ldots a_p\}$

approximate y by a linear combination of $a_1 \ldots a_p$

regress y onto $a_1 \ldots a_p$

As p increases, so the optimal residual decreases

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

Growing Sets of Regressors (2.2)

$\hat{a}_p = $ A(:, 1:p)\y. ← matlab

$A = QR$

for all $p \leq n$

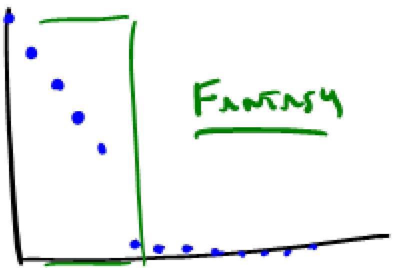
$$x_{ls}^{(p)} = (A_p^T A_p)^{-1} A_p^T y = \underline{R_p^{-1} Q_p^T y}$$

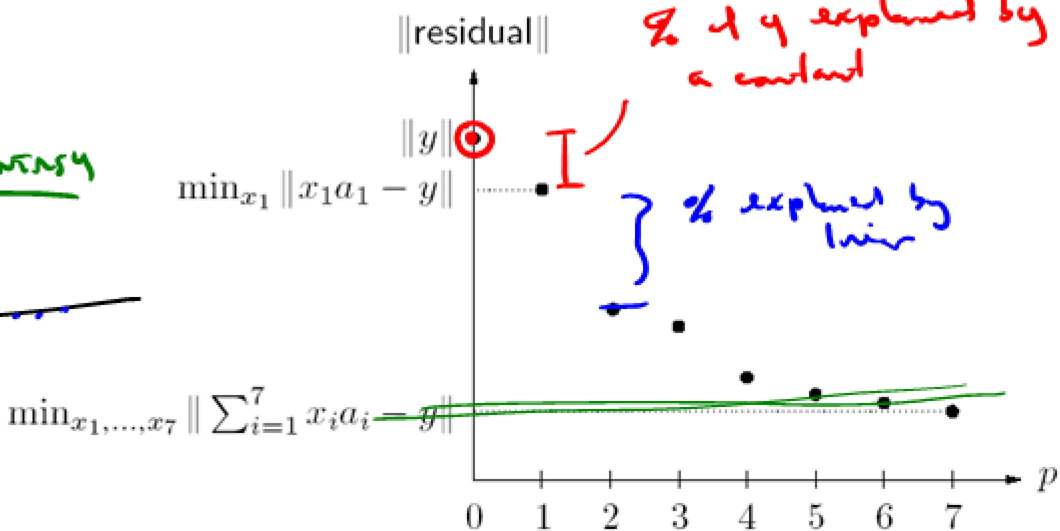
R_p is the leading $p \times p$ submatrix of R

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

Norm of Optimal Residual vs. *p*

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





Least Squares via QR Factorization

$A = QR$

$$A x_{ls} = \underbrace{QR}_{A}\,\underbrace{R^{-1}Q^T}_{A^\dagger}\, y = QQ^T y$$

not identity

proj $R(A)$

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

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

$$QQ^T y = \sum_{i=1}^{n} q_i \,(\underbrace{q_i^T y}_{\text{projects onto basis } q})$$

reconstructs it

$$y^T Q \boxed{Q^T Q} Q^T y = y^T Q Q^T y$$

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

Full QR Factorization

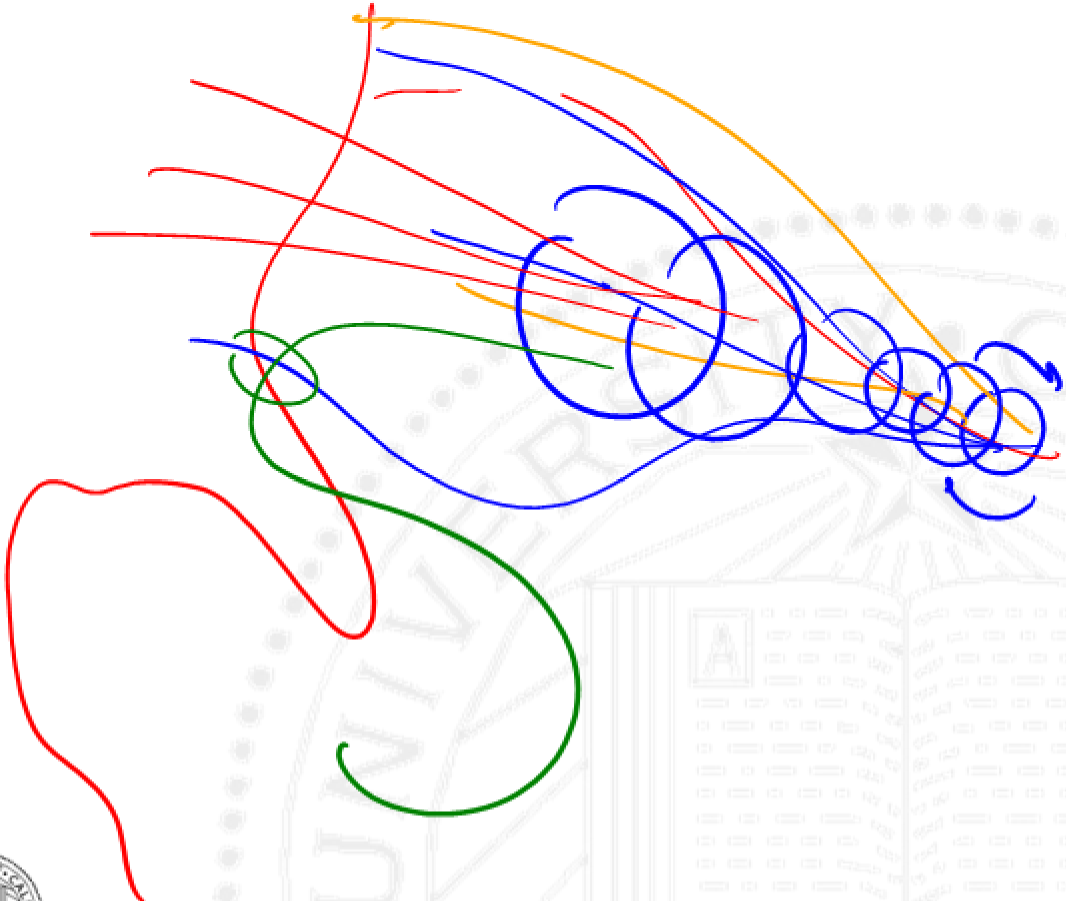
$$[A \; \vdots \; I] = QR = [Q_1 \;\; Q_2]\begin{bmatrix} R_1 \\ 0 \end{bmatrix}$$

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

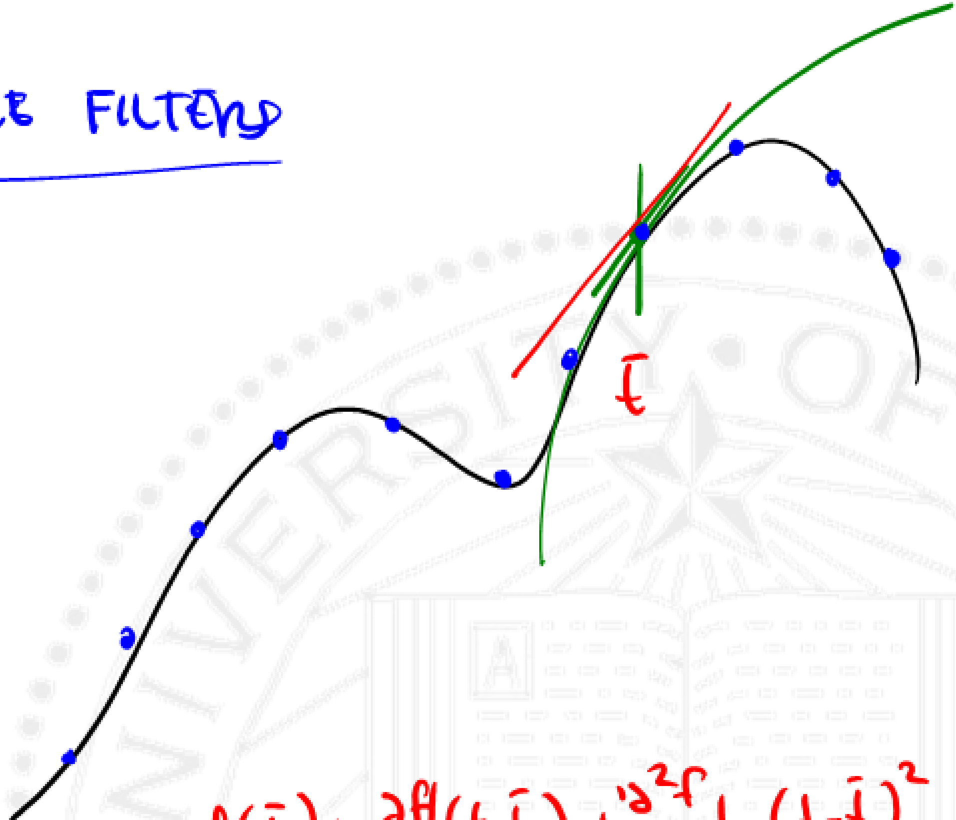
$$\|Q_2^T y\| = \sqrt{\sum_{i=1}^{m-n} (\tilde{q}_i^T y)^2} \qquad Q_2 = [\tilde{q}_1 \cdots \tilde{q}_{m-n}]$$

$Q_1 Q_1^T y$ – part of y that I can hit from Ax_{ls}.

$Q_2 Q_2^T y$ – part of y orthogonal to $R(A)$.

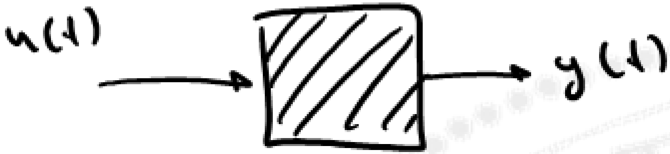


PARTICLE FILTERS



$$f(t) \simeq f(\bar{t}) + \left.\frac{\partial f}{\partial t}\right|_{\bar{t}} (t-\bar{t}) + \frac{1}{2}\left.\frac{\partial^2 f}{\partial t^2}\right|_{t=\bar{t}} (t-\bar{t})^2$$

LS System Identification (1.3)



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

FIR filter → Moving Average w/ delays

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

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

LS System Identification (2.3)

$$\begin{bmatrix} \hat{y}_n \\ \vdots \\ \hat{y}_N \end{bmatrix} = \underbrace{\begin{bmatrix} u_n & u_{n-1} & \cdots & u_0 \\ u_{n+1} & \ddots & \ddots & \\ \vdots & \ddots & \ddots & \\ u_N & & & \end{bmatrix}}_{\text{Toeplitz Matrix}} \begin{bmatrix} h_0 \\ \vdots \\ h_n \end{bmatrix}$$

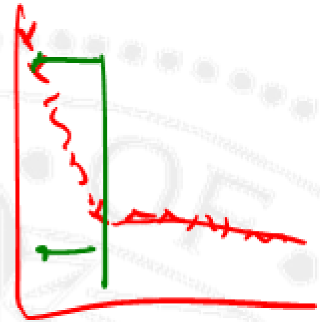
$$e = \begin{bmatrix} y_n - \hat{y}_n \\ \vdots \\ y_N - \hat{y}_N \end{bmatrix}$$

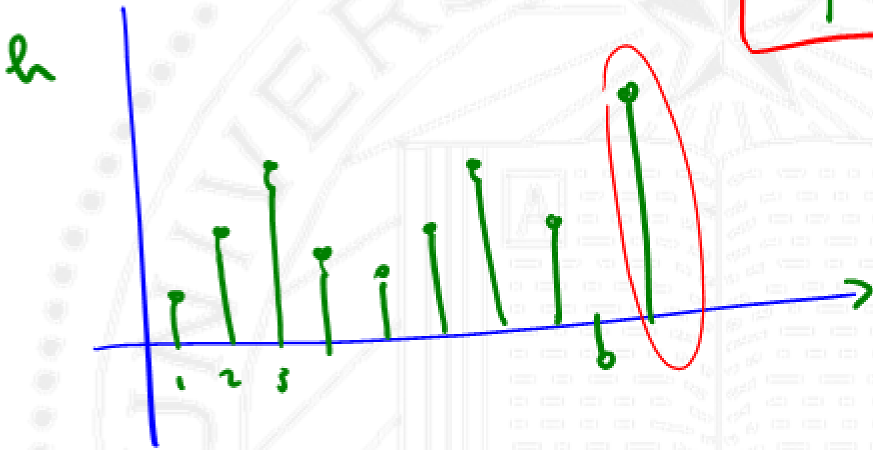
choose model (h)

$$\min_h \|e\|$$

LS System Identification (3.3)

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



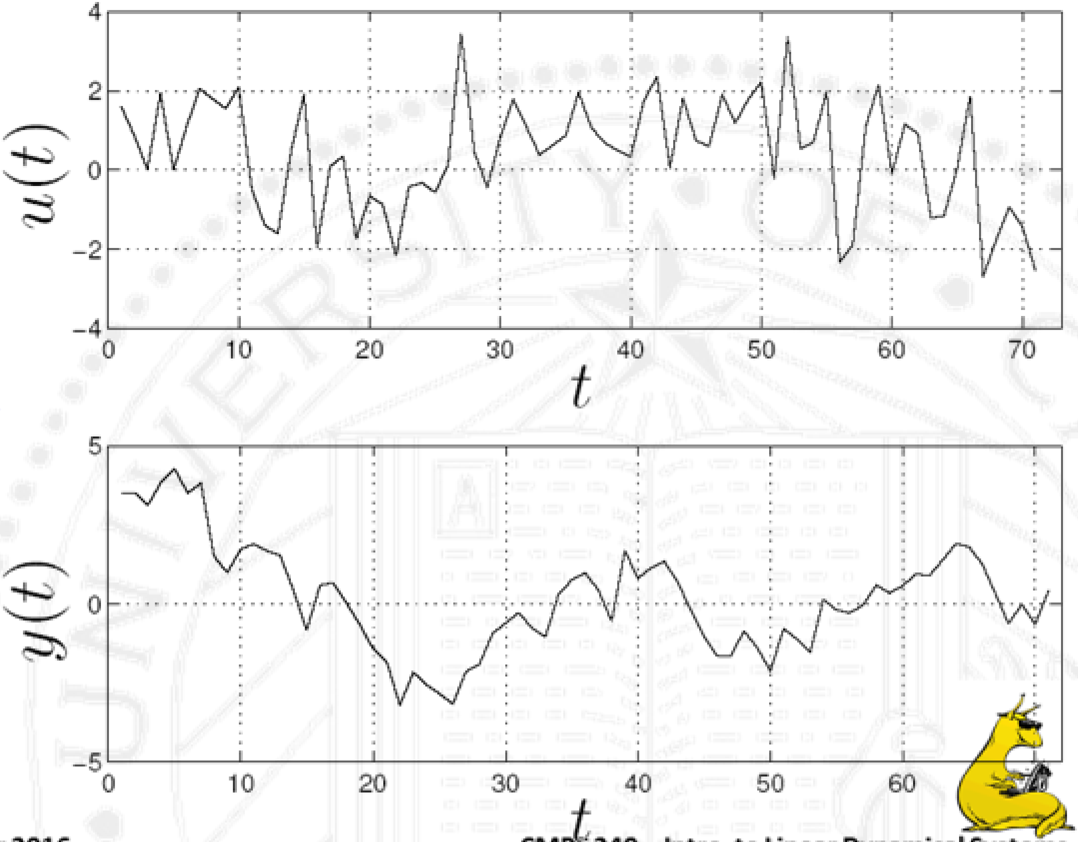


LS SysId Example (1.2)

NOT ESTIMATED BY FIR

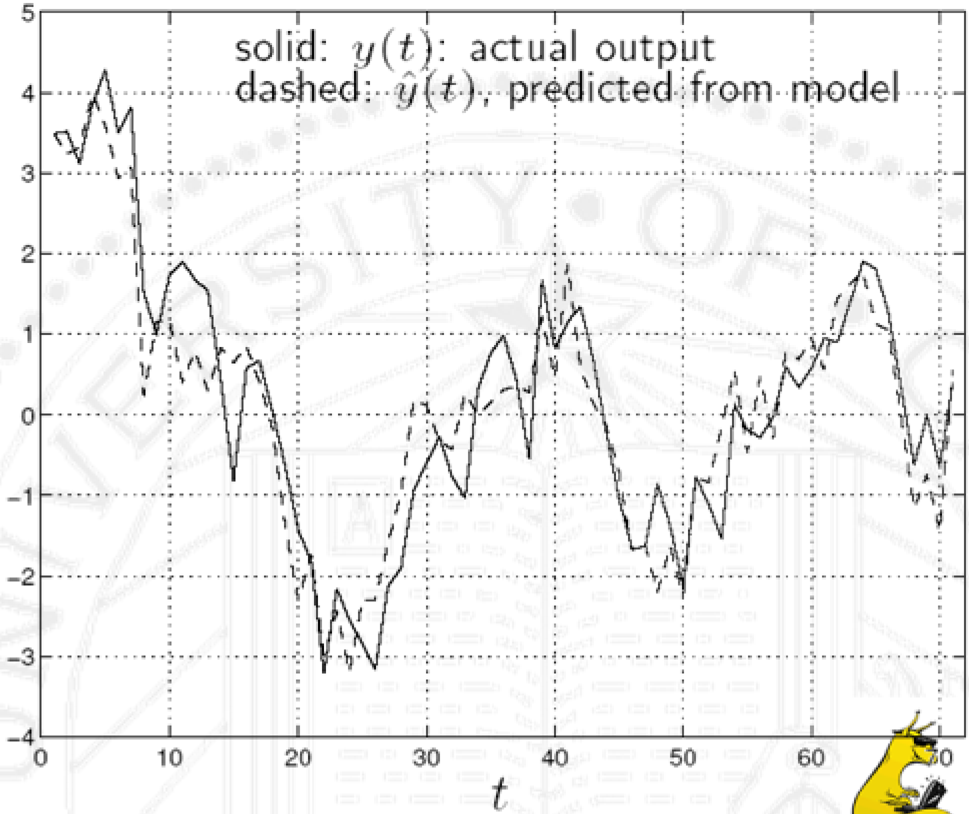
$\frac{\|e\|}{\|y\|} \approx 0.37$

$n=7$

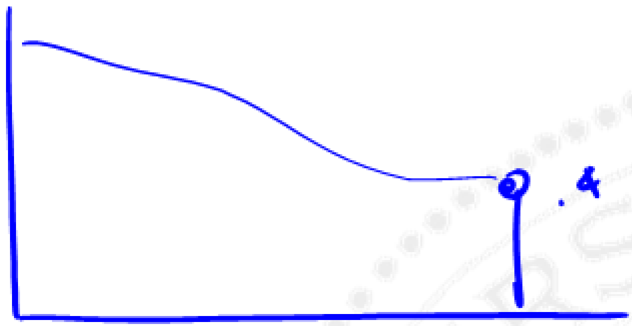


LS SysId Example (2.2)

$$\begin{bmatrix} 0.0240 \\ 0.2819 \\ 0.4176 \\ 0.3536 \\ 0.2425 \\ 0.4873 \\ 0.2084 \\ 0.4412 \end{bmatrix}$$

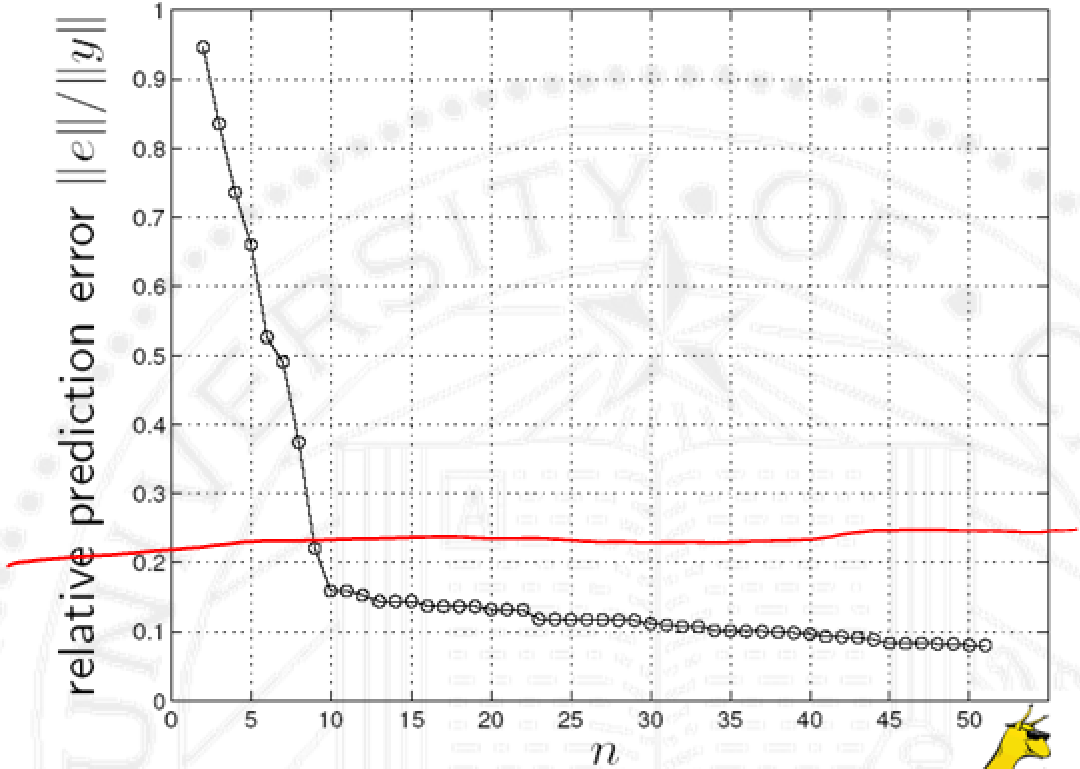


Model Order Selection (1.2)



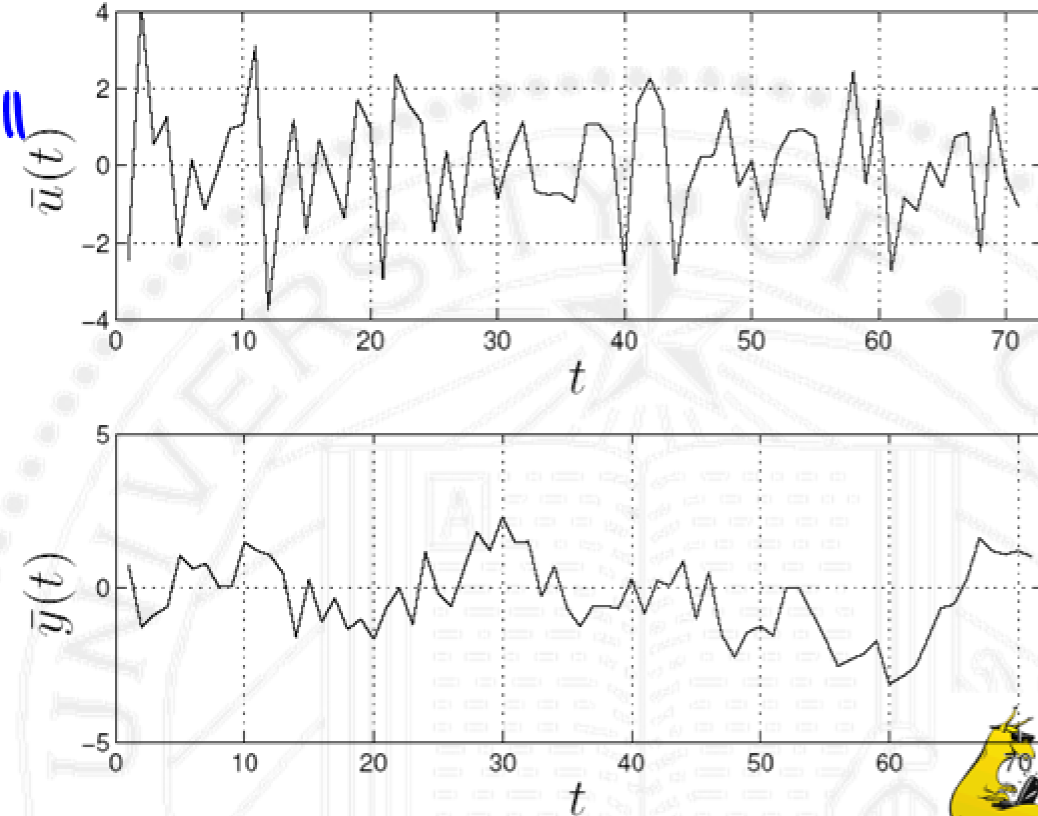
Bigger n better on data used to form the model

Model Order Selection (2.2)

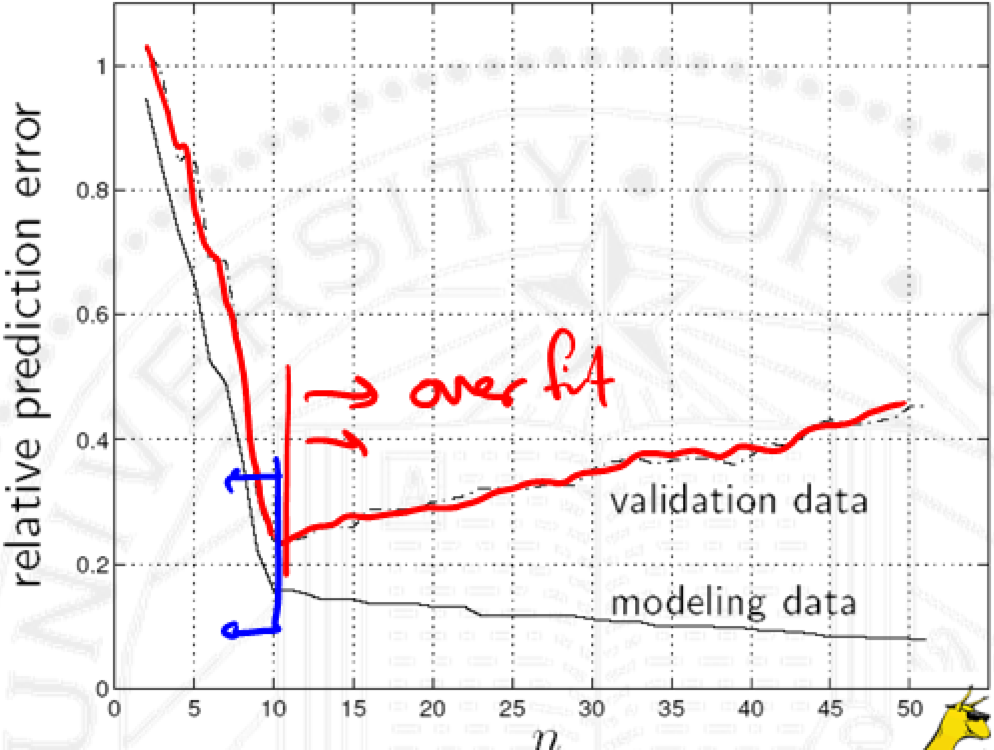


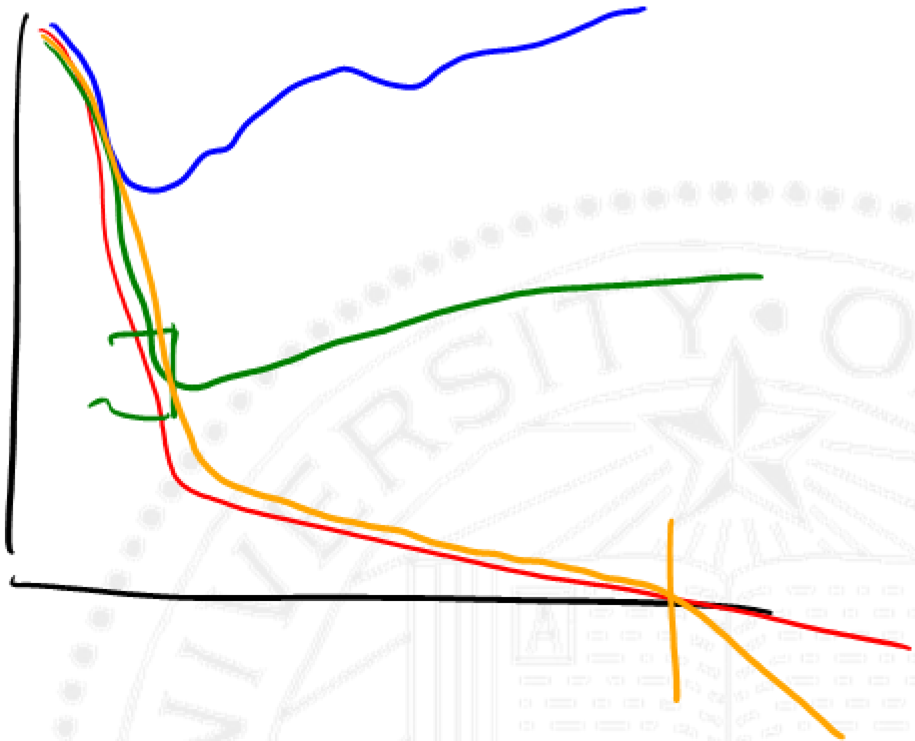
Cross-Validation (1.3)

$\|\bar{e}\| = \|\bar{y} - A x_1^{(4)}\|$



Cross-Validation (2.3)





//

Cross-Validation (3.3)

$n=10$

Training set

Validation set