Recursive Least Squares Parameter Estimation for Linear Steady State and Dynamic Models

Thomas F. Edgar
Department of Chemical Engineering
University of Texas
Austin, TX 78712

Outline

- Static model, sequential estimation
- Multivariate sequential estimation
- Example
- Dynamic discrete-time model
- Closed-loop estimation

Least Squares Parameter Estimation

Linear Time Series Models

ref: PC Young, Control Engr., p. 119, Oct, 1969

scalar example (no dynamics)

model $y = ax$

data $y^* = ax + \in$ $\in$: error

least squares estimate of a: $(\hat{a})$

$$\min_{\hat{a}} \sum_{i=1}^{k} \left(\hat{a} x_i - y_i^* \right)^2 \quad (1)$$

Simple Example

The analytical solution for the minimum (least squares) estimate is

$$\hat{a}_k = \underbrace{\left(\sum_{i-1}^{k} x_i^{\,2} \right)^{-1}}_{p_k} \underbrace{\left(\sum_{i-1}^{k} x_i y_i^{\,*} \right)}_{b_k} \qquad (2)$$

p_k, b_k are functions of the number of samples

This is the non-sequential form or non-recursive form

Sequential or Recursive Form

To update $\hat{a}_k$ based on a new data pt. (y_i , x_i), in Eq. (2) let

$$p_k^{-1} = \sum_{i=1}^{k} x_i^{\,2} = p_{k-1}^{\;-1} + x_k^{\,2} \qquad (3)$$

and

$$b_k = \sum_{i=1}^{k} x_i y_i^{\,*} = b_{k-1} + x_k y_k^{\,*} \qquad (4)$$

Recursive Form for Parameter Estimation

$$\hat{a}_k = \hat{a}_{k-1} - K_k \underbrace{\left(x_k \hat{a}_{k-1} - y_k^{\ *}\right)}_{\text{estimation error}} \qquad (5)$$

where

$$K_k = p_{k-1} x_k \left(1 + p_{k-1} x_k^{\ 2}\right)^{-1} \qquad (6)$$

To start the algorithm, need initial estimates $\hat{a}_0$ and p_0. To update p,

$$p_k = p_{k-1} - p_{k-1}^{\ 2} x_k^{\ 2} \left(1 + p_{k-1} x_k^{\ 2}\right)^{-1} \qquad (7)$$

(Set p_0 = large positive number)

Eqn. (7) shows p_k is decreasing with k

Estimating Multiple Parameters (Steady State Model)

$$y = \underline{x}^T \underline{a} = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n \qquad (8)$$

$$\min_{\underline{\hat{a}}} \sum_{i=1}^{k} \left(\underline{x}_i^T \underline{\hat{a}} - y_i^* \right)^2$$

(non-sequential solution requires n x n inverse)

To obtain a recursive form for $\underline{\hat{a}}$,

$$\underline{P}_k^{-1} = \underline{P}_{k-1}^{-1} + \underline{x}_k \ \underline{x}_k^T \qquad (9)$$

$$\underline{B}_k = \underline{B}_{k-1} + \underline{x}_k \ y_k^* \qquad (10)$$

Recursive Solution

$$\underline{\hat{a}}_k = \underline{\hat{a}}_{k-1} - \underbrace{\underline{P}_{k-1}\underline{x}_k\left(1+\underline{x}_k{}^T\underline{P}_{k-1}\underline{x}_k\right)^{-1}}_{K_k} \underbrace{\left(\underline{x}_k{}^T\underline{\hat{a}}_{k-1} - y_k{}^*\right)}_{\text{estimation error}} \quad (11)$$

$$\underline{P}_k = \underline{P}_{k-1} - \underline{P}_{k-1}\underline{x}_k\left[1+\underline{x}_k{}^T\underline{P}_{k-1}\underline{x}_k\right]^{-1}\underline{x}_k{}^T\underline{P}_{k-1} \quad (12)$$

need to assume $\underline{\hat{a}}_0$ (vector)

$\underline{P}_0$ (diagonal matrix)

$$P_0 = \begin{bmatrix} P_{11} & & 0 \\ & P_{22} & \\ 0 & & P_{nn} \end{bmatrix}$$

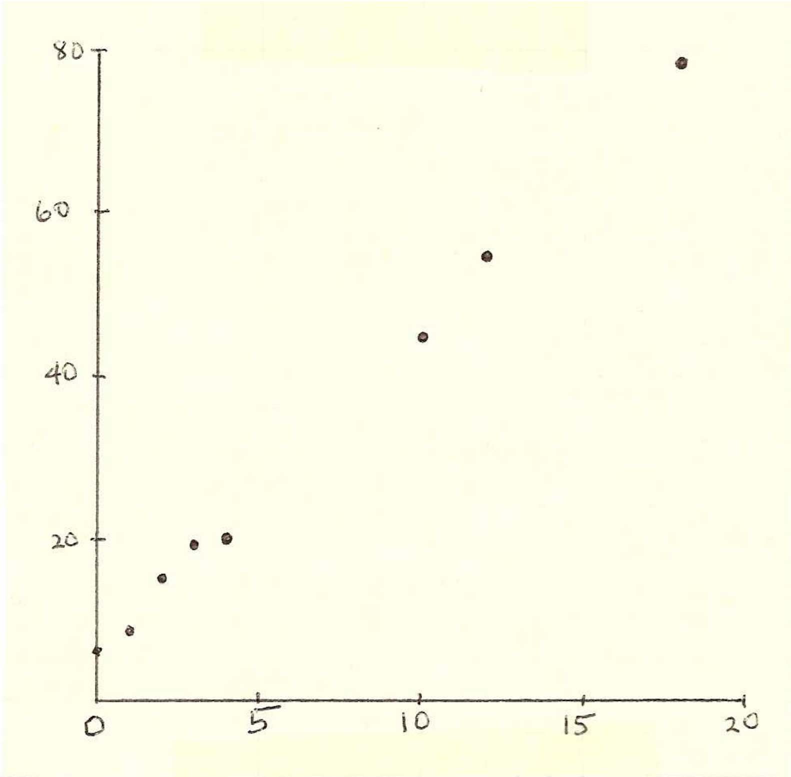
P_{ii} large

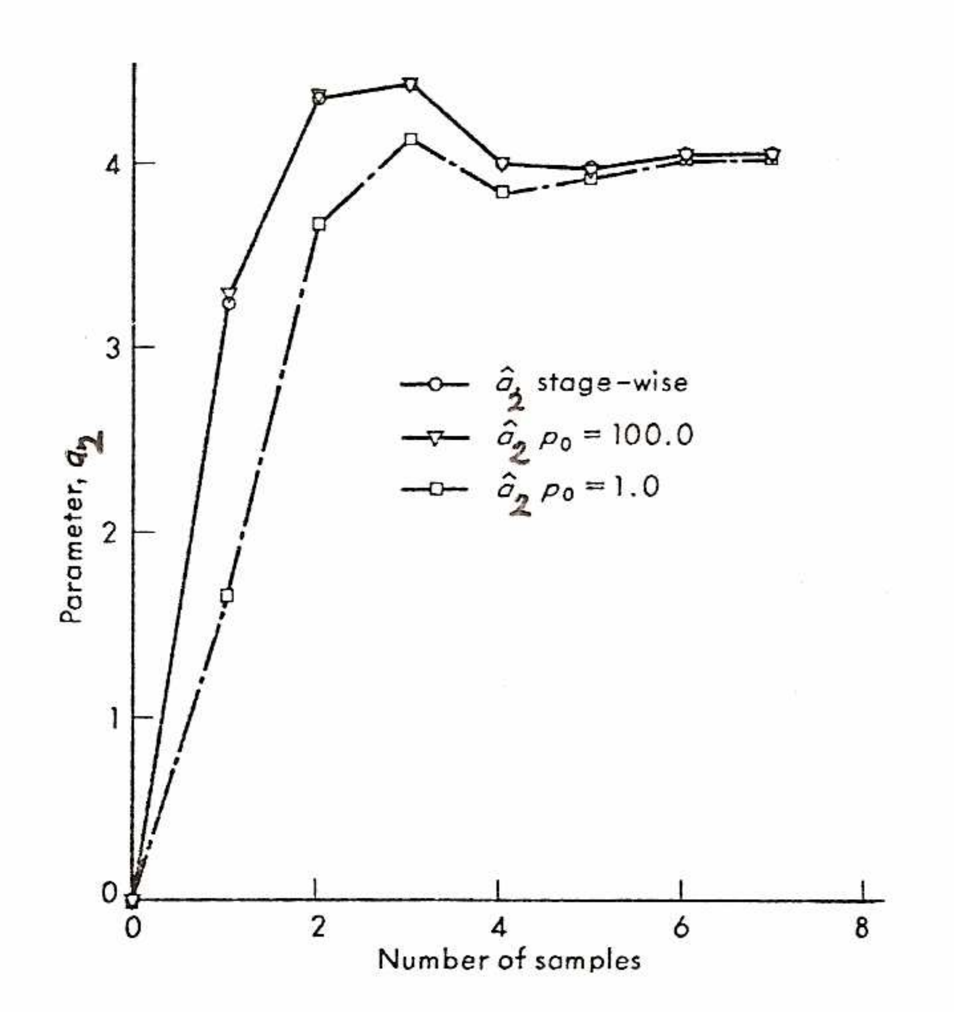
Simple Example (Estimate Slope & Intercept)

Linear parametric model

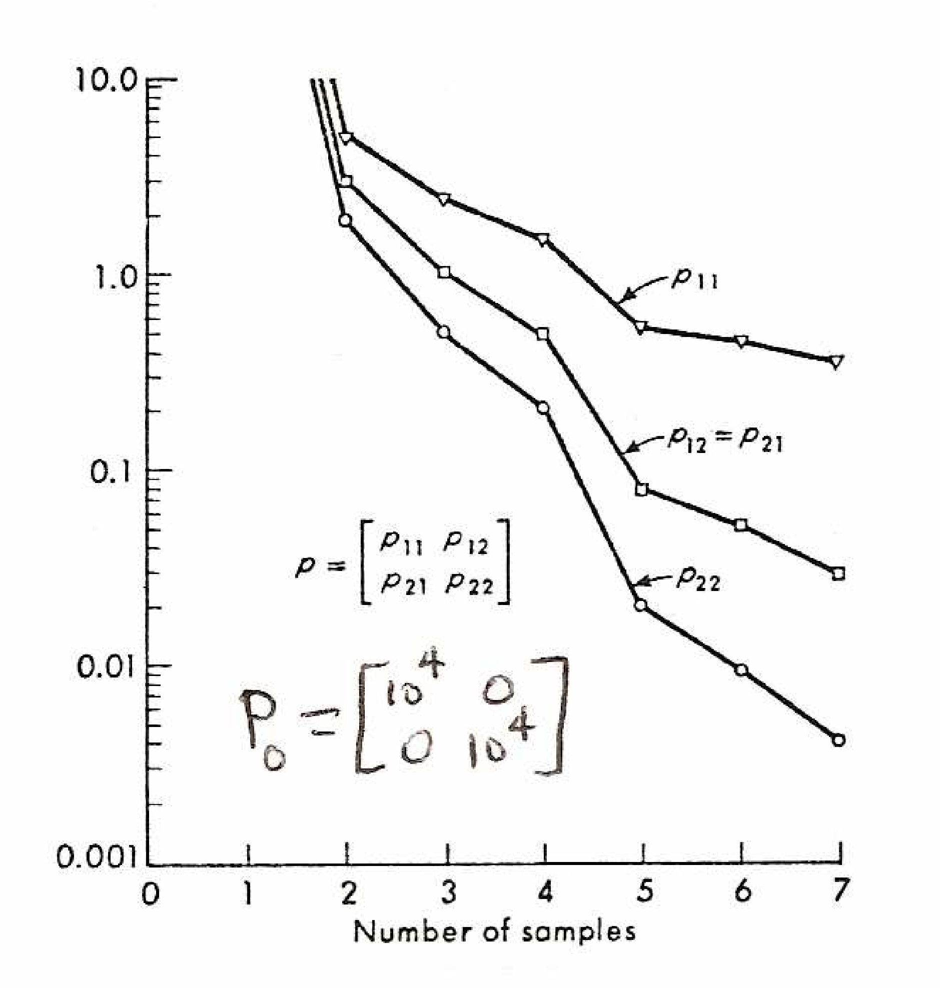
input, *u*	0	1	2	3	4	10	12	18
output, *y*	5.71	9	15	19	20	45	55	78

$$y = a_1 + a_2 u$$

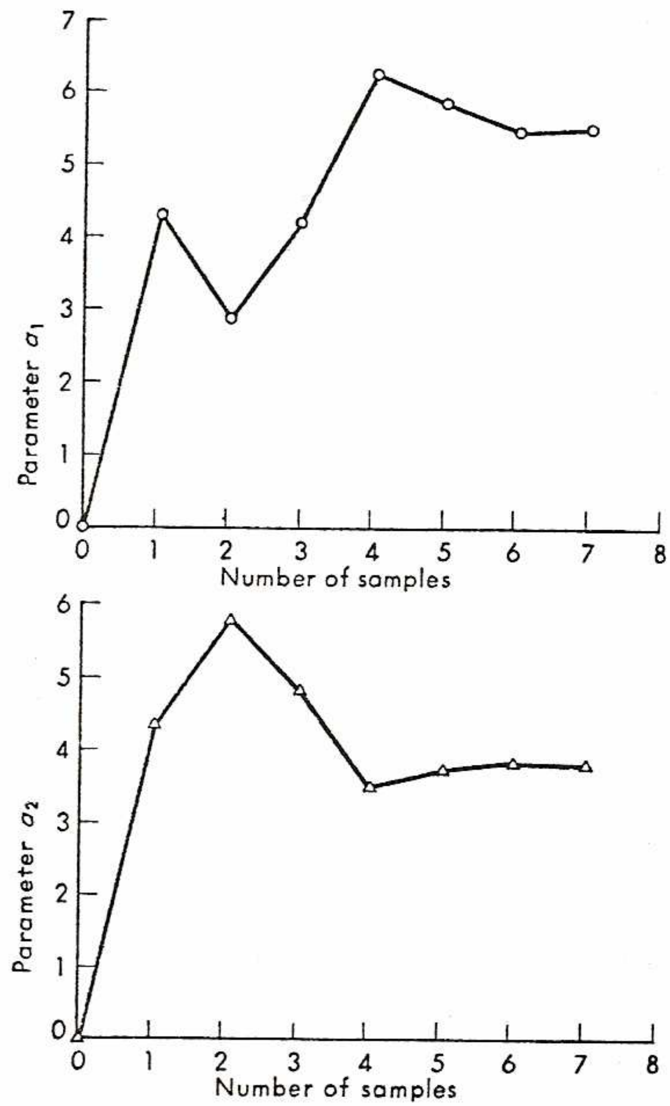




Sequential vs. Non-sequential Estimation of a_2 Only ($a_1 = 0$)



Covariance Matrix vs. Number of Samples Using Eq. (12)



Sequential Estimation of $\hat{a}_1$ and $\hat{a}_2$ Using Eq. (11)

Application to Digital Model and Feedback Control

Linear Discrete Model with Time Delay:

$$y(t) = a_1 y(t-1) + a_2 y(t-2) + \cdots + a_n y(t-n)$$

$$+ b_1 u(t-1-N) + b_2 u(t-2-N) + \cdots + b_r u(t-r-N) + d \quad (13)$$

y: output u: input d: disturbance N: time delay

Recursive Least Squares Solution

$$y(t) = \Psi^T(t-1)\theta(t-1) + \varepsilon(t) \qquad (14)$$

where

$$\Psi^T(t-1) = [y(t-1), y(t-2), \ldots\ \mathrm{y}(t-n),$$
$$u(t-1-N), \ldots\ u(t-r-N),\ 1]$$

$$\theta^T(t-1) = [a_1,\ a_2, \ldots\ a_n,\ b_1,\ b_2,\ \ldots\ b_r,\ d].$$

$$\min_{\hat{\theta}} \sum_{i=1}^{t} \underbrace{\left[\underline{\Psi}^T(i-1)\underline{\theta}(i) - y(i)\right]^2}_{\text{"least squares"}} \qquad (15)$$

"least squares"
(predicted value of y)

$$\hat{\theta}(t) = \hat{\theta}(t-1) + P(t)\Psi(t-1)[y(t) - \Psi^T(t-1)\hat{\theta}(t-1)] \qquad (16)$$

Recursive Least Squares Solution

$$P(t) = P(t-1) - P(t-1)\psi(t-1)[\psi^T(t-1)P(t-1)\Psi(t-1)+1]^{-1}$$

$$\Psi^T(t-1)P(t-1) \qquad (17)$$

$$K(t) = \frac{P(t-1)\Psi(t-1)}{1+\Psi^T(t-1)P(t-1)\Psi(t-1)} \qquad (18)$$

$$P(t) = [I - K(t)\Psi^T(t-1)]P(t-1) \qquad (19)$$

$$\hat{\theta}(t) = \hat{\theta}(t-1) + K(t)[y(t) - \hat{y}(t)] \qquad (20)$$

K: Kalman filter gain

Closed-Loop RLS Estimation

- There are three practical considerations in implementation of parameter estimation algorithms
  - covariance resetting
  - variable forgetting factor
  - use of perturbation signal

Enhance sensitivity of least squares estimation algorithms with forgetting factor λ

$$J(\theta(t)) = \sum_{i=1}^{t} \lambda^{t-i} \left[\psi^T(i-1)\theta(i) - y(i) \right]^2 \qquad (21)$$

$$P(t) = \frac{1}{\lambda}[P(t-1) - P(t-1)\psi(t-1)[\psi T(t-1)P(t-1)\psi(t-1) + \lambda]^{-1}.$$

$$\psi^T(t-1)P(t-1)] \qquad (22)$$

$$\hat{\theta}(t) = \hat{\theta}(t-1) + P(t)\psi(t-1)[y(t) - \psi^T(t-1)\hat{\theta}(t-1)] \qquad (23)$$

λ prevents elements of *P* from becoming too small (improves sensitivity) but noise may lead to incorrect parameter estimates

$0 \leq \lambda \leq 1.0$ $\qquad$ $\lambda \rightarrow 1.0$ all data weighted equally

$\lambda \sim 0.98$ typical

Closed-Loop Estimation (RLS)

Perturbation signal is added to process input (via set point) to excite process dynamics

large signal: good parameter estimates
but large errors in process output

small signal: better control but more sensitivity to noise

Guidelines: Vogel and Edgar, *Comp. Chem. Engr*., Vol. 12, pp. 15-26 (1988)

1. set forgetting factor $\lambda = 1.0$

2. use covariance resetting (add diagonal matrix *D* to *P* when tr (*P*) becomes small)

3. use PRBS perturbation signal only when estimation error is large and *P* is not small. Vary PRBS amplitude with size of elements proportional to tr *(P)*.

4. $P(0) = 10^4 \ I$

5. filter new parameter estimates

$$\theta_c(k) = \rho\theta_c(k-1) + (1-\rho)\theta(k)$$

ρ: tuning parameter
(θ_c used by controller)

6. Use other diagnostic checks such as the sign of the computed process gain