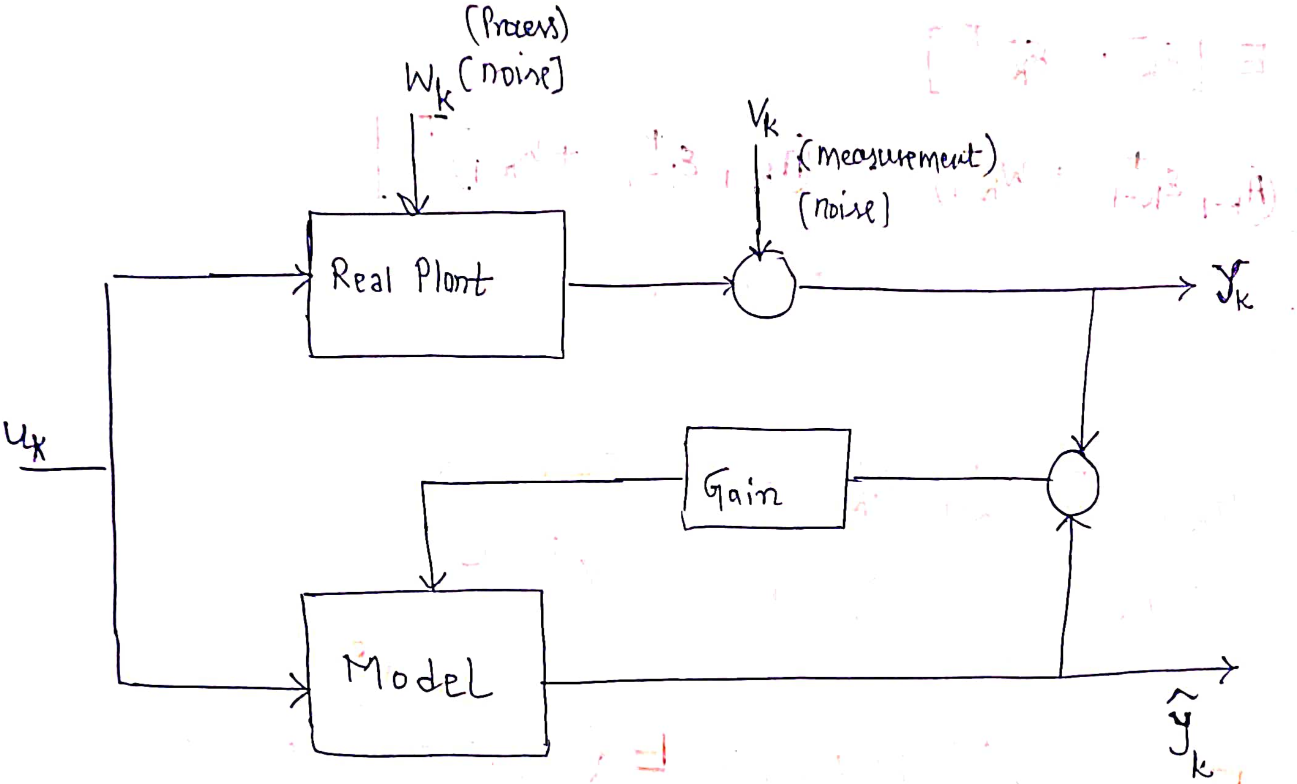


Kalman filter

u_k → Input vector

W_k → Process Noise

V_k → Measurement noise

y_k → Output Vector [Observed measurement] or Sensor output
(measurement at time step k.

$\hat{y}_k$ = Predicted output vector

x_k = Parameter vector that we want to estimate
(unknown)

Note

$x_k \rightarrow$ state vector

$y_k \rightarrow$ Measurement vector

$u_k \rightarrow$ Input control vector

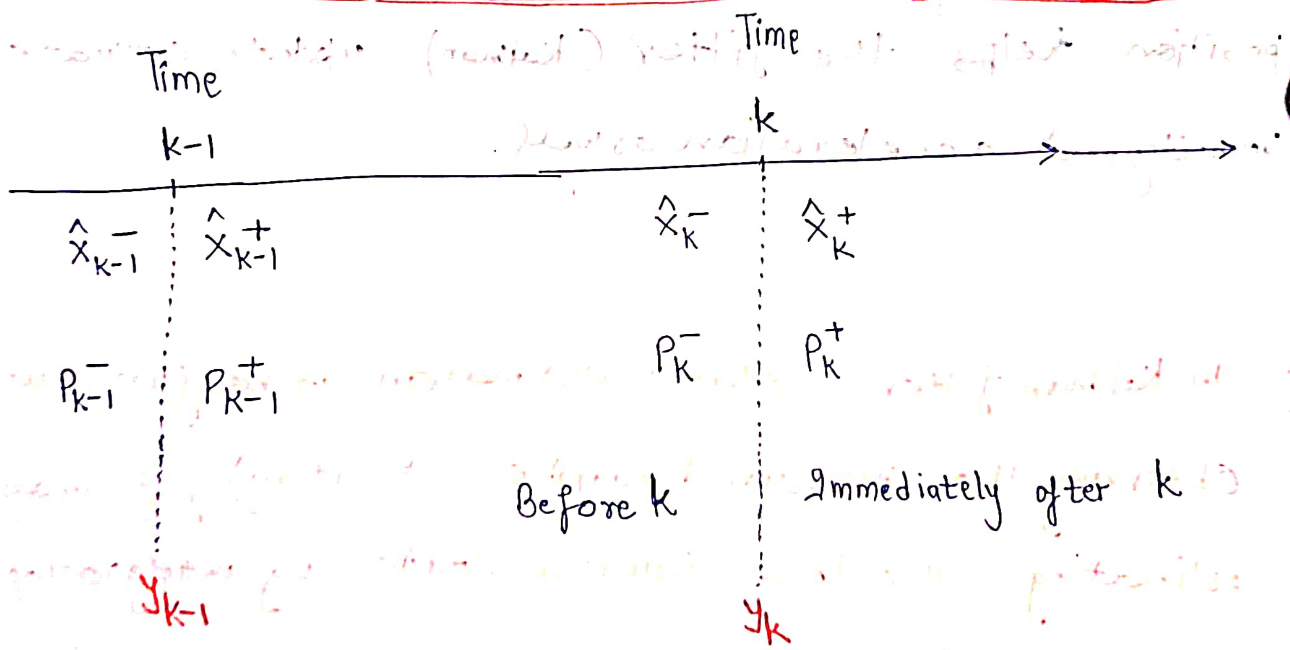
w_k : Process Noise (White noise), $E[w_k] = 0$

v_k : Measurement noise (white noise), $E[v_k] = 0$

Q_k : Symmetric, positive semidefinite, $Q \geq 0$

R_k : Symmetric, positive definite, $R > 0$

P_k : Symmetric, positive semidefinite, $P_k \geq 0$



Note: Discrete time instant $= k$

Note: The lower suffix k refers to the time step or discrete time Index.

KALMAN FILTER

$$x_k = A_{k-1} x_{k-1} + B_{k-1} u_{k-1} + w_{k-1}$$

$$y_k = C_k x_k + v_k$$

Note:

A_{k-1}, B_{k-1}, C_k, u_{k-1}, y_k } Known

$$\left.\begin{aligned} Q_k &= E\{w_k \cdot w_k^T\} \\ R_k &= E\{v_k \cdot v_k^T\} \end{aligned}\right\} \text{known}$$

state vector x_k } unknown

$\boxed{\text{A priori estimation}}$:

$$\hat{x}_k^- = A_{k-1} \hat{x}_{k-1}^+ + B_{k-1} u_{k-1}$$

$$P_k^- = E[\varepsilon_k^- \cdot \varepsilon_k^{-T}]$$

$$P_k^- = A_{k-1} P_{k-1}^+ A_{k-1}^T + Q_{k-1}$$

Proof: Let x_k is the parameter vector that we want to estimate.

$$\varepsilon_k^- = x_k - \hat{x}_k^- \quad \text{or} \quad x - \hat{x}_k^- \quad \text{(General form)}$$

$$= \{A_{k-1} x_{k-1} + \cancel{B_{k-1} u_{k-1}} + w_{k-1}\}$$

$$- \{A_{k-1} x_{k-1}^+ + \cancel{B_{k-1} u_{k-1}}\}$$

$$= A_{k-1} x_{k-1} - A_{k-1} \hat{x}_{k-1}^+ + w_{k-1}$$

$$= A_{k-1} (x_{k-1} - x_{k-1}^+) + w_{k-1} = A_{k-1} \varepsilon_{k-1}^+ + w_{k-1}$$

Note:

$$\varepsilon_k^- = x_k - \hat{x}_k^- \quad \text{or} \quad x - \hat{x}_k^-$$

$$\varepsilon_k^+ = x_k - \hat{x}_k^+ \quad \text{or} \quad x - \hat{x}_k^+$$

$$\varepsilon_{k-1}^- = x_{k-1} - \hat{x}_{k-1}^-$$

$$\varepsilon_{k-1}^+ = x_{k-1} - \hat{x}_{k-1}^+$$

A priori estimate:

$$\hat{x}_k^- = E[\, x_k \mid y_1, y_2 \ldots\ldots y_{k-1}]$$

= Estimate of x_k before we process the measurement y_k at time k.

A posteriori estimate

$$\hat{x}_k^+ = E[\, x_k \mid y_1, y_2 \ldots\ldots y_{k-1}, y_k]$$

= Estimate of x_k after we process the measurement y_k at time k.

where $\varepsilon_{k-1}^{+} = x_{k-1} - \hat{x}_{k-1}^{+}$

A priori error covariance :- P_k^{-}

$P_k^{-} = E\,[\varepsilon_k^{-} \cdot \varepsilon_k^{-T}]$

$= E[(A_{k-1}\varepsilon_{k-1}^{+} + w_{k-1}) \cdot (A_{k-1}\varepsilon_{k-1}^{+} + w_{k-1})^T]$

$= E[(A_{k-1}\varepsilon_{k-1}^{+} + w_{k-1}) \cdot (\varepsilon_{k-1}^{+T} A_{k-1}^{T} + w_{k-1}^{T})]$

$= E\{ A_{k-1}\varepsilon_{k-1}^{+}\varepsilon_{k-1}^{+T}A_{k-1}^{T} + w_{k-1}\varepsilon_{k-1}^{+T}A_{k-1}^{T}$
$+ A_{k-1}\varepsilon_{k-1}^{+}w_{k-1}^{T} + w_{k-1}w_{k-1}^{T} \}$

$= A_{k-1}\mathbf{E}\{\varepsilon_{k-1}^{+}\cdot\varepsilon_{k-1}^{+T}\}A_{k-1} + \mathbf{E}\{w_{k-1}\cdot\varepsilon_{k-1}^{+T}\}A_{k-1}$
$+ A_{k-1}\mathbf{E}\{\varepsilon_{k-1}^{+}w_{k-1}^{T}\} + \mathbf{E}\{w_{k-1}\cdot w_{k-1}^{T}\}$

$= A_{k-1}\,E\{\varepsilon_{k-1}^{+}\varepsilon_{k-1}^{+T}\}A_{k-1}^{T} + E\{w_{k-1}\cdot w_{k-1}^{T}\}$

$= A_{k-1}P_{k-1}^{+}A_{k-1} + Q_{k-1}$

$$\therefore \boxed{P_k^{-} = A_{k-1}P_{k-1}^{+}A_{k-1} + Q_{k-1}}$$

A posteriori estimation :

$$\hat{x}_k^+ = \hat{x}_k^- + K_k \left(y_k - C_k \hat{x}_k^- \right)$$

x_k is the parameter vector that we want to estimate

$$\varepsilon_k^+ = x_k - \hat{x}_k^+$$

$$= x - \hat{x}_k^+$$

(General form)

$$= x - \left[\hat{x}_k^- + K_k \left(y_k - C_k \hat{x}_k^- \right) \right]$$

$$= x - \hat{x}_k^- - K_k \left(y_k - C_k \hat{x}_k^- \right)$$

$$= (x - \hat{x}_k^-) - K_k \left(C_k x + v_k - C_k \hat{x}_k^- \right)$$

$$= (x - \hat{x}_k^-) - K_k C_k x - K_k v_k + K_k C_k \hat{x}_k^-$$

$$= (x - \hat{x}_k^-) - K_k C_k (x - \hat{x}_k^-) - K_k v_k$$

$$= (I - K_k C_k)(x - \hat{x}_k^-) - K_k v_k$$

$$\varepsilon_k^+ = (I - K_k C_k)\, \varepsilon_k^- - K_k v_k$$

A posteriori covariance of error : P_k^+

$$P_k^+ = E\left[\varepsilon_k^+ \cdot \varepsilon_k^{+T}\right]$$

$$= E\left[\left[(I - K_k C_k)\varepsilon_k^- - K_k V_k\right] \cdot \left[(I - K_k C_k)\varepsilon_k^- - K_k V_k\right]^T\right]$$

$$= E\left[\left[(I - K_k C_k)\varepsilon_k^- - K_k V_k\right] \cdot \left[\varepsilon_k^{-T}(I - K_k C_k)^T - V_k^T K_k^T\right]\right]$$

$$= E\left[(I - K_k C_k)\varepsilon_k^- \cdot \varepsilon_k^{-T}(I - K_k C_k)^T - K_k V_k \varepsilon_k^{-T}(I - K_k C_k)^T - (I - K_k C_k)\varepsilon_k^- V_k^T K_k^T + K_k V_k \cdot V_k^T K_k^T\right]$$

$$= (I - K_k C_k)\, E\left[\varepsilon_k^- \varepsilon_k^{-T}\right](I - K_k C_k)^T - K_k \underbrace{E\left[V_k \varepsilon_k^{-T}\right]}_{\to 0}(I - K_k C_k)^T - (I - K_k C_k)\underbrace{E\left[\varepsilon_k^- V_k^T\right]}_{\to 0} K_k^T + K_k E\left[V_k \cdot V_k^T\right] K_k^T$$

Note

$\because$ $E\left[\varepsilon_k^- \varepsilon_k^{-T}\right] = P_k^-$

$E\left[V_k \varepsilon_k^{-T}\right] = E[V_k] \cdot E\left[\varepsilon_k^{-T}\right] = 0$

$E\left[\varepsilon_k^- V_k^T\right] = E\left[\varepsilon_k^-\right] \cdot E\left[V_k^T\right] = 0$

$E\left[V_k V_k^T\right] = R_k$

$$P_k^+ = (I - K_k C_k) P_k^- (I - K_k C_k)^T + K_k R_k K_k^T$$

$$= (P_k^- - K_k C_k P_k^-)(I - C_k^T K_k^T) + K_k R_k K_k^T$$

$$= P_k^- - K_k C_k P_k^- - P_k^- C_k^T K_k^T + K_k C_k P_k^- C_k^T K_k^T + K_k R_k K_k^T \quad (1)$$

Note

Cost function W_k :

$W_k = \text{trace}[P_k^+]$

$P_k^+ = E[\varepsilon_k^+ \cdot \varepsilon_k^{+T}]$

Note $\varepsilon_k^+ = X - \hat{X}_k^+$

$$\begin{bmatrix} \varepsilon_{1k}^+ \\ \varepsilon_{2k}^+ \\ \vdots \\ \varepsilon_{nk}^+ \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} - \begin{bmatrix} \hat{x}_{1k}^+ \\ \hat{x}_{2k}^+ \\ \vdots \\ \hat{x}_{nk}^+ \end{bmatrix}$$

Parametervector that we want to estimate

$$\begin{bmatrix} \varepsilon_{1k}^+ \\ \vdots \\ \varepsilon_{nk}^+ \end{bmatrix} = \begin{bmatrix} x_1 - \hat{x}_{1k}^+ \\ x_2 - \hat{x}_{2k}^+ \\ \vdots \\ x_n - \hat{x}_{nk}^+ \end{bmatrix}$$

$$\therefore P_k^+ = E\left\{ \begin{bmatrix} \varepsilon_{1k}^+ \\ \vdots \\ \varepsilon_{nk}^+ \end{bmatrix} \cdot \begin{bmatrix} \varepsilon_{1k}^+ & \cdots & \varepsilon_{nk}^+ \end{bmatrix} \right\}$$

$$= E\begin{bmatrix} (\varepsilon_{1k}^+)^2 & \cdots & \varepsilon_{1k}^+\varepsilon_{nk}^+ \\ \varepsilon_{nk}^+\varepsilon_{1k}^+ & \cdots & (\varepsilon_{nk}^+)^2 \end{bmatrix} = E\begin{bmatrix} (E_{1k}^+)^2 & \cdots & \\ \vdots & \ddots & \\ & & (\varepsilon_{nk})^2 \end{bmatrix}$$

Note

$$\therefore \text{trace}\left[P_k^+\right] = E\left[\varepsilon_{1k}^{+\,2}\right] + E\left[\varepsilon_{2k}^{+\,2}\right] + E\left[\varepsilon_{3k}^{+\,2}\right] + \dots\dots\dots E\left[\varepsilon_{nk}^{+\,2}\right]$$

Cost function W_k :

$$W_k = \text{Trace}\,(P_k^+)$$

$\therefore W_k = tr(P_k^-) - tr(K_k C_k P_k^-) - tr(P_k^- C_k^T K_k^T) + tr(K_k C_k P_k^- C_k^T K_k^T) + tr(K_k R_k K_k^T)$ — (2)

We want to minimize the cost W_k.

$$\left.\begin{aligned} &\therefore \frac{\partial W_k}{\partial K_k} = 0 \\ &\frac{\partial(\text{trace } P_k^+)}{\partial K_k} = 0 \end{aligned}\right\} \text{— (3)}$$

Note: $\frac{\partial}{\partial K_k} tr(P_k^-) = 0$

$$\frac{\partial}{\partial K_k} tr(K_k C_k P_k^-) = (C_k P_k^-)^T = (P_k^-)^T C_k^T = P_k^- C_k^T$$

$$\frac{\partial}{\partial K_k} tr(P_k^- C_k^T K_k^T) = P_k^- C_k^T$$

$$\frac{\partial}{\partial K_k} tr(K_k C_k P_k^- C_k^T K_k^T) = 2 K_k C_k P_k^- C_k^T$$

$$\frac{\partial}{\partial K_k} tr(K_k R_k K_k^T) = 2 K_k R_k$$

From equation (2) and (3): $\frac{\partial [W_k]}{\partial K_k} = 0$

$$0 - P_k^- C_k^T - P_k^- C_k^T + 2 K_k C_k P_k^- C_k^T + 2 K_k R_k = 0$$

$$K_k [R_k + C_k P_k^- C_k^T] = P_k^- C_k^T$$

$$\therefore \boxed{K_k = P_k^- C_k^T [R_k + C_k P_k^- C_k^T]^{-1}}$$

$\because$ $R_k + C_k P_k^- C_k^T$ is symmetric in nature.

$$\therefore \left\{ \left(R_k + C_k P_k^- C_k^T \right)^{-1} \right\}^T = \left(R_k + C_k P_k^- C_k^T \right)^{-1}$$

Now put the value of K_k in equation ①

$$P_k^+ = P_k^- - P_k^- C_k^T (R_k + C_k P_k^- C_k^T)^{-1} C_k P_k^-$$
$$- P_k^- C_k^T \{P_k^- C_k^T (R_k + C_k P_k^- C_k^T)^{-1}\}^T$$
$$+ \{P_k^- C_k^T (R_k + C_k P_k^- C_k^T)^{-1}\} \cdot C_k P_k^- C_k^T \cdot \{P_k^- C_k^T (R_k + C_k P_k^{-1} C_k^T)^{-1}\}^T$$
$$+ \{P_k^- C_k^T (R_k + C_k P_k^- C_k^T)^{-1}\} \cdot R_k \cdot \{P_k^- C_k^T (R_k + C_k P_k^- C_k^T)^{-1}\}^T$$

$$= P_k^- - P_k^- C_k^T (R_k + C_k P_k^- C_k^T)^{-1} C_k P_k^-$$
$$- P_k^- C_k^T (R_k + C_k P_k^- C_k^T)^{-1} \cdot C_k P_k^-$$
$$+ P_k^- C_k^T (R_k + C_k P_k^- C_k^T)^{-1} \cdot C_k P_k^- C_k^T \cdot (R_k + C_k P_k^- C_k^T)^{-1} C_k P_k^-$$
$$+ P_k^- C_k^T (R_k + C_k P_k^- C_k^T)^{-1} \cdot R_k \cdot (R_k + C_k P_k^- C_k^T)^{-1} C_k \cdot P_k^-$$

$$= P_k^- - 2 P_k^- C_k^T (R_k + C_k P_k^- C_k^T)^{-1} C_k P_k^-$$
$$+ P_k^- C_k^T (R_k + C_k P_k^- C_k^T)^{-1} \cdot [C_k P_k^- \cancel{C_k^T + R_k}] \cdot (\cancel{R_k + C_k P_k^- C_k^T})^{-1} C_k P_k^-$$

$$= P_k^- - 2 P_k^- C_k^T (R_k + C_k P_k^- C_k^T)^{-1} C_k P_k^-$$
$$+ P_k^- C_k^T (R_k + C_k P_k^- C_k^T)^{-1} C_k P_k^-$$

$$= P_k^- - P_k^- C_k^T (R_k + C_k P_k^- C_k^T)^{-1} C_k P_k^-$$

$$= P_k^- - K_k \cdot C_k P_k^-$$

$$\boxed{P_k^+ = (I - K_k C_k) P_k^-}$$

Summary of the Kalman filter

$$X_k = A_{k-1} X_{k-1} + B_{k-1} u_{k-1} + W_{k-1}$$

$$Y_k = C_k X_k + V_k$$

Assume initial estimate of state $= \hat{x}_0^+$

Assume initial covariance matrix of estimation error $= P_0^+$

For $k = 1, 2, 3, 4 \ldots\ldots$ we perform the following steps :

STEP-I : A priori (prediction) step :

A priori state estimate $(\hat{x}_k^-)$:

$$\hat{x}_k^- = A_{k-1} \hat{x}_{k-1}^+ + B_{k-1} u_{k-1}$$

A priori error covariance (P_k^-) :

$$P_k^- = A_{k-1} P_{k-1}^+ A_{k-1}^T + Q_{k-1}$$

STEP-II : A posteriori (update) step :

Kalman gain (K_k) :

$$K_k = P_k^- C_k^T \left(R_k + C_k P_k^- C_k^T \right)^{-1}$$

A posteriori state estimate $(\hat{x}_k^+)$:

$$\hat{x}_k^+ = \hat{x}_k^- + K_k \left(y_k - C_k \hat{x}_k^- \right)$$

A posteriori error covariance (P_k^+) :

$$P_k^+ = (I - K_k C_k) P_k^- (I - K_k C_k)^T + K_k R_k K_k^T$$

or $$P_k^+ = (I - K_k C_k) P_k^-$$

Note : Quantification of noise

W_k = Process Noise
V_k = Measurement noise

} uncorrelated (white) noise; wide sense stationary

$Q_k = E[W_k W_k^T]$ = Process noise covariance matrix

$R_k = E[V_k \cdot V_k^T]$ = Measurement noise covariance matrix

$E[W_k] = 0$ $\quad\quad$ $E[V_k] = 0$

W_k and V_k are assumed to be independent of each other. (as is often the case physically)

$E[W_t \cdot V_s^T] = 0$ $\quad$ For all t and s

$E[W_k \cdot V_k^T] = 0$

Auto correlation: It is the correlation between two time slices of same random variable

$E[V_{t,i} \cdot V_{s,i}] = 0$

$E[W_{t,j} \cdot W_{s,j}] = 0$

for all t and s
$t \neq s$

Cross correlation: The correlation between two different components of the same noise vector or the one between measurement and process noise.

$E[V_{t,i} \; V_{s,j}] = 0$ $\quad$ for all t and s , $t \neq s$

$E[W_{t,i} \; W_{s,j}] = 0$ $\quad$ for all i and j , $i \neq j$

$E[V_{t,i} \; W_{s,j}] = 0$ $\quad$ For all t and s
for all i and j

Optimal State Estimation

Kalman, H_∞, and Nonlinear Approaches

Dan Simon
Cleveland State University

WILEY-INTERSCIENCE
A JOHN WILEY & SONS, INC., PUBLICATION

2.160 | Spring 2006 | Graduate

Identification, Estimation, And Learning

Syllabus
Calendar
Lecture Notes
Assignments
Exams
Projects

Course Description

This course provides a broad theoretical basis for system identification, estimation, and learning. Students will study least squares estimation and its convergence properties, Kalman filters, noise dynamics and system representation, function approximation theory, neural nets, radial basis functions, wavelets, ... Show more

Course Info

INSTRUCTOR

Prof. Harry Asada

DEPARTMENTS

Mechanical Engineering

TOPICS

- Engineering
  - Computer Science
    - Algorithms and Data Structures
    - Artificial Intelligence
  - Electrical Engineering
    - Signal Processing
  - Mechanical Engineering
    - Dynamics and Control
  - Systems Engineering
    - Numerical Simulation

LEARNING RESOURCE TYPES

Problem Sets Exams Lecture Notes

The Mars rover relies on sophisticated identification and estimation techniques to navigate the Martian terrain. (Image courtesy of NASA.)

Download Course

Feedback

2.160 | Spring 2006 | Graduate

Identification, Estimation, And Learning

Syllabus
Calendar
Lecture Notes
Assignments
Exams
Projects

Lecture Notes

Lecture Notes Table of Contents (PDF)

Available lecture notes are listed below.

LEC #	TOPICS
1	Introduction (PDF 1) (PDF 2)
Part I: Estimation	
2	Recursive Least Square (RLS) Algorithms (PDF)
3	Properties of RLS (PDF)
4	Random Processes, Active Noise Cancellation (PDF)
5	Discrete Kalman Filter-1 (PDF)
6	Discrete Kalman Filter-2 (PDF)
7	Continuous Kalman Filter (PDF)
8	Extended Kalman Filter (PDF)
Part 2: Representation and Learning	
9	Prediction Modeling of Linear Systems (PDF)
10	Model Structure of Linear Time-invariant Systems (PDF)
11	Time Series Data Compression, Laguerre Series Expansion (PDF)
12	Non-linear Models, Function Approximation Theory, Radial Basis Functions (PDF)
13	Neural Networks (PDF)

Course Info

INSTRUCTOR

Prof. Harry Asada

DEPARTMENTS

Mechanical Engineering

AS TAUGHT IN

Spring 2006

LEVEL

Graduate

TOPICS

- Engineering
  - Computer Science
    - Algorithms and Data Structures
    - Artificial Intelligence
  - Electrical Engineering
    - Signal Processing
  - Mechanical Engineering
    - Dynamics and Control
  - Systems Engineering
    - Numerical Simulation

LEARNING RESOURCE TYPES