Optimal State Estimation: Kalman Approach

Harini Kulatunga

Department of Computer Science
Imperial College London

08 June 2007 / AESOP Meeting

First Application

First publicly acknowledged application was made for the APOLLO Mission . . .

- In 1960s NASA Ames Research center was studying navigation and control of the Apollo space capsule.
  - analysis of the trajectory of the shuttle.
  - simulate mid-course correction - Linear Perturbation Methods (LPM).
- Lunar navigation - use of pilot observations to estimate the state
  - such data needs to be processed efficiently in order not to burden the IBM 704 that was used.
  - nature of the problem is non-linear and require irregular series of discrete measurements.
- Wiener filtering - was used for missile navigation.
- Kalman filtering - Linear and Sequential

First Application

Successful Kalman formulation for lunar navigation

- LPM concepts used in guidance produces a linear system to which Kalman theory can be applied
- solves the problem faced by speed and storage limitations of the IBM 704
- original Kalman formulation [Kalman, 1960] was decomposed into two-parts to suit the practical considerations of the lunar mission
- re-linearisation about the current trajectory estimate as opposed to about a reference trajectory - estimated state is closer to the actual "on the average"

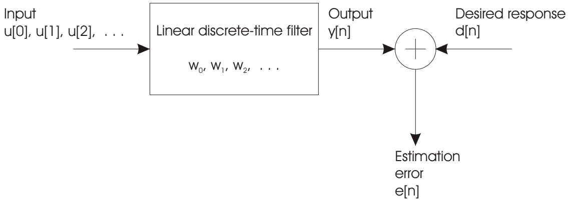
Norbert Wiener

Kalman's seminal paper is based on 'solving' the Wiener problem.

- pioneer in the study of stochastic and noise processes, contributing to electronic engineering, electronic communication, and control systems.
- best known as the founder of cybernetics, a field that formalizes the notion of feedback and has implications for engineering, systems control, computer science, biology, philosophy, and the organization of society.
- arguably the most distinguished scientist to decline an invitation to join the Manhattan Project

Optimal filter

- Filtering problem - extracting a waveform buried in noise.
- Optimal filtering - how the two can 'best' be separated.



- the filter is specified in terms of its *impulse response* $w_0, w_1, \ldots$

Optimal filtering

- The filter output $y[n]$ provides an estimate of the desired response $d[n]$.
- 'tune' the filter coefficients to give the minimum estimation error $e[n] = y[n] - d[n]$
- Criterion for statistical optimisation,
  - mean-square value of $e[n]$, MSE
  - Expectation of the absolute value of $e[n]$
  - Expectation of third or higher powers of the absolute value of $e[n]$

Wiener filters

Are a class of optimum linear discrete-time filters formulated for complex-valued *time-series*. Filter design minimises the MSE.

- optimum filter coefficients that minimise the MSE
- The cost function $J = E[|e[n]|^2]$
- The gradient operator ∇_k is defined in terms of each filter coefficient w_k
- ∇ operator is used to find the stationary points of the cost function J

Wiener filtering

- The partial derivative calculations proves the important *Principle of Orthogonality*,

$$J_{min} = E[|e_o[n]|^2] \text{ iff } E\left[u[n-k]e_o^*[n]\right] = 0 \qquad (1)$$

  and lead to the well known Wiener-Hopf equations,

$$\sum_{i=0}^{\infty} w_{oi} \underbrace{r[i-k]}_{E[u[n-k]u*[n-i]]} = \underbrace{p[-k]}_{E[u[n-k]d*[n]]}, \quad \forall k \qquad (2)$$

- The solution of Wiener-Hopf equations gives the optimal filter design $\mathbf{w}_o$, $\mathbf{w}_o = \mathbf{R}^{-1}\mathbf{p}$

Kalman Filter

- Wiener estimates the time-domain impulse response of the filter based on the entire past observed data
- Kalman filter is described in terms of state-space approach to control engineering problems
  - state estimates are computed recursively
  - reduced storage requirements
    - only previous state estimate requires storage
    - eliminates the need for storing the entire past observation data
- suited for implementation on a digital computer

State-space model

- A description of the internal and external characteristics of a linear, finite-dimensional stochastic system.

$$\begin{aligned} \text{process equation:} \quad \mathbf{x}[n+1] &= \Phi[n+1, n]\mathbf{x}[n] + \mathbf{v}_1[n] \\ \text{measurement equation:} \quad \mathbf{y}[n] &= \mathbf{C}[n]\mathbf{x}[n] + \mathbf{v}_2[n] \end{aligned}$$

- $\mathbf{v}_1[n]$, $\mathbf{v}_2[n]$ noise vectors which are statistically independent
- $\Phi[n+1, n]$ state transition matrix
- $\mathbf{C}[n]$ measurement matrix
- Any stochastic system described by a difference equation can be recast in the above form.

Kalman filtering problem

- Process equation

$$\mathbf{x}[n+1] = \Phi[n+1, n]\mathbf{x}[n] + \mathbf{v}_1[n] \tag{3}$$

$\mathbf{v}_1[n]$ is modelled as $\aleph(0, \mathbf{Q}_1[n])$

- Measurement equation

$$\mathbf{y}[n] = \mathbf{C}[n]\mathbf{x}[n] + \mathbf{v}_2[n] \tag{4}$$

$\mathbf{v}_2[n]$ is modelled as $\aleph(0, \mathbf{Q}_2[n])$

- $E\left[\mathbf{v}_1[n]\mathbf{v}_2^H[k]\right] = \mathbf{0} \quad \forall\ n, k$
- Matrices $\Phi[n+1, n]$, $\mathbf{C}[n]$ are assumed *known*

Kalman filtering problem

- Goal is to use the observed data $\mathcal{Y}_n = (\mathbf{y}[1], \ldots, \mathbf{y}[n])$ to find the MMSE estimate of the state vector $\mathbf{x}[i]$
- Problem can be,
  - if $i = n$, a filtering problem
  - if $i > n$, a prediction problem
  - if $1 \leq i < n$, a smoothing problem
- This problem can be solved using the innovations approach [Kailath, 1968, 1970]

Innovation process

- The forward prediction error,

$$\alpha[n] \triangleq \mathbf{y}[n] - \hat{\mathbf{y}}[n|\mathcal{Y}_{n-1}], \qquad n = 1, 2, \ldots \tag{5}$$

  where $\hat{\mathbf{y}}[n|\mathcal{Y}_{n-1}]$ is the predictable part of $\mathbf{y}[n]$ determined by past observations

- From the principle of orthogonality (derived from the condition for minimising the MSE cost function),

$$\underbrace{\alpha[n]}_{new\ information} \quad \text{orthogonal to} \quad \mathcal{Y}_{n-1} \tag{6}$$

- $\alpha[n]$ is a measure of new information in the random variable $y[n]$; Hence the name "innovation"

Recursive MMSE Estimate I

MMSE estimate of $\mathbf{x}[n-1]$ is $\hat{\mathbf{x}}[n-1|\mathcal{Y}_{n-1}]$

- Store previous estimate $\hat{\mathbf{x}}[n-1|\mathcal{Y}_{n-1}]$ and exploit it to compute *updated* estimate $\hat{\mathbf{x}}[n|\mathcal{Y}_n]$ in the light of new observation $y[n]$
- However there is one-to-one correspondence between the observation data and innovations (from Gram-Schmidt orthogonalisation)
- The updated MMSE state estimate given the innovations,

$$\hat{\mathbf{x}}[i|\mathcal{Y}_n] = \sum_{k=1}^{n} \mathbf{B}(k)\alpha(k) \tag{7}$$

Recursive MMSE Estimate II

- The recursion,

$$\hat{\mathbf{x}}[i|\mathcal{Y}_n] = \sum_{k=1}^{n-1} \mathbf{B}(k)\alpha(k) + \mathbf{B}(n)\alpha(n) \quad (8)$$

where

$$\mathbf{B}(k) = \frac{E\left[\mathbf{x}[i]\alpha^H[k]\right]}{E\left[\alpha[k]\alpha^H[k]\right]} \quad (9)$$

chosen to minimise the MSE error $\mathbf{x}[i] - \hat{\mathbf{x}}[i|\mathcal{Y}_n]$ (**??**).

Correlation Matrices

- predicted state error $\epsilon[n, n-1] = \mathbf{x}[n] - \hat{\mathbf{x}}[n|\mathcal{Y}_{n-1}]$
- predicted state error correlation matrix

$$\mathbf{P}[n, n-1] = E\left[\epsilon[n, n-1]\epsilon^H[n, n-1]\right] \tag{10}$$

- Now $\alpha[n] = \mathbf{y}[n] - \mathbf{C}[n]\hat{\mathbf{x}}[n|\mathcal{Y}_{n-1}] = \mathbf{C}[n]\epsilon[n, n-1] + \mathbf{v}_2[n]$
- innovation correlation matrix

$$\Sigma[n] = \mathbf{C}[n]\mathbf{P}[n, n-1]\mathbf{C}^H[n] + \mathbf{Q}_2[n] \tag{11}$$

Recursive MMSE Estimate III

- For $i = n + 1$, the predictive state estimate,

$$\hat{\mathbf{x}}[n+1|\mathcal{Y}_n] = \sum_{k=1}^{n-1} E\left[\mathbf{x}[n+1]\alpha^H[k]\right]\Sigma^{-1}[k]\alpha[k] + E\left[\mathbf{x}[n+1]\alpha^H[n]\right]\Sigma^{-1}[n]\alpha[n] \quad (12)$$

- Using the process equation of the Kalman problem

$$\hat{\mathbf{x}}[n+1|\mathcal{Y}_n] = \Phi(n+1,n)\hat{\mathbf{x}}[n|\mathcal{Y}_{n-1}] + \underbrace{\mathbf{K}[n]}_{Kalman\ Gain}\alpha[n] \quad (13)$$

where $\hat{\mathbf{x}}[n|\mathcal{Y}_{n-1}] = E\left[\mathbf{x}[n]\alpha^H[k]\right]\Sigma^{-1}[k]\alpha[k]$ and $\mathbf{K}[n] = E\left[\mathbf{x}[n+1]\alpha^H[n]\right]\Sigma^{-1}[n]$ and $\mathbf{v}_1[n]$ orthogonal $\alpha[n]$.

Computation of $\mathbf{K}[n]$

$$E\left[\mathbf{x}[n+1]\alpha^H[n]\right] = \Phi[n+1,n]E\left[\mathbf{x}[n]\alpha^H[n]\right] \tag{14}$$

- But $\alpha[n] = \mathbf{C}[n]\epsilon[n,n-1] + \mathbf{v}_2[n]$, $E\left[\mathbf{x}[0]\mathbf{v}_2^H[n]\right] = 0$ and $\mathbf{x}[n] = \epsilon[n,n-1] + \hat{\mathbf{x}}[n|\mathcal{Y}_{n-1}]$, $\hat{\mathbf{x}}[n|\mathcal{Y}_{n-1}]$ orthogonal $\epsilon[n,n-1]$
- Therefore,

$$\begin{aligned} E\left[\mathbf{x}[n+1]\alpha^H[n]\right] &= \Phi[n+1,n]E\left[\mathbf{x}[n]\epsilon^H[n,n-1]\right]\mathbf{C}^H[n] \\ &= \Phi[n+1,n] \\ &\qquad E\left[\epsilon[n,n-1]\epsilon^H[n,n-1]\right]\mathbf{C}^H[n] \\ &= \Phi[n+1,n]\mathbf{P}[n,n-1]\mathbf{C}^H[n] \end{aligned} \tag{15}$$

Computation of $\mathbf{K}[n]$

- Therefore

$$\mathbf{K}[n] = \Phi[n+1, n]\mathbf{P}[n, n-1]\mathbf{C}^H[n]\Sigma^{-1}[n] \quad (16)$$

- The Kalman gain is expressed in terms of the predicted state error correlation and innovation correlation coefficient respectively.

Recursive computation of $\mathbf{P}[n, n-1]$

- Substituting the state-space equations and recursive state estimation equation in the state estimation error gives,

$$
\begin{aligned}
\epsilon[n+1, n] &= \mathbf{x}[n+1] - \hat{\mathbf{x}}[n+1|\mathcal{Y}_n] \\
&= \Phi[n+1, n]\left[\mathbf{x}[n] - \hat{\mathbf{x}}[n|\mathcal{Y}_{n-1}]\right] \\
&\quad -\mathbf{K}[n]\left[\mathbf{y}[n] - \mathbf{C}[n]\hat{\mathbf{x}}[n|\mathcal{Y}_{n-1}]\right] + \mathbf{v}_1[n] \\
&= \left[\Phi[n+1, n] - \mathbf{K}[n]\mathbf{C}[n]\right]\epsilon[n, n-1] \\
&\quad -\mathbf{K}[n]\mathbf{v}_2[n] + \mathbf{v}_1[n]
\end{aligned}
$$

- Since $\mathbf{P}[n+1, n] = E\left[\epsilon[n+1, n]\epsilon^H[n+1, n]\right]$,

$$
\begin{aligned}
\mathbf{P}[n+1, n] &= \left[\Phi[n+1, n] - \mathbf{K}[n]\mathbf{C}[n]\right]\mathbf{P}[n, n-1]\left[\cdot\right] \\
&\quad +\mathbf{K}[n]\mathbf{Q}_2[n]\mathbf{K}^H[n] + \mathbf{Q}_1[n]
\end{aligned}
$$

Recursive computation of $\mathbf{P}[n, n-1]$

- Substituting for $\mathbf{K}[n]$, $\Sigma[n]$ gives,

$$\mathbf{P}[n+1, n] = \Phi[n+1, n]\mathbf{P}[n]\Phi^H[n+1, n] + \mathbf{Q}_1[n] \tag{17}$$

  where $\mathbf{P}[n] = \{\mathbf{I} - \Phi[n, n+1]\mathbf{K}[n]\mathbf{C}[n]\}\,\mathbf{P}[n, n-1]$
  and $\Phi[n+1, n]\Phi[n, n+1] = \mathbf{I}$ (property of state transition matrix)
- Riccati difference equation for recursive computation of the predicted state error correlation

Kalman's one-step prediction algorithm

Summarizing ...

- $\alpha[n] = \mathbf{y}[n] - \mathbf{C}[n]\hat{\mathbf{x}}[n|\mathcal{Y}_{n-1}]$ (Innovation)
- $\Sigma[n] = \mathbf{C}[n]\mathbf{P}[n, n-1]\mathbf{C}^H[n] + \mathbf{Q}_2[n]$ (Innovation Correlation)
- $\mathbf{K}[n] = \Phi[n+1, n]\mathbf{P}[n, n-1]\mathbf{C}^H[n]\Sigma^{-1}[n]$ (Kalman Gain)
- $\hat{\mathbf{x}}[n+1|\mathcal{Y}_n] = \Phi(n+1, n)\hat{\mathbf{x}}[n|\mathcal{Y}_{n-1}] + \mathbf{K}[n]\alpha[n]$ (one-step prediction)
- $\mathbf{P}[n] = \{\mathbf{I} - \Phi[n, n+1]\mathbf{K}[n]\mathbf{C}[n]\}\,\mathbf{P}[n, n-1]$ (error correlation)
- $\mathbf{P}[n+1, n] = \Phi[n+1, n]\mathbf{P}[n]\Phi^H[n+1, n] + \mathbf{Q}_1[n]$ (predicted error correlation)

Kalman Filtering

Filtered state estimate is $\hat{\mathbf{x}}[n|\mathcal{Y}_n]$

- From the process equation, $\hat{\mathbf{x}}[n+1|\mathcal{Y}_n] = \Phi[n+1,n]\hat{\mathbf{x}}[n|\mathcal{Y}_n] + \underbrace{\mathbf{v}_1(n|\mathcal{Y}_n)}_{=0}$
- $\hat{\mathbf{x}}[n|\mathcal{Y}_n] = \Phi^{-1}(n+1,n)\hat{\mathbf{x}}[n+1|\mathcal{Y}_n]$
- Substituting for the prediction estimate,

$$\hat{\mathbf{x}}[n|\mathcal{Y}_n] = \hat{\mathbf{x}}[n|\mathcal{Y}_{n-1}] + \Phi^{-1}(n+1,n)\mathbf{K}[n]\alpha[n] \tag{18}$$

where $\hat{\mathbf{x}}[n|\mathcal{Y}_{n-1}] = \Phi[n,n-1]\hat{\mathbf{x}}[n-1|\mathcal{Y}_{n-1}]$

Kalman filter algorithm

- Initial conditions,

$$\hat{\mathbf{x}}[0|\mathcal{Y}_0] = \mathbf{0}$$
$$\mathbf{P}[0] = \mathbf{P}_0$$

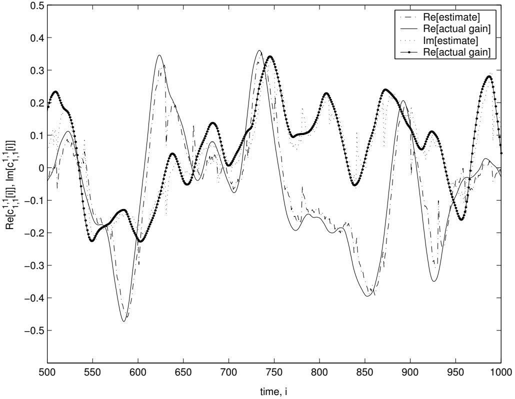
- Observations, $(\mathbf{y}[1], \mathbf{y}[2], \ldots, \mathbf{y}[n])$
- Known parameters, $\Phi[n+1, n]$, $\mathbf{C}[n]$, $\mathbf{Q}_1[n]$, $\mathbf{Q}_2[n]$

Prediction-Correction Steps

- Prediction,
  $\hat{\mathbf{x}}[n|\mathcal{Y}_{n-1}] = \Phi[n, n-1]\hat{\mathbf{x}}[n-1|\mathcal{Y}_{n-1}]$
  $\mathbf{P}[n, n-1] = \Phi[n, n-1]\mathbf{P}[n-1]\Phi^H[n, n-1] + \mathbf{Q}_1[n-1]$
- Correction,
  $\alpha[n] = \mathbf{y}[n] - \mathbf{C}[n]\hat{\mathbf{x}}[n|\mathcal{Y}_{n-1}]$
  $\Sigma[n] = \mathbf{C}[n]\mathbf{P}[n, n-1]\mathbf{C}^H[n] + \mathbf{Q}_2[n]$
  $\mathbf{K}[n] = \Phi[n, n-1]\mathbf{P}[n, n-1]\mathbf{C}^H[n]\Sigma^{-1}[n]$
  $\hat{\mathbf{x}}[n|\mathcal{Y}_n] = \hat{\mathbf{x}}[n|\mathcal{Y}_{n-1}] + \Phi^{-1}(n+1, n)\mathbf{K}[n]\alpha[n]$
  $\mathbf{P}[n] = \{\mathbf{I} - \Phi[n, n+1]\mathbf{K}[n]\mathbf{C}[n]\}\,\mathbf{P}[n, n-1]$

Channel Tracking Example in Wireless Communication

- Kalman filter tracking of a rapidly varying complex wireless channel between a Tx-Rx pair.



Summarising ...

- Prediction, Filtering and Smoothing
- Estimation of hidden states
- Measure of estimation quality
- Robust

Variations of the Filter

- Discrete-Discrete - considered here
- Continuous-Discrete
- Extended Kalman filter - non-linear parameters
- Square-root Kalman algorithms - combats filter divergence

Performance Applications

- Soule, A. Salamatian, K. Nucci, A. Taft, N. 2005. "Traffic matrix tracking using Kalman filters", Special issue on the first ACM SIGMETRICS workshop on large scale network inference (LSNI 2005), December 2005.
- Litoiu, M. 2007. "A performance analysis method for autonomic computing systems", ACM Transactions on Autonomous and Adaptive Systems (TAAS), March 2007.
- Soule, A. Lakhina, A. Taft, N. Papagiannaki, K. Salamatian, K. Nucci, A. Crovella, M. Diot, C. 2005. "Traffic matrices: balancing measurements, inference and modeling", Proceedings of the 2005 ACM SIGMETRICS international conference on Measurement and modeling of computer systems SIGMETRICS '05, June 2005.

Performance Applications

- Alouf, S. Altman, E. Barakat, C. Nain, P. 2003. "Estimating membership in a multicast session", Proceedings of the 2003 ACM SIGMETRICS international conference on Measurement and modeling of computer systems SIGMETRICS '03, June 2003.
- Dwyer, M. B. Pasareanu, C. S. 1998. "Filter-based model checking of partial systems", ACM SIGSOFT Software Engineering Notes , Proceedings of the 6th ACM SIGSOFT international symposium on Foundations of software engineering SIGSOFT '98/FSE-6, November 1998.

The End.