4F7 Adaptive Filters (and Spectrum Estimation)

Kalman Filter

Sumeetpal Singh
Email : sss40@eng.cam.ac.uk

1 Outline

- State space model
- Kalman filter
- Examples

2 Parameter Estimation

- We have repeated observations of a random variable x through

$$y(n) = x + v(n) \quad \text{for } n = 1, 2, \ldots$$

where $\{v(n)\}$ is a zero-mean scalar noise sequence, $E(v(n)v(l)) = 0$ for $n \neq l$, $E(v(n)x) = 0$, $E\left\{v(n)^2\right\} = \sigma_v^2$, $\sigma_0^2 = E\left(x^2\right)$, $E\left(x\right) = 0$

- Aim: at time n, $n > 0$, compute the optimum linear estimator for x using $\{y(1), \ldots, y(n)\}$. That is we want a linear estimator of x of the form

$$\widehat{x}(n) = \sum_{i=1}^{n} a_i y(i)$$

where the coeffecients a_i are chosen so that the mean square error

$$E\left\{(x - \widehat{x}(n))^2\right\}$$

is minimised

- Then, find a recursion for $\widehat{x}(n)$, i.e. relate $\widehat{x}(n+1)$ with $\widehat{x}(n)$

- The solution to this problem is

$$\widehat{x}(n+1) = \frac{\sigma(n)^2}{\sigma(n)^2 + \sigma_v^2}\left(y(n+1) - \widehat{x}(n)\right) + \widehat{x}(n)$$

where $\sigma(n)^2 = E\left(\left(x - \widehat{x}(n)\right)^2\right)$ satisfies the recursion

$$\sigma(n)^2 = \sigma(n-1)^2 - \frac{\sigma(n-1)^4}{\sigma(n-1)^2 + \sigma_v^2}$$

- Initialise by setting $\sigma(0)^2 = E\left(x^2\right)$ and $\widehat{x}(0) = 0$
- This model is too simple though

3 Linear Estimator with state dynamics

- We now extend the linear estimation problem to

$$\begin{aligned} x(n+1) &= x(n) + w(n+1) \\ y(n) &= x(n) + v(n) \end{aligned}$$

where $\{w(n)\}$ and $\{v(n)\}$ independent zero-mean random variables satisfying

$$\begin{aligned} E\left(v(n)v(l)\right) &= \sigma_v^2 \delta_{n,l}, \\ E\left(w(n)w(l)\right) &= \sigma_w^2 \delta_{n,l}, \end{aligned}$$

where $\delta_{n,l} = 0$ for $n \neq l$, $\delta_{n,l} = 1$ for $n = l$, $E\left(x(0)^2\right) = \sigma(0)^2$, $E\left(x(0)\right) = 0$. Noises $\{w(n)\}$ and $\{v(n)\}$ are independent of $x(0)$.

- Aim: at time n, compute the optimum linear estimator for $x(n)$ using $\{y(1), \ldots, y(n)\}$, i.e., $\widehat{x}(n) = \sum_{i=1}^{n} a_i y(i)$. Then, find a recursion relating $\widehat{x}(n+1)$ with $\widehat{x}(n)$

- Same problem as before except $x(n)$ has dynamics
- The solution to this problem is

$$\begin{aligned}\widehat{x}(n+1) &= \frac{\sigma(n)^2+\sigma_w^2}{\sigma(n)^2+\sigma_w^2+\sigma_v^2}\left(y(n+1)-\widehat{x}(n)\right)+\widehat{x}(n)\\ &= K(n)\left(y(n+1)-\widehat{x}(n)\right)+\widehat{x}(n)\end{aligned}$$

where

$$\sigma(n)^2 = \sigma(n-1)^2+\sigma_w^2-\frac{\left(\sigma(n-1)^2+\sigma_w^2\right)^2}{\sigma(n-1)^2+\sigma_w^2+\sigma_v^2}$$

(Initialisation as before.)

4 A State-Space Model

- A simple **state-space** model is the following

$$x(n+1) = x(n) + w(n+1)$$
$$y(n) = x(n) + v(n)$$

where $\{w(n)\}$ and $\{v(n)\}$ independent zero-mean random variables satisfying

$$E\left(v(n)v(k)\right) = \sigma_v^2 \delta_{n,k},$$
$$E\left(w(n)w(k)\right) = \sigma_w^2 \delta_{n,k},$$

- We gave a recursion for the optimum linear estimator of $x(n)$ using $\{y(1), \ldots, y(n)\}$, i.e. linear in the sense $\widehat{x}(n) = \sum_{i=1}^{n} a_i y(i)$.
- The recursion that related $\widehat{x}(n+1)$ to $\widehat{x}(n)$ is the Kalman filter

- Consider a state space model with vector valued states and observations

$$\begin{array}{ll} \textit{State Equation} & \mathbf{x}(n) = \mathbf{A}(n)\mathbf{x}(n-1) + \mathbf{w}(n) \\ \textit{Observation Equation} & \mathbf{y}(n) = \mathbf{C}(n)\mathbf{x}(n) + \mathbf{v}(n) \end{array}$$

  where $\{\mathbf{w}(n)\}$ and $\{\mathbf{v}(n)\}$ are independent zero-mean random variables (vectors) with

$$E\left(\mathbf{v}(n)\mathbf{v}^{\mathrm{T}}(k)\right) = \mathbf{Q}_v(n)\delta_{n,k},$$
$$E\left(\mathbf{w}(n)\mathbf{w}^{\mathrm{T}}(k)\right) = \mathbf{Q}_w(n)\delta_{n,k},$$

  and $\mathbf{x}(0)$ has mean $\mathbf{m}(0)$ and covariance matrix $\mathbf{P}(0)$. $\{\mathbf{w}(n)\}$, $\{\mathbf{v}(n)\}$ and $\mathbf{x}(0)$ are independent

- The aim is to derive the Kalman filter for this more complicated model. The previous solutions can be recovered as a special case

5 An example of a state-space model

- We want to track a target but only have access to noisy measurements of its position
- A **1D target model**
  - the target's position at time t is $p(t)$
  - the target's velocity at time t is $\dot{p}(t)$
  - we know that $\frac{d}{dt}p(t) = \dot{p}(t)$, and $\frac{d}{dt}\dot{p}(t) = \ddot{p}(t)$ where $\ddot{p}(t)$ is the acceleration
- We are able to measure the target's position with error
- We will discretise this model with a time step T

$$\begin{aligned} p(nT) &= p((n-1)\,T) + T\dot{p}((n-1)\,T), \\ \dot{p}(nT) &= \dot{p}((n-1)\,T) + T\ddot{p}((n-1)\,T) \end{aligned}$$

- The discretised observation model is

$$y(nT) = p(nT) + v(nT)$$

where $v(nT)$ is zero mean white noise

- Now define the following variables

$$\mathbf{x}(n) = [p(nT), \dot{p}(nT)]^{\mathrm{T}},$$
$$y(n) = y(nT)$$

- We can write

$$\mathbf{x}(n) = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \mathbf{x}(n-1) + \begin{bmatrix} 0 \\ T \end{bmatrix} w(n),$$
$$y(n) = [1, 0]\mathbf{x}(n) + v(n)$$

The accelleration is being modelled as noise $w(n)$

- In the literature, the state dynamics is also defined as

$$\mathbf{x}(n) = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \mathbf{x}(n-1) + \begin{bmatrix} T/2 \\ T \end{bmatrix} w(n)$$

6 Derivation of the Kalman Filter

- The optimum linear estimator of $\mathbf{x}(n)$ using all the measurement vectors up to time n may be expressed as

$$\hat{\mathbf{x}}(n) = \mathbf{K}'(n)\hat{\mathbf{x}}(n-1) + \mathbf{K}(n)\mathbf{y}(n)$$

  where $\hat{\mathbf{x}}(n-1)$ is the best linear unbiased estimate of $\mathbf{x}(n-1)$ based on the observations $\mathbf{y}(1), \ldots, \mathbf{y}(n-1)$

- $\mathbf{K}'(n)$, $\mathbf{K}(n)$ are the gain matrices to be derived
- We will resolve $\mathbf{K}'(n)$ be requiring unbiasedness of $\hat{\mathbf{x}}(n)$ and $\mathbf{K}(n)$ by minimising the MSE
- Define the estimation error

$$\mathbf{e}(n) = \mathbf{x}(n) - \hat{\mathbf{x}}(n)$$

  and the corresponding error covariance matrix

$$\mathbf{P}(n) = E\left\{\mathbf{e}(n)\mathbf{e}(n)^{\mathrm{T}}\right\}$$

- We will need the following matrix differentiation formulae

$$\frac{d}{d\mathbf{K}}\mathrm{tr}(\mathbf{KA}) = \mathbf{A}^{\mathrm{T}}$$
$$\frac{d}{d\mathbf{K}}\mathrm{tr}(\mathbf{AK}^{\mathrm{T}}) = \mathbf{A}$$
$$\frac{d}{d\mathbf{K}}\mathrm{tr}(\mathbf{KAK}^{\mathrm{T}}) = 2\mathbf{KA}$$

- Assuming $\mathbf{K} = [k_{i,j}]$, understand $\frac{dg(\mathbf{K})}{d\mathbf{K}}$ as the matrix with i, j element equal to $\frac{dg(\mathbf{K})}{dk_{i,j}}$

- Now check this for $\frac{d}{d\mathbf{K}}\mathrm{tr}(\mathbf{KA})$

$$\frac{d\mathrm{tr}(\mathbf{KA})}{dk_{i,j}} = \frac{d}{dk_{i,j}}\sum_{l,m} k_{l,m}a_{m,l} = a_{j,i}.$$

- Let $\hat{\mathbf{x}}(n|n-1) = \mathbf{A}(n)\hat{\mathbf{x}}(n-1)$ and

$$\mathbf{e}(n|n-1) = \mathbf{x}(n) - \hat{\mathbf{x}}(n|n-1)$$

and the corresponding error covariance matrix

$$\mathbf{P}(n|n-1) = E\left\{\mathbf{e}(n|n-1)\mathbf{e}(n|n-1)^{\mathrm{T}}\right\}$$

- Now calculate $\mathbf{P}(n|n-1)$

$$\begin{aligned}
&(\mathbf{x}(n) - \hat{\mathbf{x}}(n|n-1))\,(\mathbf{x}(n) - \hat{\mathbf{x}}(n|n-1))^{\mathrm{T}} \\
&= (\mathbf{A}(n)\left[\mathbf{x}(n-1) - \hat{\mathbf{x}}(n-1)\right] + \mathbf{w}(n)) \\
&\times (\mathbf{A}(n)\left[\mathbf{x}(n-1) - \hat{\mathbf{x}}(n-1)\right] + \mathbf{w}(n))^{\mathrm{T}} \\
&= \mathbf{A}(n)\left[\mathbf{x}(n-1) - \hat{\mathbf{x}}(n-1)\right] \\
&\times \left[\mathbf{x}(n-1) - \hat{\mathbf{x}}(n-1)\right]^{\mathrm{T}}\mathbf{A}(n)^{\mathrm{T}} \\
&+ \mathbf{w}(n)\mathbf{w}(n)^{\mathrm{T}} + (\text{terms involving one } \mathbf{w}(n))
\end{aligned}$$

- Take the expectation

$$\mathbf{P}(n|n-1) = \mathbf{A}(n)\mathbf{P}(n-1)\mathbf{A}(n)^{\mathrm{T}} + \mathbf{Q}_w(n)$$

- Once $\mathbf{y}(n)$ is received, using gain matrices $\mathbf{K}'(n)$, $\mathbf{K}(n)$, the estimate is updated to

$$\hat{\mathbf{x}}(n) = \mathbf{K}'(n)\hat{\mathbf{x}}(n-1) + \mathbf{K}(n)\mathbf{y}(n)$$

- What restrictions are imposed on $\mathbf{K}'(n)$, $\mathbf{K}(n)$ when we require $\hat{\mathbf{x}}(n)$ to be unbiased?

$$\hat{\mathbf{x}}(n) = \mathbf{K}'(n)\hat{\mathbf{x}}(n-1) + \mathbf{K}(n)\left[\mathbf{C}(n)\mathbf{x}(n) + \mathbf{v}(n)\right]$$

Take the expectation

$$\begin{aligned} E\left\{\hat{\mathbf{x}}(n)\right\} &= \mathbf{K}'(n)E\left\{\hat{\mathbf{x}}(n-1)\right\} \\ &\quad + \mathbf{K}(n)E\left\{\left[\mathbf{C}(n)\mathbf{x}(n) + \mathbf{v}(n)\right]\right\} \\ &= \mathbf{K}'(n)E\left\{\mathbf{x}(n-1)\right\} + \mathbf{K}(n)\mathbf{C}(n)\mathbf{A}(n)E\left\{\mathbf{x}(n-1)\right\} \end{aligned}$$

Thus for unbiasedness, i.e. $E\left\{\hat{\mathbf{x}}(n)\right\} = E\left\{\mathbf{x}(n)\right\} = \mathbf{A}(n)E\left\{\mathbf{x}(n-1)\right\}$

$$\mathbf{K}'(n) + \mathbf{K}(n)\mathbf{C}(n)\mathbf{A}(n) = \mathbf{A}(n)$$

- Thus $\hat{\mathbf{x}}(n)$ is

$$\hat{\mathbf{x}}(n) = \hat{\mathbf{x}}(n|n-1) + \mathbf{K}(n)\left[\mathbf{y}(n) - \mathbf{C}(n)\hat{\mathbf{x}}(n|n-1)\right]$$

- Next step is to compute $E\left\{\mathbf{e}(n)\mathbf{e}(n)^{\mathrm{T}}\right\}$. By subtracting $\mathbf{x}(n)$ from both sides of the previous equation, the error is

$$\mathbf{e}(n) = [\mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)]\,\mathbf{e}(n|n-1) - \mathbf{K}(n)\mathbf{v}(n)$$

- Expand the error $\mathbf{e}(n)\mathbf{e}(n)^{\mathrm{T}}$ to get

$$\begin{aligned}\mathbf{e}(n)\mathbf{e}(n)^{\mathrm{T}} = {} & [\mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)]\,\mathbf{e}(n|n-1) \\ & \times \mathbf{e}(n|n-1)^{\mathrm{T}}\,[\mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)]^{\mathrm{T}} \\ & + \mathbf{K}(n)\mathbf{v}(n)\mathbf{v}(n)^{\mathrm{T}}\mathbf{K}(n)^{\mathrm{T}} \\ & + (\text{terms involving one } \mathbf{v}(n))\end{aligned}$$

- Take the expectation to get

$$\begin{aligned}\mathbf{P}(n) = {} & [\mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)]\,\mathbf{P}(n|n-1)\,[\mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)]^{\mathrm{T}} \\ & + \mathbf{K}(n)\mathbf{Q}_v(n)\mathbf{K}(n)^{\mathrm{T}}\end{aligned}$$

- Now separate $\mathbf{K}(n)$ terms, take the trace and differentiate

$$\begin{aligned}\frac{d}{d\mathbf{K}(n)}\mathrm{tr}(\mathbf{P}(n)) = {} & -2\,[\mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)]\,\mathbf{P}(n|n-1)\mathbf{C}(n)^{\mathrm{T}} \\ & + 2\mathbf{K}(n)\mathbf{Q}_v(n)\end{aligned}$$

- Set the derivative to 0 and solve for $\mathbf{K}(n)$

$$[\mathbf{I}-\mathbf{K}(n)\mathbf{C}(n)]\,\mathbf{P}(n|n-1)\mathbf{C}(n)^{\mathrm{T}} = \mathbf{K}(n)\mathbf{Q}_v(n)$$

$$\mathbf{P}(n|n-1)\mathbf{C}(n)^{\mathrm{T}} = \mathbf{K}(n)\mathbf{Q}_v(n) + \mathbf{K}(n)\mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}(n)^{\mathrm{T}}$$

$$\mathbf{K}(n) = \mathbf{P}(n|n-1)\mathbf{C}(n)^{\mathrm{T}}\left[\mathbf{Q}_v(n) + \mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}(n)^{\mathrm{T}}\right]^{-1}$$

- Having found $\mathbf{K}(n)$, simplify the above expression for $\mathbf{P}(n)$

$$\begin{aligned}\mathbf{P}(n) &= [\mathbf{I}-\mathbf{K}(n)\mathbf{C}(n)]\,\mathbf{P}(n|n-1)\,[\mathbf{I}-\mathbf{K}(n)\mathbf{C}(n)]^{\mathrm{T}} \\ &\quad + \mathbf{K}(n)\mathbf{Q}_v(n)\mathbf{K}(n)^{\mathrm{T}} \\ &= [\mathbf{I}-\mathbf{K}(n)\mathbf{C}(n)]\,\mathbf{P}(n|n-1) \\ &\quad - [\mathbf{I}-\mathbf{K}(n)\mathbf{C}(n)]\,\mathbf{P}(n|n-1)\mathbf{C}(n)^{\mathrm{T}}\mathbf{K}(n)^{\mathrm{T}} \\ &\quad + \mathbf{K}(n)\mathbf{Q}_v(n)\mathbf{K}(n)^{\mathrm{T}} \\ \mathbf{P}(n) &= [\mathbf{I}-\mathbf{K}(n)\mathbf{C}(n)]\,\mathbf{P}(n|n-1)\end{aligned}$$

7 Summary of Kalman filter equations

$$\begin{array}{ll} \textit{State Equation} & \mathbf{x}(n) = \mathbf{A}(n)\mathbf{x}(n-1) + \mathbf{w}(n) \\ \textit{Observation Equation} & \mathbf{y}(n) = \mathbf{C}(n)\mathbf{x}(n) + \mathbf{v}(n) \end{array}$$

and $\mathbf{x}(0)$ has mean $\mathbf{m}(0)$ and covariance matrix $\mathbf{P}(0)$

Initialization: $\hat{\mathbf{x}}(0) = \mathbf{m}(0)$ and $\mathbf{P}(0)$

Computation: for $n = 1, 2, \ldots$

Prediction step:

$$\hat{\mathbf{x}}(n|n-1) = \mathbf{A}(n)\hat{\mathbf{x}}(n-1)$$

$$\mathbf{P}(n|n-1) = \mathbf{A}(n)\mathbf{P}(n-1)\mathbf{A}(n)^{\mathrm{T}} + \mathbf{Q}_w(n)$$

Gain calculation:

$$\mathbf{K}(n) = \mathbf{P}(n|n-1)$$
$$\times \, \mathbf{C}(n)^{\mathrm{T}} \left[\mathbf{Q}_v(n) + \mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}(n)^{\mathrm{T}}\right]^{-1}$$

Update step:

$$\hat{\mathbf{x}}(n) = \hat{\mathbf{x}}(n|n-1) + \mathbf{K}(n)\left[\mathbf{y}(n) - \mathbf{C}(n)\hat{\mathbf{x}}(n|n-1)\right]$$
$$\mathbf{P}(n) = \left[\mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)\right]\mathbf{P}(n|n-1)$$

Apply the Kalman filter to the following model:

$$x(n+1) = x(n) + w(n+1)$$
$$y(n) = x(n) + v(n)$$

where $\{w(n)\}$ and $\{v(n)\}$ independent zero-mean random variables satisfying

$$E\left(v(n)v(k)\right) = \sigma_v^2 \delta_{n,k},$$
$$E\left(w(n)w(k)\right) = \sigma_w^2 \delta_{n,k},$$

Also $E\left(x(0)^2\right) = \sigma(0)^2$, $E\left(x(0)\right) = 0$

Prediction step:

$$\hat{x}(n|n-1) = \hat{x}(n-1)$$
$$\sigma(n|n-1)^2 = \sigma(n-1)^2 + \sigma_w^2$$

Gain calculation:

$$K(n) = \frac{\sigma(n-1)^2 + \sigma_w^2}{\sigma_v^2 + \sigma(n-1)^2 + \sigma_w^2}$$

Update step:

$$\hat{x}(n) = \hat{x}(n-1) + K(n)\left[y(n) - \hat{x}(n-1)\right]$$

$$\sigma(n)^2 = [1 - K(n)]\left(\sigma(n-1)^2 + \sigma_w^2\right)$$