13.4 Scalar Kalman Filter

Data Model

To derive the Kalman filter we need the data model:

$$s[n] = as[n-1] + u[n] \quad \text{< State Equation >}$$

$$x[n] = s[n] + w[n] \quad \text{< Observation Equation >}$$

Assumptions

1. $u[n]$ is zero mean Gaussian, White, $E\{u^2[n]\} = \sigma_u^2$
2. $w[n]$ is zero mean Gaussian, White, $E\{w^2[n]\} = \sigma_n^2$ (Can vary with time)
3. The initial state is $s[-1] \sim N(\mu_s, \sigma_s^2)$
4. $u[n]$, $w[n]$, and $s[-1]$ are all independent of each other

To simplify the derivation: let $\mu_s = 0$ (we'll account for this later)

Goal and Two Properties

Goal: **Recursively** compute $\hat{s}[n \mid n] = E\{s[n] \mid x[0], x[1], \ldots, x[n]\}$

$$\mathbf{X}[n] = [x[0], x[1], \ldots, x[n]]^T$$

Notation:
$\mathbf{X}[n]$ is set of all observations
$\mathbf{x}[n]$ is a single vector-observation

Two Properties We Need

1. For the jointly Gaussian case, the MMSE estimator of zero mean based on two uncorrelated data vectors $\mathbf{x}_1$ & $\mathbf{x}_2$ is (see p. 350 of text)

$$\hat{\theta} = E\{\theta \mid \mathbf{x}_1, \mathbf{x}_2\} = E\{\theta \mid \mathbf{x}_1\} + E\{\theta \mid \mathbf{x}_2\}$$

2. If $\theta = \theta_1 + \theta_2$ then the MSEE estimator is

$$\hat{\theta} = E\{\theta \mid \mathbf{x}\} = E\{\theta_1 + \theta_2 \mid \mathbf{x}\} = E\{\theta_1 \mid \mathbf{x}\} + E\{\theta_2 \mid \mathbf{x}\}$$

(a result of the linearity of $E\{.\}$ operator)

Derivation of Scalar Kalman Filter

Recall from Section 12.6… Innovation: $\tilde{x}[n] = x[n] - \underbrace{\hat{x}[n \mid n-1]}$

MMSE estimate of $x[n]$ given $\mathbf{X}[n-1]$ (prediction!!)

By MMSE Orthogonality Principle

$$E\left\{ \tilde{x}[n]\mathbf{X}[n-1] \right\} = \mathbf{0}$$

$\tilde{x}[n]$ is part of $x[n]$ that is uncorrelated with the previous data

Now note: $\mathbf{X}[n]$ is equivalent to $\left\{ \mathbf{X}[n-1], \tilde{x}[n] \right\}$

Why? Because we can get get $\mathbf{X}[n]$ from it as follows:

$$\begin{bmatrix} \mathbf{X}[n-1] \\ \tilde{x}[n] \end{bmatrix} \rightarrow \begin{bmatrix} \mathbf{X}[n-1] \\ x[n] \end{bmatrix} = \mathbf{X}[n]$$

$$x[n] = \tilde{x}[n] + \underbrace{\sum_{k=0}^{n-1} a_k x[k]}_{\hat{x}[n|n-1]}$$

What have we done so far?

- Have shown that $\mathbf{X}[n] \leftrightarrow \underbrace{\left\{ \mathbf{X}[n-1], \tilde{x}[n] \right\}}_{\text{uncorrelated}}$

$\Rightarrow$ Have split current data set into 2 parts:

1. Old data
2. Uncorrelated part of new data ("just the new facts")

$$\Rightarrow \hat{s}[n\,|\,n] = E\left\{ s[n]\,|\,\mathbf{X}[n] \right\} = E\left\{ s[n]\,|\,\mathbf{X}[n-1], \tilde{x}[n] \right\}$$

Because of this

So what??!! Well… can now exploit Property #1!!

$$\Rightarrow \hat{s}[n\,|\,n] = \underbrace{E\left\{ s[n]\,|\,\mathbf{X}[n-1] \right\}}_{\stackrel{\Delta}{=} \hat{s}[n\,|\,n-1]} + \underbrace{E\left\{ s[n]\,|\,\tilde{x}[n] \right\}}$$

prediction of $s[n]$ based on past data

Update based on innovation part of new data

Now need to look more closely at each of these!

Look at Prediction Term: $\hat{s}[n \mid n-1]$

Use the Dynamical Model… it is the key to prediction because it tells us how the state should progress from instant to instant

$$\hat{s}[n \mid n-1] = E\left\{s[n] \mid \mathbf{X}[n-1]\right\} = E\left\{\underbrace{as[n-1]+u[n]}\mid \mathbf{X}[n-1]\right\}$$

Now use Property #2:

$$\hat{s}[n \mid n-1] = a\underbrace{E\left\{s[n-1] \mid \mathbf{X}[n-1]\right\}}_{=\hat{s}[n-1|n-1]} + \underbrace{E\left\{u[n] \mid \mathbf{X}[n-1]\right\}}_{=E\{u[n]\}=0}$$

By Definition

By independence of $u[n]$ & $\mathbf{X}[n\text{-}1]$… See bottom of p. 433 in textbook.

$$\hat{s}[n \mid n-1] = a\hat{s}[n-1 \mid n-1]$$

The Dynamical Model provides the update from estimate to prediction!!

Look at Update Term: $E\{s[n] \mid \tilde{x}[n]\}$

Use the form for the Gaussian MMSE estimate:

$$E\{s[n] \mid \tilde{x}[n]\} = \underbrace{\left[\frac{E\{s[n]\tilde{x}[n]\}}{E\{\tilde{x}^2[n]\}}\right]}_{\triangleq k[n]} \tilde{x}[n]$$

$$\tilde{x}[n] = x[n] - \hat{x}[n \mid n-1]$$

So… $E\{s[n] \mid \tilde{x}[n]\} = k[n]\big(x[n] - \hat{x}[n \mid n-1]\big)$

by Prop. #2 $= \underbrace{\hat{s}[n \mid n-1]} + \underbrace{\hat{w}[n \mid n-1]}_{=0}$

Prediction Shows Up Again!!!

Because $w[n]$ is indep. of $\{x[0], \dots, x[n-1]\}$

Put these Results Together:

$$\hat{s}[n \mid n] = \underbrace{\hat{s}[n \mid n-1]}_{= a\hat{s}[n-1 \mid n-1]} + k[n]\big[x[n] + \hat{s}[n \mid n-1]\big]$$

This is the Kalman Filter

How to get the gain?

Look at the Gain Term:

Need two properties…

A. $E\{s[n](x[n]-\hat{s}[n|n-1])\} = E\{(s[n]-\hat{s}[n|n-1])(x[n]-\hat{s}[n|n-1])\}$

Aside
$\langle \mathbf{x},\mathbf{y}\rangle = \langle \mathbf{x}+\mathbf{z},\mathbf{y}\rangle$
for any $\mathbf{z} \perp \mathbf{y}$

Linear combo of past data… thus ⊥ w/ innovation

$= x[n]-\hat{x}[n|n-1]$

$= \tilde{x}[n]$

The innovation

B. $E\{w[n](s[n]-\hat{s}[n|n-1])\} = 0$

" proof "

- $w[n]$ is the measurement noise and by assumption is indep. of the "dynamical driving noise" $u[n]$ and $s[-1]$… In other words: $w[n]$ is indep. of everything dynamical… So $E\{w[n]s[n]\} = 0$

- $\hat{s}[n|n-1]$ is based on past data, which include $\{w[0], \ldots, w[n-1]\}$, and since the measurement noise has indep. samples we get $\hat{s}[n|n-1] \perp w[n]$

So… we start with the gain as defined above:

Plug in for innovation

(★)

$$k[n] = \frac{E\{s[n]\tilde{x}[n]\}}{E\{\tilde{x}^2[n]\}} = \frac{E\left\{s[n]\left[x[n] - \hat{s}[n|n-1]\right]\right\}}{E\left\{\left[x[n] - \hat{s}[n|n-1]\right]^2\right\}}$$

Use Prop. A in num.
Use $x[n] = s[n] + w[n]$
in denominator

$$= \frac{E\left\{\left[s[n] - \hat{s}[n|n-1]\right]\left[x[n] - \hat{s}[n|n-1]\right]\right\}}{E\left\{\left[s[n] - \hat{s}[n|n-1] + w[n]\right]^2\right\}}$$

(★★)

Use
$x[n] = s[n] + w[n]$
in numerator

$$= \frac{E\left\{\left[s[n] - \hat{s}[n|n-1]\right]\left[s[n] - \hat{s}[n|n-1] + w[n]\right]\right\}}{E\left\{\left[s[n] - \hat{s}[n|n-1] + w[n]\right]^2\right\}}$$

Expand

$$= \frac{E\left\{\left[s[n] - \hat{s}[n|n-1]\right]^2\right\} + E\left\{\left[s[n] - \hat{s}[n|n-1]\right]w[n]\right\}}{E\left\{\left[s[n] - \hat{s}[n|n-1]\right]^2\right\} + \sigma_n^2 + 2E\left\{\left[s[n] - \hat{s}[n|n-1]\right]w[n]\right\}}$$

$E\left\{\left[s[n] - \hat{s}[n|n-1]\right]^2\right\} \triangleq M[n|n-1]$

= 0 by Prop. B

MSE when $s[n]$ is estimated by 1-step prediction

This gives a form for the gain:

$$k[n] = \frac{M[n|n-1]}{\sigma_n^2 + M[n|n-1]}$$

This balances…

- the quality of the measured data
- against the predicted state

In the Kalman filter the prediction acts like the prior information about the state at time *n* before we observe the data at time *n*

Look at the Prediction MSE Term:

But now we need to know how to find $M[n|n-1]$!!!

$$M[n|n-1] = E\left\{\left[s[n] - \hat{s}[n|n-1]\right]^2\right\}$$

$$= E\left\{\left[as[n-1] + u[n] - a\hat{s}[n-1|n-1]\right]^2\right\}$$

$$= E\left\{\left[a\left(s[n-1] - \hat{s}[n-1|n-1]\right) + u[n]\right]^2\right\}$$

Use dynamical model & exploit form for prediction

Cross-terms = 0

Est. Error at previous time

$$\Rightarrow \quad M[n|n-1] = a^2 M[n-1|n-1] + \sigma_u^2$$

Why are the cross-terms zero? Two parts:

1. $s[n-1]$ depends on $\{u[0] \ldots u[n-1], s[-1]\}$, which are indep. of $u[n]$
2. $\hat{s}[n-1|n-1]$ depends on $\{s[0]+w[0] \ldots s[n-1]+w[n-1]\}$, which are indep. of $u[n]$

Look at a Recursion for MSE Term: $M[n|n]$

By def.: $M[n|n] = E\left\{[s[n]-\hat{s}[n|n]]^2\right\} = E\left\{[\underbrace{s[n]-\hat{s}[n|n-1]}_{\text{Term A}} - \underbrace{k[n](x[n]-\hat{s}[n|n-1])}_{\text{Term B}}]^2\right\}$

Now we'll get three terms:

$E\{A^2\}$, $E\{AB\}$, $E\{B^2\}$

$E\left\{A^2\right\} = M[n|n-1]$ by definition

$2E\left\{AB\right\} = -2k[n]\underbrace{E\left\{[s[n]-\hat{s}[n|n-1]][x[n]-\hat{s}[n|n-1]]\right\}}$

$= -2k[n]M[n|n-1]$ from (★★)… is num. $k[n]$

$E\left\{B^2\right\} = k^2[n]E\left\{\underbrace{[x[n]-\hat{s}[n|n-1]]^2}\right\}$

$= k^2[n][\text{Den. of } k[n]]$ from (★)… is den. $k[n]$

$= k[n][\text{Num. of } k[n]] = k[n]M[n|n-1]$

Recall: $k[n] = \dfrac{M[n|n-1]}{\sigma_n^2 + M[n|n-1]}$

So this gives…

$$M[n|n] = M[n|n-1] - 2k[n]M[n|n-1] + k[n]M[n|n-1]$$

$$\Rightarrow \quad M[n|n] = \big(1 - k[n]\big)M[n|n-1]$$

Putting all of these results together gives some very simple equations to iterate… Called the Kalman Filter

We just derived the form for Scalar State & Scalar Observation.
On the next three charts we give the Kalman Filter equations for:

- Scalar State & Scalar Observation
- Vector State & Scalar Observation
- Vector State & Vector Observation

Kalman Filter: Scalar State & Scalar Observation

State Model: $s[n] = as[n-1] + u[n]$ $u[n]$ WGN; WSS; $\sim N(0, \sigma_u^2)$

Observation Model: $x[n] = s[n] + w[n]$ $w[n]$ WGN; $\sim N(0, \sigma_n^2)$

Varies with *n*

Must Know: μ_s, σ^2_s , a, σ^2_u, σ^2_n

Initialization:

$$\hat{s}[-1|-1] = E\{s[-1]\} = \mu_s$$

$$M[-1|-1] = E\{(s[-1]\} - \hat{s}[-1|-1])^2\} = \sigma_s^2$$

Prediction:

$$\hat{s}[n|n-1] = a\hat{s}[n-1|n-1]$$

Pred. MSE:

$$M[n|n-1] = a^2 M[n-1|n-1] + \sigma_u^2$$

Kalman Gain:

$$K[n] = \frac{M[n|n-1]}{\sigma_n^2 + M[n|n-1]}$$

Update:

$$\hat{s}[n|n] = \hat{s}[n|n-1] + K[n]\big(x[n] - \hat{s}[n|n-1]\big)$$

Est. MSE:

$$M[n|n] = \big(1 - K[n]\big)M[n|n-1]$$

Kalman Filter: Vector State & Scalar Observation

State Model: $\mathbf{s}[n] = \mathbf{A}\mathbf{s}[n-1] + \mathbf{B}\mathbf{u}[n] \quad \mathbf{s}\ p\times 1; \mathbf{A}\ p\times p; \mathbf{B}\ p\times r; \mathbf{u} \sim N(\mathbf{0},\mathbf{Q})\ r\times 1$

Observation Model: $x[n] = \mathbf{h}^T[n]\mathbf{s}[n] + w[n]; \quad \mathbf{h}^T[n]\ p\times 1$ $\quad w[n]$ WGN; $\sim N(0,\sigma_n^2)$

Must Know: $\boldsymbol{\mu}_s$, $\mathbf{C}_s$, **A**, **B**, **h**, **Q**, σ^2_n

Initialization:

$$\hat{\mathbf{s}}[-1|-1] = E\{\mathbf{s}[-1]\} = \boldsymbol{\mu}_s$$

$$\mathbf{M}[-1|-1] = E\left\{(\mathbf{s}[-1]\} - \hat{\mathbf{s}}[-1|-1])(\mathbf{s}[-1]\} - \hat{\mathbf{s}}[-1|-1])^T\right\} = \mathbf{C}_s$$

Prediction:

$$\hat{\mathbf{s}}[n|n-1] = \mathbf{A}\hat{\mathbf{s}}[n-1|n-1]$$

Pred. MSE ($p\times p$):

$$\mathbf{M}[n|n-1] = \mathbf{A}\mathbf{M}[n-1|n-1]\mathbf{A}^T + \mathbf{B}\mathbf{Q}\mathbf{B}^T$$

Kalman Gain ($p\times 1$):

$$\mathbf{K}[n] = \frac{\mathbf{M}[n|n-1]\mathbf{h}[n]}{\sigma_n^2 + \underbrace{\mathbf{h}^T[n]\mathbf{M}[n|n-1]\mathbf{h}[n]}_{1\times 1}}$$

Update:

$$\hat{\mathbf{s}}[n|n] = \hat{\mathbf{s}}[n|n-1] + \mathbf{K}[n]\underbrace{(x[n] - \underbrace{\mathbf{h}^T[n]\hat{\mathbf{s}}[n|n-1]}_{\hat{x}[n|n-1]})}_{\tilde{x}[n]:\ \text{innovations}}$$

Est. MSE ($p\times p$):

$$\mathbf{M}[n|n] = \left(\mathbf{I} - \mathbf{K}[n]\mathbf{h}^T[n]\right)\mathbf{M}[n|n-1]$$

Kalman Filter: Vector State & Vector Observation

State Model: $\mathbf{s}[n] = \mathbf{A}\mathbf{s}[n-1] + \mathbf{B}\mathbf{u}[n] \quad \mathbf{s}\ p\times 1; \mathbf{A}\ p\times p; \mathbf{B}\ p\times r; \mathbf{u} \sim \mathrm{N}(\mathbf{0},\mathbf{Q})\ r\times 1$

Observation: $\mathbf{x}[n] = \mathbf{H}[n]\mathbf{s}[n] + \mathbf{w}[n]; \quad \mathbf{x}\ M\times 1; \mathbf{H}[n]\ M\times p; \mathbf{w}[n] \sim \mathrm{N}(\mathbf{0},\mathbf{C}[n])\ M\times 1$

Must Know: μ_s, $\mathbf{C}_s$, **A**, **B**, **H**, **Q**, $\mathbf{C}[n]$}

Initialization:

$$\hat{\mathbf{s}}[-1|-1] = E\{\mathbf{s}[-1]\} = \boldsymbol{\mu}_s$$

$$\mathbf{M}[-1|-1] = E\left\{(\mathbf{s}[-1]\} - \hat{\mathbf{s}}[-1|-1])(\mathbf{s}[-1]\} - \hat{\mathbf{s}}[-1|-1])^T\right\} = \mathbf{C}_s$$

Prediction:

$$\hat{\mathbf{s}}[n|n-1] = \mathbf{A}\hat{\mathbf{s}}[n-1|n-1]$$

Pred. MSE ($p\times p$):

$$\mathbf{M}[n|n-1] = \mathbf{A}\mathbf{M}[n-1|n-1]\mathbf{A}^T + \mathbf{B}\mathbf{Q}\mathbf{B}^T$$

Kalman Gain ($p\times M$):

$$\mathbf{K}[n] = \mathbf{M}[n|n-1]\mathbf{H}^T[n]\left(\mathbf{C}[n] + \underbrace{\mathbf{H}[n]\mathbf{M}[n|n-1]\mathbf{H}^T[n]}_{M\times M}\right)^{-1}$$

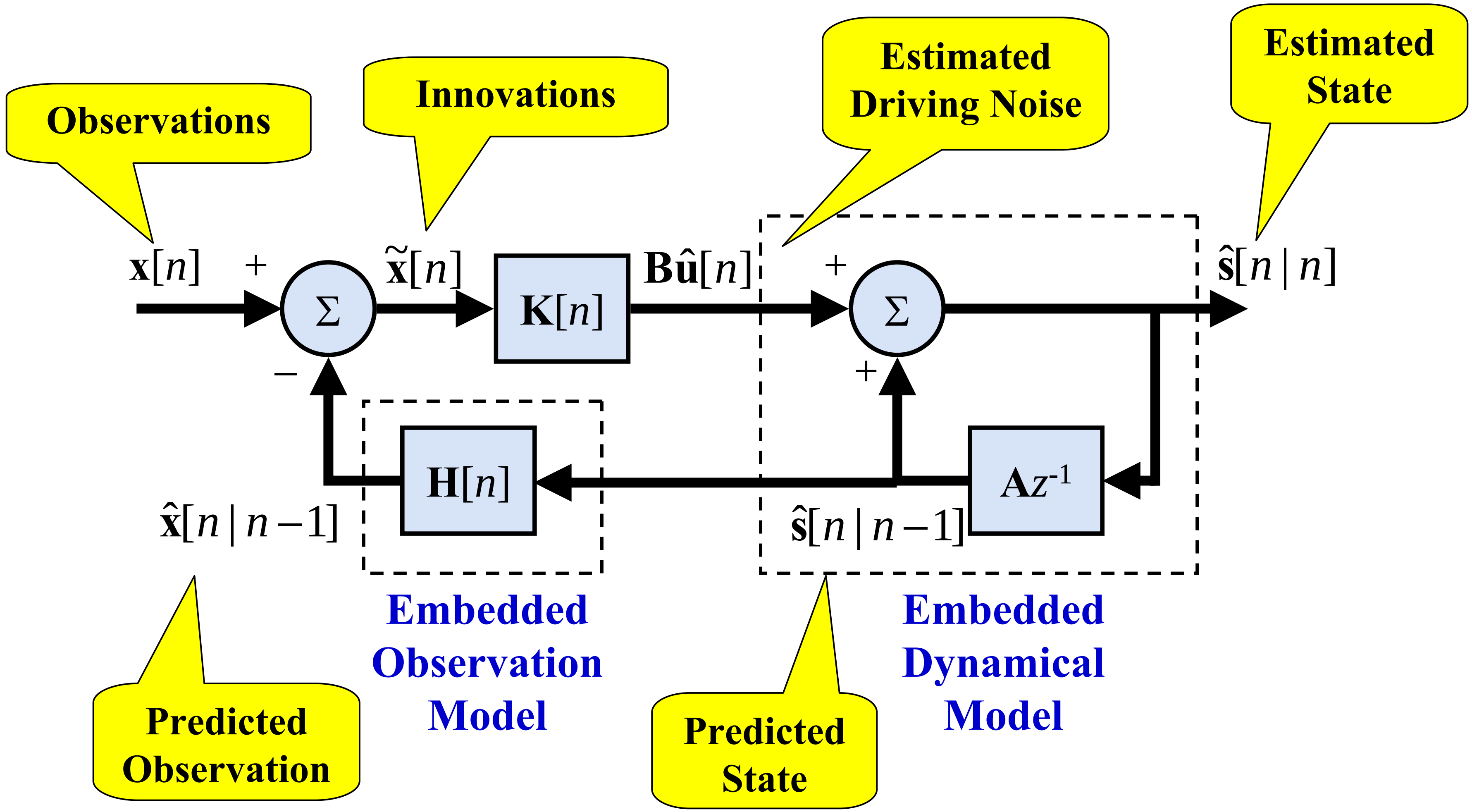
Update:

$$\hat{\mathbf{s}}[n|n] = \hat{\mathbf{s}}[n|n-1] + \mathbf{K}[n]\underbrace{(\mathbf{x}[n] - \underbrace{\mathbf{H}[n]\hat{\mathbf{s}}[n|n-1]}_{\hat{\mathbf{x}}[n|n-1]})}_{\tilde{\mathbf{x}}[n]:\ \textit{innovations}}$$

Est. MSE ($p\times p$): :

$$\mathbf{M}[n|n] = \big(\mathbf{I} - \mathbf{K}[n]\mathbf{H}[n]\big)\mathbf{M}[n|n-1]$$

Kalman Filter Block Diagram



Looks a lot like Sequential LS/MMSE except it has the Embedded Dynamical Model!!!

Overview of MMSE Estimation