KALMAN FILTERING

The Kalman filter is an important generalization of the Wiener filter.

Wiener filter	Kalman filter
• WSS process	• Gauss-Markov process
• Data stream goes back into the infinite past	• Data stream starts at a specific point in time
• scalar signals	• vector signals
• Nonadaptive	• Adaptive: model may evolve over time

Problem Statement

The goal is to estimate a time-varying <u>state vector</u> $\underline{s}(n) \in \mathbb{R}^p$ based on observations

$\underline{x}(0), \underline{x}(1), \ldots, \underline{x}(n) \in \mathbb{R}^M$

Linear Dynamical Model

We assume the following model:

$$\boxed{\begin{aligned} \underline{s}(n) &= A\,\underline{s}(n-1) + B\,\underline{u}(n) \\ \underline{x}(n) &= H\,\underline{s}(n) + \underline{w}(n) \end{aligned} \qquad n \geq 0}$$

where

A: $p \times p$ <u>state transition matrix</u>

$\underline{u}(n)$: $r \times 1$ zero-mean <u>excitation</u> or <u>driving noise</u> with covariance Q, uncorrelated for different n

B: $p \times r$ <u>state noise matrix</u>

H: $M \times 1$ <u>observation matrix</u>

$\underline{w}(n)$: $M \times 1$ <u>observation noise</u> with diagonal covariance matrix $R_w(n)$, uncorrelated for different n

$\boxed{\text{time varying}}$

We assume

- $A, B, R_w(n), H$ and Q are known
- A, B, H, Q may be time-varying also

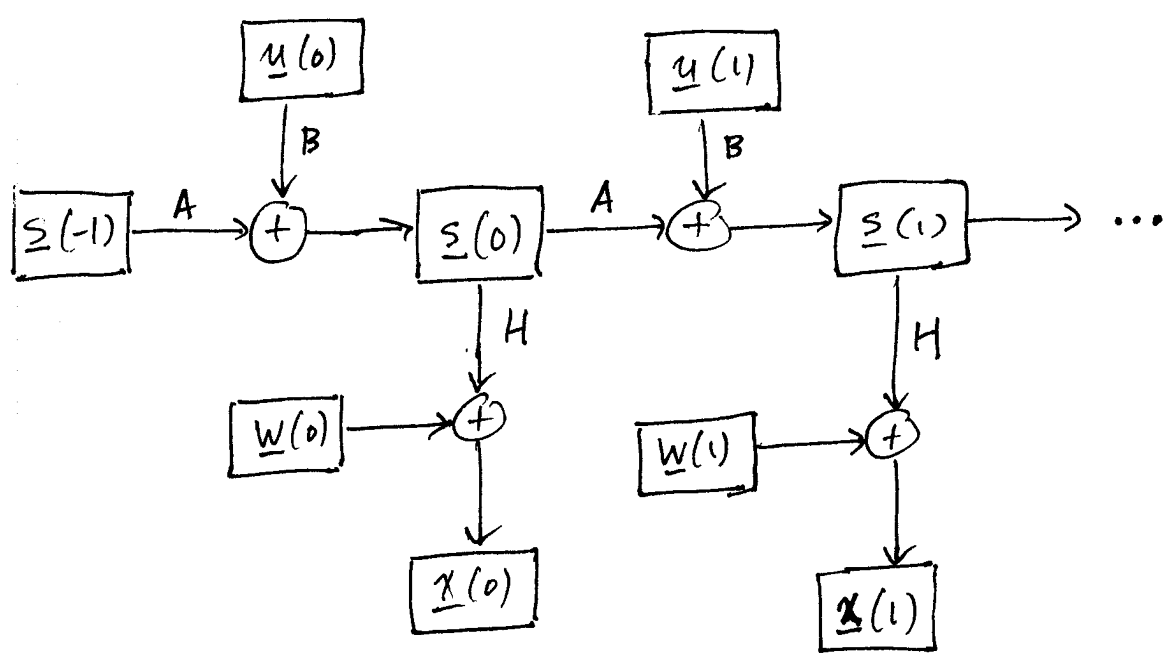
The Kalman filter is the LMMSE estimator of $\underline{s}(n)$ based on $\underline{X}(n) := \{ \underline{x}(0), \ldots, \underline{x}(n) \}$

The LMMSE estimator is a Bayesian estimator: we may think of $\underline{\theta} = \underline{s}(n)$ as the parameter and

state equation $\Longleftrightarrow$ prior
observation equation $\Longleftrightarrow$ likelihood

Initial Condition

Our linear dynamical model assumes the system starts at time $n=0$. Hence it is necessary to specify an initial condition $\underline{s}(-1)$, which may be fixed or random with mean $\underline{\mu}_0$, covariance Σ_0.



Example 1st order Gauss-Markov process

Suppose we observe a signal in noise

$$x(n) = s(n) + w(n) \qquad \text{(scalar equation)}$$

where $w(n) \overset{\text{iid}}{\sim} N(0, \sigma_w^2)$.

Furthermore, assume the signal obeys

$$s(n) = a \cdot s(n-1) + u(n)$$

where $u(n) \overset{\text{iid}}{\sim} N(0, \sigma_u^2)$, with the initial condition

$$s(-1) \sim N(\mu_0, \sigma_0^2) \qquad \left\{ \begin{array}{l} \text{independent of} \\ \{u(n)\}. \end{array} \right.$$

We say $\{s(n)\}$ is a <u>first-order Gauss-Markov process.</u>

To write this in the linear dynamical model form,

$$\underline{s}(n) = [s(n)] \quad (1 \times 1) \qquad \underline{u}(n) = u(n)$$

$$A = [a] \quad (1 \times 1) \qquad B = [1] \quad (1 \times 1)$$

$$H = [1] \quad (1 \times 1) \qquad R_w(n) = [\sigma_w^2] \quad (1 \times 1)$$

$$Q = [\sigma_u^2] \quad (1 \times 1)$$

Exercise p^{th} order Gauss Markov process

Assume the same setup as before, but now

$\{s(n)\}$ is described by

$$s(n) = \sum_{k=1}^{p} a_k s(n-k) + u(n)$$

Express this problem in the linear dynamical model form.

Hint: For the state vector, take

$$\underline{s}(n) = \left[s(n-p+1) \ \cdots \ s(n) \right]^T$$

Be careful not to confuse the <u>state</u> $\underline{s}(n)$ with the <u>signal</u> $s(n)$

Solution

$$\underbrace{\begin{bmatrix} s(n-p+1) \\ s(n-p+2) \\ \vdots \\ s(n-1) \\ s(n) \end{bmatrix}}_{\underset{\sim}{s}(n)} = \underbrace{\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ a_p & a_{p-1} & a_{p-2} & \cdots & a_1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} s(n-p) \\ s(n-p+1) \\ \vdots \\ s(n-2) \\ s(n-1) \end{bmatrix}}_{\underset{\sim}{s}(n-1)} + \underbrace{\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}}_{B} \underbrace{u(n)}_{\underset{\sim}{u}(n)}$$

$$\underbrace{x(n)}_{\underset{\sim}{x}(n)} = \underbrace{\begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \end{bmatrix}}_{H} \underbrace{\begin{bmatrix} s(n-p+1) \\ s(n-p+2) \\ \vdots \\ s(n-1) \\ s(n) \end{bmatrix}}_{\underset{\sim}{s}(n)} + \underbrace{w(n)}_{\underset{\sim}{w}(n)}$$

<u>Note</u> For a p^{th} order G. M. process, the initial condition would be

$$\underset{\sim}{s}(-1) \sim N(\underset{\sim}{\mu}_0, \Sigma_0)$$

independent of $\{u(n)\}$

Example 1: Array processing

$$\underline{x}(n) = \begin{bmatrix} x_1(n) \\ \vdots \\ x_M(n) \end{bmatrix}$$

$x_i(n)$ = signal intensity at receiver i

$$\underline{x}(n) = H \cdot \underline{s}(n) + \underline{w}(n)$$

Distortion-free measurements

$$x_i(n) = s(n) + w_i(n)$$

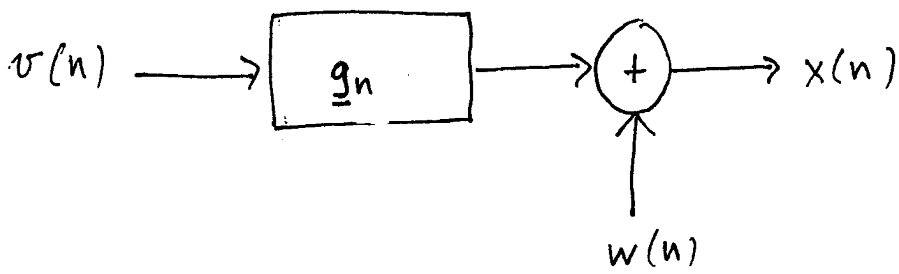
$$\Longleftrightarrow \quad H = [1 \ 1 \ \cdots \ 1]^T, \quad \underset{(1\times 1)}{\underline{s}(n) = [s(n)]}$$

Distorted measurements (attenuation / blurring)

$$x_i(n) = \sum_{k=0}^{p-1} h_i(k)\, s(n-k) + w_i(n)$$

$$\Longleftrightarrow \quad H = \underset{(M \times p)}{\begin{bmatrix} \underline{h}_1^T \\ \underline{h}_2^T \\ \vdots \\ \underline{h}_M^T \end{bmatrix}}, \quad \underline{s}(n) = \underset{(p \times 1)}{\begin{bmatrix} s(n) \\ s(n-1) \\ \vdots \\ s(n-p+1) \end{bmatrix}}$$

Example System Identification / Channel Estimation



$$x(n) = \sum_{k=0}^{p-1} g_n(k)\, v(n-k)$$

$v(n)$ = known signal (system probe), $\underline{g}_n$ unknown, time-varying impulse response

$$x(n) = H \cdot \underline{s}(n) + w(n)$$

(a)

$H =$

$\underline{s}(n) =$

The Kalman Recursions

The Kalman filter is given by

$$\hat{\underline{s}}(n) = R_{\underline{s}(n)\underline{X}(n)} R^{-1}_{\underline{X}(n)\underline{X}(n)} \underline{X}(n)$$

The covariance matrices can in principle be computed from the linear dynamical model, but there are two problems:

- these calculations are tedious
- we want an <u>on line</u> estimator, so that $\hat{\underline{s}}(n)$ can be efficiently updated from $\hat{\underline{s}}(n-1)$

<u>Approach</u>

There are several different ways to derive the Kalman recursions:

- innovations approach
- repeated use of the orthogonality principle (see Kay or Scharf)
- assume joint Gaussianity and apply properties of jointly Gaussian random vectors.

Joint Gaussianity

Since the LMMSE depends only on first and second order moments, we can assume the higher order moments are such that everything is jointly Gaussian.

For our linear dynamical model, this means

$$\left.\begin{array}{l} \underline{u}(n) \sim N(\underline{0}, Q) \\[2ex] \underline{w}(n) \sim N(\underline{0}, R_w(n)) \\[2ex] \underline{s}(-1) \sim N(\underline{\mu}_0, \Sigma_0) \end{array}\right\} \text{independent}$$

With these assumptions, the state variable

$$\underline{s}(n) = A\,\underline{s}(n-1) + B\,\underline{u}(n)$$

is said to be a <u>vector Gauss-Markov process</u>.

When deriving the Kalman filter, we will assume

$E[\underline{s}(n)] = \underline{0}$, $E[\underline{x}(n)] = \underline{0}$. The equations we derive will also be valid for non-zero means.

This can be verified by applying the Kalman filter to $\underline{s}'(n) = \underline{s}(n) - E[\underline{s}(n)]$ and $\underline{x}'(n) = \underline{x}(n) - E[\underline{x}(n)]$, and invoking linearity properties.

Notation

MMSE estimator (our goal)

$$s(n) \mid \underline{X}(n) \sim \mathcal{N}(\hat{\underline{s}}(n|n), M(n|n))$$

$$s(n) \mid \underline{X}(n-1) \sim \mathcal{N}(\hat{\underline{s}}(n|n-1), M(n|n-1))$$

MMSE predictor

$$\underline{e}(n) = \underline{s}(n) - \hat{\underline{s}}(n|n-1)$$ "prediction error"

$$R(n) = E[\underline{s}(n) \cdot \underline{s}(n)^T]$$ "state covariance"

Kalman Filter Derivation

Step 1a: $\underline{s}(n) \mid \underline{X}(n-1) \sim \underline{s}(n) \mid \hat{\underline{s}}(n|n-1)$

This follows from a more general fact:

If $(\underline{\theta}, \underline{y})$ are jointly Gaussian, zero mean, then

$$\underline{\theta} \mid \underline{y} \sim \underline{\theta} \mid E[\underline{\theta} \mid \underline{y}].$$

This seems intuitive because

$$\underline{\theta} \mid \underline{y} \sim \mathcal{N}\left(E[\underline{\theta} \mid \underline{y}],\ Cov(\underline{\theta} \mid \underline{y}) \right)$$

dependence on $\underline{y}$ manifested in posterior mean

independent of the actual value of $\underline{y}$

Let's prove it rigorously. Recall

$$E[\underline{\theta} \mid \underline{y}] = R_{\theta y} R_{yy}^{-1} \mathbf{y}$$

$$Cov(\underline{\theta} \mid \underline{y}) = R_{\theta\theta} - R_{\theta y} R_{yy}^{-1} R_{y\theta}$$

Note that $E[\underline{\theta} \mid \underline{y}]$ is a linear transformation of a Gaussian, so it is also Gaussian.

We have

$$\begin{bmatrix} \underline{\theta} \\ E[\underline{\theta}|\underline{Y}] \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & R_{\theta Y} R_{YY}^{-1} \end{bmatrix} \begin{bmatrix} \underline{\theta} \\ \underline{Y} \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} \underline{\theta} \\ E[\underline{\theta}|\underline{Y}] \end{bmatrix} \sim N\left(\underline{0}, \begin{bmatrix} R_{\theta\theta} & R_{\theta Y} R_{YY}^{-1} R_{Y\theta} \\ R_{\theta Y} R_{YY}^{-1} R_{Y\theta} & R_{\theta Y} R_{YY}^{-1} R_{Y\theta} \end{bmatrix} \right)$$

$$\Rightarrow \underline{\theta} \,|\, E[\underline{\theta}|\underline{Y}] \sim N\left(E[\underline{\theta}|\underline{Y}],\ R_{\theta\theta} - R_{\theta Y} R_{YY}^{-1} R_{Y\theta} \cdot \left(R_{\theta Y} R_{YY}^{-1} R_{Y\theta}\right)^{-1} R_{\theta Y} R_{YY}^{-1} R_{Y\theta} \right)$$

$$\sim N\left(E[\underline{\theta}|\underline{Y}],\ R_{\theta\theta} - R_{\theta Y} R_{YY}^{-1} R_{Y\theta} \right)$$

$$\sim \underline{\theta} \,|\, \underline{Y}$$

Now apply this result with $\underline{\theta} = \underline{s}(n)$, $\underline{Y} = \underline{X}(n-1)$,

and $E[\underline{\theta}|\underline{Y}] = \hat{\underline{s}}(n|n-1) \equiv E\left[\underline{s}(n) \,|\, \underline{X}(n-1)\right]$ $\blacksquare$

<u>Remark</u>

As a byproduct, we also showed

$$E\left[\hat{\underline{s}}(n|n-1)\, \underline{s}(n)^T\right] = R_{\theta Y} R_{YY}^{-1} R_{Y\theta}$$

$$= R_{\theta\theta} - \left(R_{\theta\theta} - R_{\theta Y} R_{YY}^{-1} R_{Y\theta} \right)$$

$$= R(n) - M(n|n-1)$$

Step 1b: $\quad \underline{s}(n) \mid \underline{X}(n) \quad \sim \quad \underline{s}(n) \mid \hat{\underline{s}}(n|n-1), \underline{x}(n)$

We just showed $\underline{s}(n) \mid \underline{X}(n-1)$ and $\underline{s}(n) \mid \hat{\underline{s}}(n|n-1)$ have the same distribution. Now condition both random variables on $\underline{x}(n)$ ∎

Conclusion from Step 1:

$$\hat{\underline{s}}(n|n) = E\left[\underline{s}(n) \mid \underline{X}(n)\right]$$

$$= E\left[\underline{s}(n) \mid \hat{\underline{s}}(n|n-1), \underline{x}(n)\right]$$

We will focus on computing

Step 2

$$\begin{bmatrix} \underline{s}(n) \\ \hat{\underline{s}}(n|n-1) \\ \underline{x}(n) \end{bmatrix} = \begin{bmatrix} I & I & 0 \\ 0 & I & 0 \\ H & H & I \end{bmatrix} \begin{bmatrix} \underline{e}(n) \\ \hat{\underline{s}}(n|n-1) \\ \underline{w}(n) \end{bmatrix}$$

because

$$\underline{s}(n) = \left(\underline{s}(n) - \hat{\underline{s}}(n|n-1)\right) + \hat{\underline{s}}(n|n-1)$$

$$= \underline{e}(n) + \hat{\underline{s}}(n|n-1)$$

Outline of rest of derivation:

Step 3: determine covariance of $\begin{bmatrix} \underline{e}(n) \\ \hat{\underline{s}}(n|n-1) \\ \underline{w}(n) \end{bmatrix}$

Step 4: determine covariance of $\begin{bmatrix} \underline{s}(n) \\ \hat{\underline{s}}(n|n-1) \\ \underline{x}(n) \end{bmatrix}$

Step 5: compute $\hat{\underline{s}}(n|n) = E\left[\underline{s}(n) \mid \hat{\underline{s}}(n|n-1), \underline{x}(n)\right]$

Step 6: identify recursive structure in

Step 3

$$E\left\{\begin{bmatrix}\underline{e}(n)\\ \hat{\underline{s}}(n|n-1)\\ \underline{w}(n)\end{bmatrix}\begin{bmatrix}\underline{e}(n)^T & \hat{\underline{s}}(n|n-1)^T & \underline{w}(n)^T\end{bmatrix}\right\} = \begin{bmatrix}M(n|n-1) & 0 & 0\\ 0 & R(n)-M(n|n-1) & 0\\ 0 & 0 & R_w(n)\end{bmatrix}$$

There are nine smaller covariances to compute.

The five involving $\underline{w}(n)$ are easy.

$$E\left[\underline{w}(n)\,\underline{w}(n)^T\right] = R_w(n) \quad \text{by definition of } R_w(n)$$

The other four are 0 because $\underline{e}(n)$ and $\hat{\underline{s}}(n|n-1)$ are linear combinations of $\{\underline{s}(-1), \underline{y}(0), \dots, \underline{y}(n), \underline{w}(0), \dots, \underline{w}(n-1)\}$, all of which are independent of $\underline{w}(n)$.

Next, lets show $E\left[\underline{e}(n)\,\hat{\underline{s}}(n|n-1)^T\right] = 0$.

Any guesses?

By the <u>orthogonality principle</u>, the LMMSE estimator is orthogonal to its prediction error. Equivalently, this follows from the Wiener-Hopf equations. If you're still not convinced, you can use properties of Gaussians to show

$$E\left[\underline{s}(n)\,\hat{\underline{s}}(n|n-1)^T\right] = E\left[\hat{\underline{s}}(n|n-1)\,\hat{\underline{s}}(n|n-1)^T\right]$$

using arguments like those in step 1a.

By symmetry, $E\left[\hat{\underline{s}}(n|n-1)\,\underline{e}(n)^T\right] = 0$, so only two terms remain.

Next, we have

$$E\left[\hat{\underline{s}}(n|n-1)\,\hat{\underline{s}}(n|n-1)^T\right] = E\left[\hat{\underline{s}}(n|n-1)\,\underline{s}(n)^T\right] \qquad \text{(from above)}$$

$$= R(n) - M(n|n-1) \qquad \boxed{\text{byproduct of step 1a}}$$

Finally,

$$E\left[\underline{e}(n)\,\underline{e}(n)^T\right] = E\left[\left(\underline{s}(n) - \hat{\underline{s}}(n|n-1)\right)\left(\underline{s}(n) - \hat{\underline{s}}(n|n-1)\right)^T\right]$$

$$= E\left[\left(\underline{s}(n) - \hat{\underline{s}}(n|n-1)\right)\underline{s}(n)^T\right] \qquad \text{(from above)}$$

$$= R(n) - \left(R(n) - M(n|n-1)\right) = M(n|n-1)$$

Step 4

$$E\left\{ \begin{bmatrix} \underline{s}(n) \\ \hat{\underline{s}}(n|n-1) \\ \underline{x}(n) \end{bmatrix} \begin{bmatrix} \underline{s}(n)^T & \hat{\underline{s}}(n|n-1)^T & \underline{x}(n)^T \end{bmatrix} \right\} =$$

$$\begin{bmatrix} I & I & 0 \\ 0 & I & 0 \\ H & H & I \end{bmatrix} \begin{bmatrix} M(n|n-1) & 0 & 0 \\ 0 & R(n)-M(n|n-1) & 0 \\ 0 & 0 & R_w(n) \end{bmatrix} \begin{bmatrix} I & 0 & H^T \\ I & I & H^T \\ 0 & 0 & I \end{bmatrix}$$

$$= \begin{bmatrix} R(n) & \hat{M}(n|n-1) & R(n)\underline{H}^T \\ \hat{M}(n|n-1) & \hat{M}(n|n-1) & \hat{M}(n|n-1)H^T \\ H R(n) & H\hat{M}(n|n-1) & H R(n)\underline{H}^T + R_w(n) \end{bmatrix}$$

$$\hat{M}(n|n-1) = R(n) - M(n|n-1)$$

$$= \left[\begin{array}{c:cc} R(n) & \hat{M}(n|n-1) & R(n)\underline{H}^T \\ \hdashline \hat{M}(n|n-1) & \multicolumn{2}{c}{\multirow{2}{*}{D(n)}} \\ H R(n) & & \end{array} \right]$$

Step 5 : Up to this point we have

$$\underline{s}(n) \mid \underline{X}(n) \sim \mathcal{N}\left(\hat{\underline{s}}(n|n),\ M(n|n) \right)$$

$$\sim \underline{s}(n) \,\Big|\, \hat{\underline{s}}(n|n-1),\ \underline{x}(n) \qquad \boxed{\text{Step 1.}}$$

and

$$\begin{bmatrix} \underline{s}(n) \\ \hat{\underline{s}}(n|n-1) \\ \underline{x}(n) \end{bmatrix} \sim \mathcal{N}\left(\underline{0},\ \left[\begin{array}{c:cc} R(n) & \hat{M}(n|n-1) & R(n)H^T \\ \hdashline \hat{M}(n|n-1) & \multicolumn{2}{c}{\multirow{2}{*}{$D(n)$}} \\ H\,R(n) & & \end{array}\right] \right) \quad \boxed{\text{Steps 2–4}}$$

Therefore

$$\hat{\underline{s}}(n|n) = \left[\hat{M}(n|n-1) \quad R(n)H^T \right] \cdot D(n)^{-1} \cdot \begin{bmatrix} \hat{\underline{s}}(n|n-1) \\ \underline{x}(n) \end{bmatrix}$$

$$M(n|n) = R(n) - \left[\hat{M}(n|n-1) \quad R(n)H^T \right] D(n)^{-1} \begin{bmatrix} \hat{M}(n|n-1) \\ H\,R(n) \end{bmatrix}$$

Let's compute $D(n)^{-1}$!

Due to the nice structure of $D(n)$,

$$D(n)^{-1} = \begin{bmatrix} \hat{M}(n|n-1)^{-1} & 0 \\ 0 & 0 \end{bmatrix} +$$

$$\begin{bmatrix} -H^T \\ I \end{bmatrix} \left[H M(n|n-1) H^T + R_w(n) \right]^{-1} \begin{bmatrix} -H & I \end{bmatrix}$$

You may verify this yourself.

Plugging in, we get

$$\hat{s}(n|n) = \hat{\underline{s}}(n|n-1) + K(n) \cdot \left(\underline{x}(n) - H \hat{\underline{s}}(n|n-1) \right)$$

$$M(n|n) = \left(I - K(n) H \right) M(n|n-1)$$

where

$$K(n) = M(n|n-1) \cdot H^T \cdot \left(H M(n|n-1) H^T + R_w(n) \right)^{-1}$$

is called the <u>Kalman gain</u>.

Step 6 Identify recursion.

We are almost there. All we lack are

$$\hat{\underline{s}}(n|n-1) = E\left[\underline{s}(n) \mid \underline{X}(n-1)\right]$$

$$= E\left[A \cdot \underline{s}(n-1) + B\underline{u}(n) \mid \underline{X}(n-1)\right]$$

$$= A\, E\left[\underline{s}(n-1) \mid \underline{X}(n-1)\right] + B\, E\left[\underline{u}(n) / \underline{X}(n-1)\right]$$

$$= A\, \hat{\underline{s}}(n-1|n-1) + \underline{0}.$$

and

$$M(n|n-1) = A\, M(n-1|n-1) A^T + B Q B^T$$

by a similar application of the definition of the vector Gauss-Markov process $\underline{s}(n)$

In particular

$$M(n|n-1) = E\left[\left(\underline{s}(n) - \hat{\underline{s}}(n|n-1)\right)\left(\underline{s}(n) - \hat{\underline{s}}(n|n-1)\right)^T \middle| \underline{X}(n-1)\right]$$

$$= E\left[\left(A\underline{s}(n-1) + B\underline{u}(n) - A\cdot\hat{\underline{s}}(n-1|n-1)\right)\cdot\right.$$

$$\left.\left(A\underline{s}(n-1) + B\underline{u}(n) - A\hat{\underline{s}}(n-1|n-1)\right)^T \middle| \underline{X}(n-1)\right]$$

$$= E\left\{\left(A\left[\underline{s}(n-1) - \hat{\underline{s}}(n-1|n-1)\right] + B\underline{u}(n)\right)\cdot\right.$$

$$\left.\left(A\left[\underline{s}(n-1) - \hat{\underline{s}}(n-1|n-1)\right] + B\underline{u}(n)\right)^T \middle| \underline{X}(n-1)\right\}$$

$$= A\cdot E\left[\left(\underline{s}(n-1) - \hat{\underline{s}}(n-1|n-1)\right)\cdot\left(\underline{s}(n-1) - \hat{\underline{s}}(n-1|n-1)\right)^T \middle| \underline{X}(n-1)\right]$$

$$+ B\cdot E\left[\underline{u}(n)\cdot\underline{u}(n)^T\right]B^T$$

independence assumption

$$= A\,M(n-1|n-1)A^T + BQB^T$$

Kalman Filter Equations

1. $\hat{\underline{s}}(n|n-1) = A\,\hat{\underline{s}}(n-1|n-1)$

2. $M(n|n-1) = A \cdot M(n-1|n-1)\,A^T + B Q B^T$.

3. $K(n) = M(n|n-1)\,H^T \left(H\,M(n|n-1)\,H^T + R_w(n) \right)^{-1}$

4. $\hat{\underline{s}}(n|n) = \hat{\underline{s}}(n|n-1) + K(n) \cdot \left(\underline{x}(n) - H\,\hat{\underline{s}}(n|n-1) \right)$

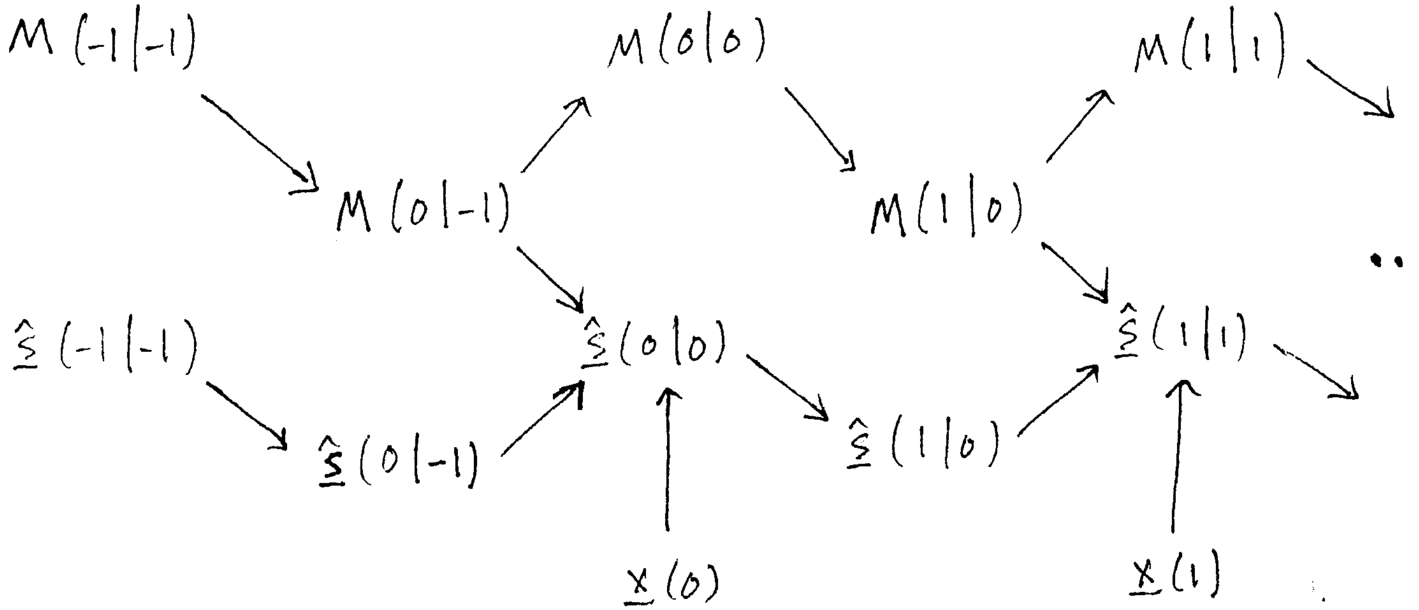
5. $M(n|n) = \left(I - K(n) H \right) M(n|n-1)$

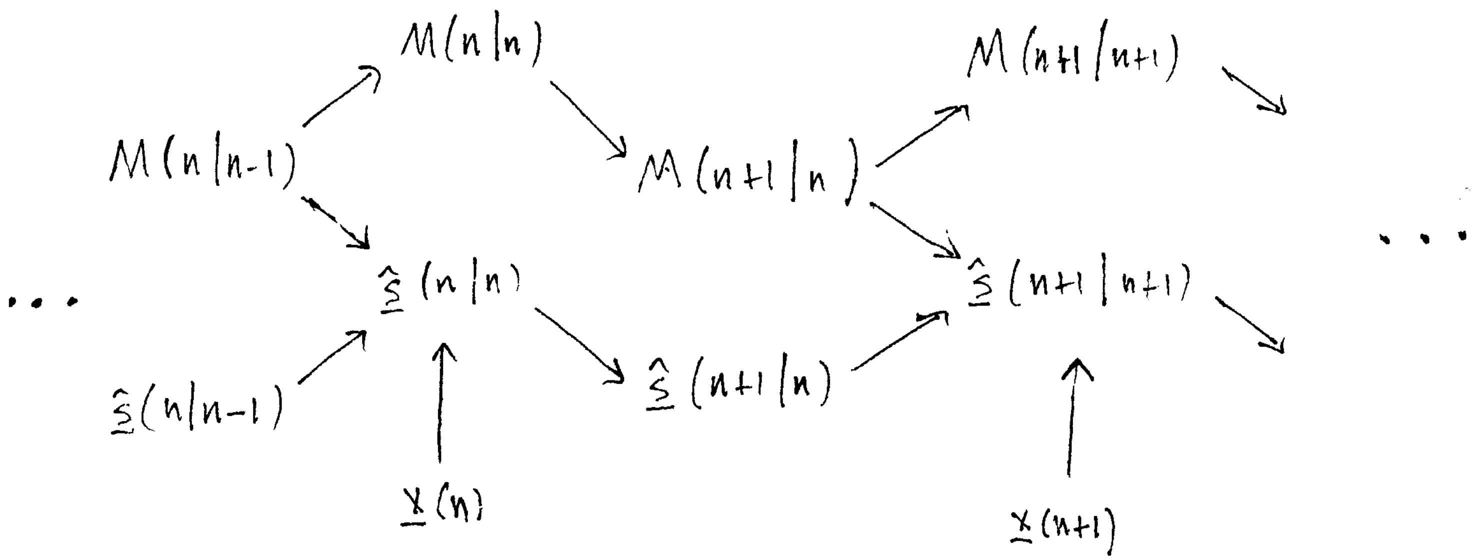
Initialization

$$\hat{\underline{s}}(-1|-1) = \underline{\mu}_0$$

$$\hat{M}(-1|-1) = \Sigma_0$$

Flow Diagram





Two stages:

I. **Update** predictor variables $\hat{\underline{s}}(n|n-1)$, $M(n|n-1)$

II. **Update** estimator variables $\hat{\underline{s}}(n|n)$, $M(n|n)$

Remarks

- The Kalman filter is the MMSE estimator when the data / state are jointly Gaussian.
- If not, the Kalman filter is still the LMMSE estimator.
- In practice, the state transition model or observation model may be innaccurate. The Kalman filter may still return a useful estimator.
- The Kalman recursions are still valid when we allow A, B, H, and Q to be time varying.

Key

a. $H = [v(n-p+1) \quad \cdots \quad v(n)]$, $\underline{s}(n) = \begin{bmatrix} g_n(0) \\ g_n(1) \\ \vdots \\ g_n(p-1) \end{bmatrix}$