I225E Statistical Signal Processing

14. Signal Processing II

MAWALIM and UNOKI

candylim@jaist.ac.jp and unoki@jaist.ac.jp

School of Information Science

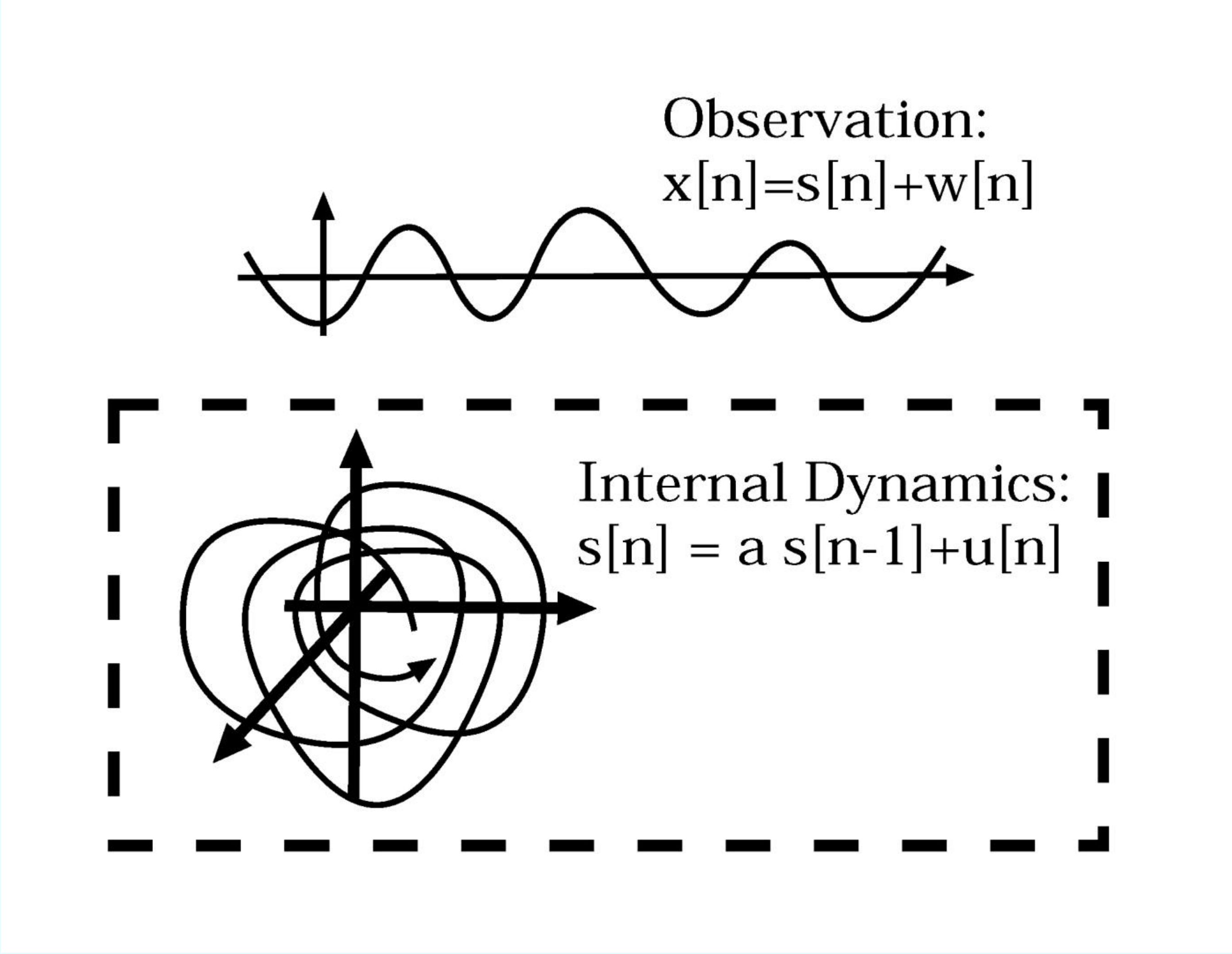
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY,
HOKURIKU 1990

Wiener filter and Kalman filter

Kalman filter is an important generalization of Wiener filter.

- ***Wiener Filter***
  - WSS (Wide-sense-stationary) Process
  - Data from infinite past
  - Scalar signals
  - Non-adaptive
- ***Kalman Filter***
  - Gauss-Markov Process
  - Data from a specific point in time
  - Vector signals
  - Adaptive (model may evolve over time)

Estimate internal system state from observed data



2. Scalar Kalman Filter

First-order Gauss-Markov model

$$x[n] = s[n] + w[n],$$
$$s[n] = as[n-1] + u[n].$$

From observed data $\boldsymbol{X}[n] = [x[0], x[1], \cdots, x[n]]^T$, estimate $s[n]$ $(n \geq 0)$

- Constant a is known ($|a| < 1$).
- $u[n] \sim N(0, \sigma_u^2)$, $w[n] \sim N(0, \sigma_w^2)$, $s[-1] \sim N(0, \sigma_s^2)$.
- $s[-1]$, $u[n]$, $w[n]$ are all independent from each other.
- Denote estimate of $s[n]$ based on $\boldsymbol{X}[m] = [x[0], x[1], \cdots, x[m]]^T$ by $\hat{s}[n|m]$.

Find estimator $\hat{s}[n|n]$ that minimizes mean square error

$$E[(s[n] - \hat{s}[n|n])^2].$$

Computational Procedure of Kalman Filter

- **Prediction**

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

- **Minimum prediction error**

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

- **Kalman gain**

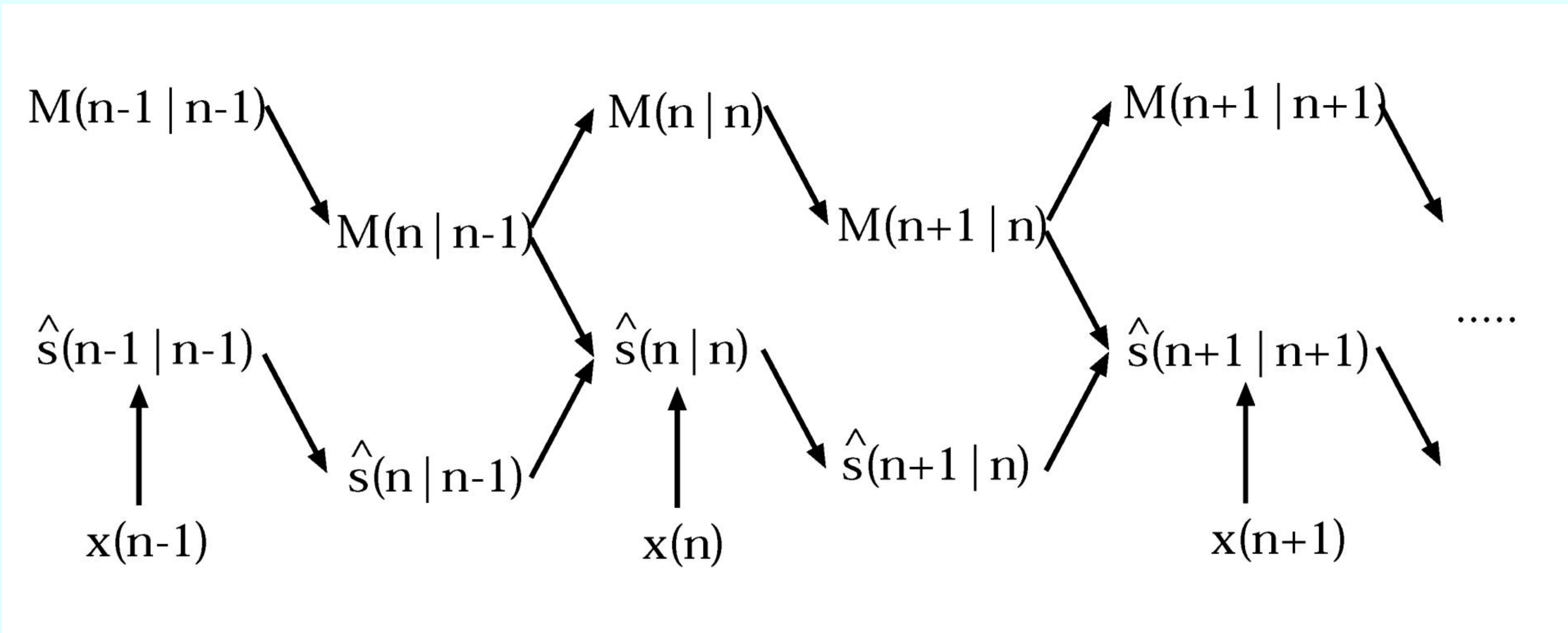
$$K[n] = \frac{M[n|n-1]}{\sigma_w^2 + M[n|n-1]}$$

- **Correction**

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

- **Minimum mean square error**

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


M(n-1 | n-1)
M(n | n-1)
M(n | n)
M(n+1 | n)
M(n+1 | n+1)
ŝ(n-1 | n-1)
ŝ(n | n-1)
ŝ(n | n)
ŝ(n+1 | n)
ŝ(n+1 | n+1)
.....
x(n-1)
x(n)
x(n+1)

3. Dynamical Model

Theorem: Gauss-Markov model

$$\boldsymbol{s}[n] = \boldsymbol{A}\boldsymbol{s}[n-1] + \boldsymbol{B}\boldsymbol{u}[n] \qquad n \geq 0.$$

- $\boldsymbol{A} \in \Re^{p \times p}$, $\boldsymbol{B} \in \Re^{p \times r}$

  Eigenvalues of $\boldsymbol{A}$ have amplitude smaller than 1.
- $\boldsymbol{s}[n] \in \Re^{p \times 1}$, $\boldsymbol{s}[-1] \sim N(\mu_s, \boldsymbol{C}_s)$
- $\boldsymbol{u}[n] \in \Re^{r \times 1}$, $\boldsymbol{u}[n] \sim N(0, \boldsymbol{Q})$
- $\boldsymbol{s}[-1]$, $\boldsymbol{u}[-1]$ are all independent from each other.

Then, mean and covariance of signal $\boldsymbol{s}[n]$ are

- **Mean**: $E\{\boldsymbol{s}[n]\} = \boldsymbol{A}^{n+1}\mu_s$
- **Covariance**: $\boldsymbol{C}_s[m,n] =$

$$E\left\{(\boldsymbol{s}[m] - E\{\boldsymbol{s}[m]\})\boldsymbol{s}[n] - E\{\boldsymbol{s}[n]\}\right\}$$

$(m \geq n)$ $\quad = \boldsymbol{A}^{m+1}\boldsymbol{C}_s(\boldsymbol{A}^{n+1})^T + \sum_{k=m-n}^{m} \boldsymbol{A}^k\boldsymbol{B}\boldsymbol{Q}\boldsymbol{B}^T\left(\boldsymbol{A}^{n-m+k}\right)^T$

$(m < n)$ $\quad \boldsymbol{C}_s[m,n] = \boldsymbol{C}_s^T[n,m]$

$(m = n)$ $\quad \boldsymbol{C}[n] = \boldsymbol{C}_s[n,n]$

$$= \boldsymbol{A}^{n+1}\boldsymbol{C}_s(\boldsymbol{A}^{n+1})^T + \sum_{k=0n} \boldsymbol{A}^k\boldsymbol{B}\boldsymbol{Q}\boldsymbol{B}^T\left(\boldsymbol{A}^{k}\right)^T$$

Time evolution of mean and covariance is

$$E\{\boldsymbol{s}[n]\} = \boldsymbol{A}E\{\boldsymbol{s}[n-1]\},$$
$$\boldsymbol{C}[n] = \boldsymbol{A}\boldsymbol{C}[n-1]\boldsymbol{A}^T + \boldsymbol{B}\boldsymbol{Q}\boldsymbol{B}^T.$$

4. Derivation of scalar Kalman filter

First-order Gauss-Markov model

$$x[n] = s[n] + w[n]$$
$$s[n] = as[n-1] + u[n]$$

From observed data $\boldsymbol{X}[n] = [x[0], x[1], \cdots, x[n]]^T$, estimate $s[n] (n \geq 0)$.

- Constant a is known ($|a| < 1$).
- $u[n] \sim N(0, \sigma_u^2)$, $w[n] \sim N(0, \sigma_w^2)$, $s[-1] \sim N(0, \sigma_s^2)$.
- $s[-1], u[n], w[n]$ are all independent from each other.
- Denote estimate of $s[n]$ based on $\boldsymbol{X}[m] = [x[0], x[1], \cdots, x[m]]^T$ by $\hat{s}[n|m]$.
- Denote error by $\tilde{x}[n] = x[n] - \hat{x}[n|n-1]$.

With respect to minimum mean square error (MMSE)

$$E\{(s[n] - \hat{s}[n|n])^2\},$$

Corresponding minimum mean square estimation is

$$\hat{s}[n|n] = E\{s[n]|\boldsymbol{X}[n]\}.$$

Basic properties of minimum mean square error estimator used for derivation:

- With respect to uncorrelated data x_1, x_2, minimum mean square error estimator $\hat{\theta}$ is (in case of $E\{\theta\} = 0$),

$$\hat{\theta} = E\{\theta|x_1, x_2\} = E\{\theta|x_1\} + E\{\theta|x_2\}.$$

- For $\theta = \theta_1 + \theta_2$, minimum mean square error estimator $\hat{\theta}$ is,

$$\hat{\theta} = E\{\theta|x\}$$
$$= E\{\theta_1 + \theta_2|x\}$$
$$= E\{\theta_1|x\} + E\{\theta_2|x\}.$$

$$\tilde{s}[n|n] = E\{s[n]|\boldsymbol{X}[n]\} = E\{s[n]|\boldsymbol{X}[n-1], x[n]\}$$
$$= E\{s[n]|\boldsymbol{X}[n-1], \tilde{x}[n] + \hat{x}[n]\}$$

because $\hat{x}[n]$ is represented by linear summation of $\{x[0], x[1], \cdots, x[n-1]\}$

$$= E\{s[n]|\boldsymbol{X}[n-1], \tilde{x}[n]\}$$
$$= E\{s[n]|\boldsymbol{X}[n-1]\} + E\{s[n]|\tilde{x}[n]\}.$$

$$
\begin{aligned}
\tilde{s}[n|n-1] &= E\{s[n]|\boldsymbol{X}[n-1]\} \\
&= E\{as[n-1]+u[n]|\boldsymbol{X}[n-1]\} \\
&= aE\{s[n-1]|\boldsymbol{X}[n-1]\,\}
\end{aligned}
$$

$$\text{because } E\{u[n]|X[n-1]\} = E\{u[n]\} = 0$$

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

$$\hat{s}[n|n] = \hat{s}[n|n-1] + E\{s[n]|\tilde{x}[n]\},$$

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

Since $E\{s[n]|\tilde{x}[n]\}$ is MMSE estimator of $s[n]$ based on $\tilde{x}[n]$,

$$
\begin{aligned}
E\{s[n]|\tilde{x}[n]\} &= K[n]\tilde{x}[n] \\
&= K[n](x[n] - \hat{x}[n|n-1])
\end{aligned}
$$

$$K[n] = \frac{E\{s[n]\tilde{x}[n]\}}{E\{\tilde{x}^2[n]\}}$$

$$
\begin{aligned}
\hat{x}[n|n-1] &= E\{x[n]|\boldsymbol{X}[n-1]\} \\
&= E\{s[n] + w[n]|\boldsymbol{X}[n-1]\} \\
&= E\{s[n]|\boldsymbol{X}[n-1]\} + E\{w[n]|\boldsymbol{X}[n-1]\} \\
&= \hat{s}[n|n-1] + \widehat{w}[n|n-1] \\
&= \hat{s}[n|n-1]. \ \text{(from } \widehat{w}[n|n-1] = 0\text{)}
\end{aligned}
$$

Summarizing above

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

Denominator and numerator of $K[n] = \frac{E\{s[n]\tilde{x}[n]\}}{E\{\tilde{x}^2[n]\}}$ are

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

from orthogonality principle, $E\{\hat{s}[n|n-1]\tilde{x}[n]\} = 0$

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

$$= E\{(s[n] - \hat{s}[n|n-1])(x[n] - \hat{s}[n|n-1])\}$$
$$= E\{(s[n] - \hat{s}[n|n-1])(s[n] + w[n] - \hat{x}[n|n-1])\}$$
$$\text{using } E\{(s[n] - \hat{s}[n|n-1])w[n]\} = 0$$
$$= E\{(s[n] - \hat{s}[n|n-1])(s[n] - \hat{s}[n|n-1])\}$$
$$= E\{(s[n] - \hat{s}[n|n-1])^2\}.$$

$$E\{\tilde{x}^2[n]\} = E\{(x[n] - \hat{x}[n|n-1])^2\}$$
$$= E\{(s[n] - \hat{s}[n|n-1] + w[n])^2\}$$
$$= \sigma_w^2 + E\{(s[n] - \hat{s}[n|n-1])^2\}$$

Hence
$$K[n] = \frac{E\left\{(s[n]-\hat{s}[n|n-1])^2\right\}}{\sigma_w^2+E\{(s[n]-\hat{s}[n|n-1])^2\}}$$
$$= \frac{M[n|n-1]}{\sigma_w^2+M[n|n-1]},$$
where
$$M[n|n-1] = E\{(s[n]-\hat{s}[n|n-1])^2\}.$$

$$M[n|n-1] = E\{(s[n]-\hat{s}[n|n-1])^2\}$$
$$= E\{(as[n-1]+u[n]-\hat{s}[n|n-1])^2\}$$
$$= E\{(a(s[n-1]-\hat{s}[n|n-1])+u[n])^2\}$$
because $E\{(s[n-1]-\hat{s}[n-1|n-1])u[n]\} = 0$
$$= a^2M[n-1|n-1]+\sigma_u^2.$$

$$
\begin{aligned}
M[n|n] &= E\{(s[n]-\hat{s}[n|n])^2\} \\
&= E\{(s[n]-\hat{s}[n|n-1]-K[n](x[n]-\hat{s}[n|n-1]))^2\} \\
&= E\{(s[n]-\hat{s}[n|n-1])^2\} \\
&\qquad -2K[n]E\{(s[n]-\hat{s}[n|n-1])(x[n]-\hat{s}[n|n-1])\} \\
&\qquad\quad +K^2[n]E\{(x[n]-\hat{s}[n|n-1])^2\} \\
&= M[n|n-1]-2K[n]M[n|n-1] \\
&\qquad\qquad +K[n]M[n|n-1] \\
&= (1-K[n])M[n|n-1].
\end{aligned}
$$

In summary, scalar Kalman filter is obtained as:

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

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

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

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

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

4. Extension to vector form

Gauss-Markov model

$$x[n] = \boldsymbol{h}^T s[n-1] + w[n]$$
$$\boldsymbol{s}[n] = \boldsymbol{A}\boldsymbol{s}[n-1] + \boldsymbol{B}\boldsymbol{u}[n], \quad n \geq 0$$

Estimate $s[n]$ from observed data $[x[0], x[1], \cdots, x[n]]^T$

- $\boldsymbol{A} \in \Re^{p\times p}$, $\boldsymbol{B} \in \Re^{p\times r}$, $\boldsymbol{h}[n] \in \Re^{p\times 1}$
- $\boldsymbol{s}[n] \in \Re^{p\times 1}$, $\boldsymbol{s}[-1] \sim N(\mu_s, \boldsymbol{C}_s)$.
- $\boldsymbol{u}[n] \in \Re^{r\times 1}$, $\boldsymbol{u}[n] \sim N(0, \boldsymbol{Q})$
- $w[n] \in \Re^{1\times 1}$, $w[n] \sim N(0, \sigma_w^2)$

Vector Kalman filter:

- ***Prediction*** $$\hat{\boldsymbol{s}}[n|n-1] = \boldsymbol{A}\hat{\boldsymbol{s}}[n-1|n-1],$$
- ***Minimum prediction error***

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

- ***Kalman gain*** $$\boldsymbol{K}[n] = \frac{\boldsymbol{M}[n|n-1]\boldsymbol{h}[n]}{\sigma_w^2 + \boldsymbol{h}^T[n]\boldsymbol{M}[n|n-1]\boldsymbol{h}[n]},$$
- ***Correction***

$$\hat{\boldsymbol{s}}[n|n] = \hat{\boldsymbol{s}}[n|n-1] + \boldsymbol{K}[n](x[n] - \boldsymbol{h}^T[n]\hat{s}[n|n-1]),$$

- ***Minimun mean square error***

$$\boldsymbol{M}[n|n] = (\boldsymbol{I} - \boldsymbol{K}[n]\boldsymbol{h}^T[n])\boldsymbol{M}[n|n-1].$$

Applications

- Satellite control
- Autopilot
- Economy (especially macroeconomics), Time Series Econometrics
- Inertial navigation system
- Car navigation
- Weather forecast

Kalman filter

- Process and measurement equations:

$$\begin{cases} \mathbf{x}_{n+1} = \mathbf{A}\mathbf{x}_n + \mathbf{w}_n \\ \mathbf{z}_n = \mathbf{C}\mathbf{x}_n + \mathbf{v}_n \end{cases}$$

where **w** and **v** are Gaussian noises with mean **0** and covariance **Q** and **R**.

$$\mathbf{w}_n \sim N\left(\mathbf{0}, \mathbf{Q}\right), \quad \mathbf{v}_n \sim N\left(\mathbf{0}, \mathbf{R}\right)$$

Problem: Given the posterior probability at step *n*

$$p\left(\mathbf{x}_n \middle| \mathbf{Z}_{1:n}\right) = p\left(\mathbf{x}_n \middle| \mathbf{z}_1, \cdots, \mathbf{z}_n\right),$$

Compute the posterior probability at step *n*+1

$$p\left(\mathbf{x}_{n+1} \middle| \mathbf{Z}_{1:n+1}\right) = p\left(\mathbf{x}_{n+1} \middle| \mathbf{z}_1, \cdots, \mathbf{z}_n, \mathbf{z}_{n+1}\right).$$

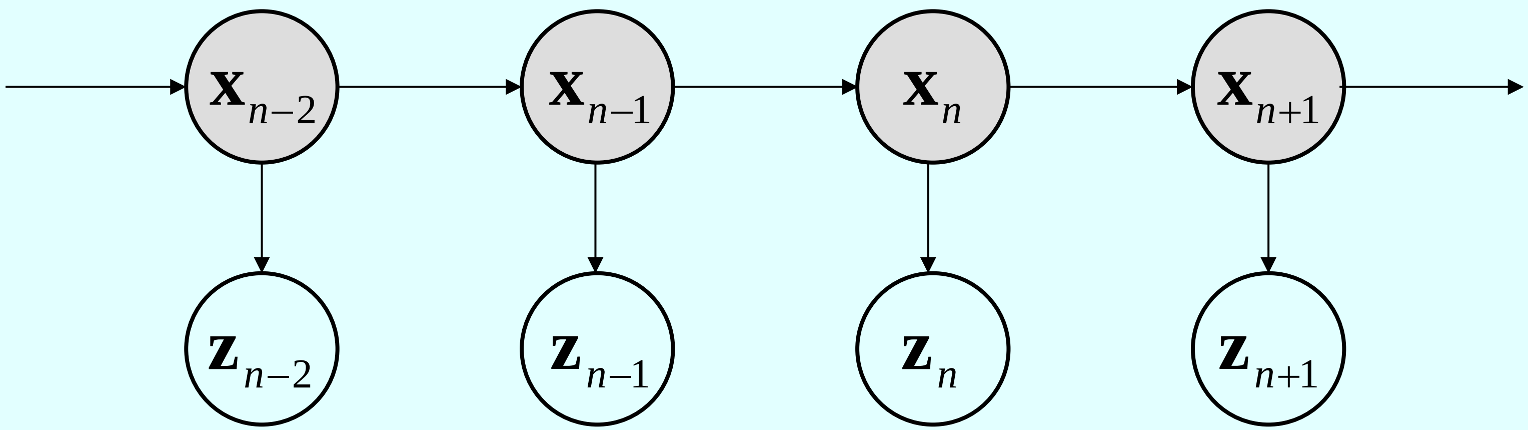
Kalman filter

- Process and measurement equations:

$$\begin{cases} \mathbf{x}_{n+1} = \mathbf{A}\mathbf{x}_n + \mathbf{w}_n \\ \mathbf{z}_n = \mathbf{C}\mathbf{x}_n + \mathbf{v}_n \end{cases}$$

where **w** and **v** are Gaussian noises with mean **0** and covariance **Q** and **R**.

$$\mathbf{w}_n \sim N(\mathbf{0}, \mathbf{Q}), \quad \mathbf{v}_n \sim N(\mathbf{0}, \mathbf{R})$$



Kalman filter

- Problem: Given the posterior probability at step *n*

$$p\left(\mathbf{x}_n \middle| \mathbf{Z}_{1:n}\right) = p\left(\mathbf{x}_n \middle| \mathbf{z}_1, \cdots, \mathbf{z}_n\right),$$

Compute the posterior probability at step *n*+1

$$p\left(\mathbf{x}_{n+1} \middle| \mathbf{Z}_{1:n+1}\right) = p\left(\mathbf{x}_{n+1} \middle| \mathbf{z}_1, \cdots, \mathbf{z}_n, \mathbf{z}_{n+1}\right).$$

$$p\left(\mathbf{x}_n \middle| \mathbf{Z}_{1:n}\right) \xrightarrow[\text{prediction}]{} p\left(\mathbf{x}_{n+1} \middle| \mathbf{Z}_{1:n}\right) \xrightarrow[\text{filtering}]{} p\left(\mathbf{x}_{n+1} \middle| \mathbf{Z}_{1:n+1}\right)$$

$$\begin{matrix} \hat{\mathbf{x}}_{n|n} \\ \Sigma_{n|n} \end{matrix} \xrightarrow[\text{prediction}]{} \begin{matrix} \hat{\mathbf{x}}_{n+1|n} \\ \Sigma_{n+1|n} \end{matrix} \xrightarrow[\text{filtering}]{} \begin{matrix} \hat{\mathbf{x}}_{n+1|n+1} \\ \Sigma_{n+1|n+1} \end{matrix}$$

Kalman filter

- Kalman filter: result

$$\hat{\mathbf{x}}_{n|n},\ \Sigma_{n|n} \xrightarrow{\text{prediction}} \hat{\mathbf{x}}_{n+1|n},\ \Sigma_{n+1|n} \xrightarrow{\text{filtering}} \hat{\mathbf{x}}_{n+1|n+1},\ \Sigma_{n+1|n+1}$$

prediction step:

$$\boxed{\begin{aligned}\hat{\mathbf{x}}_{n+1|n} &= \mathbf{A}\hat{\mathbf{x}}_{n|n}\\ \Sigma_{n+1|n} &= \mathbf{A}\Sigma_{n|n}\mathbf{A}^{\mathrm{T}} + \mathbf{Q}\end{aligned}}$$

filtering step:

$$\boxed{\begin{aligned}\hat{\mathbf{x}}_{n+1|n+1} &= \hat{\mathbf{x}}_{n+1|n} + \mathbf{K}_{n+1}\left(\mathbf{z}_{n+1} - \mathbf{C}\hat{\mathbf{x}}_{n+1|n}\right)\\ \Sigma_{n+1|n+1} &= \left(\mathbf{I} - \mathbf{K}_{n+1}\mathbf{C}\right)\Sigma_{n+1|n}\end{aligned}}$$

$$\mathbf{K}_{n+1} = \Sigma_{n+1|n}\mathbf{C}^{\mathrm{T}}\left(\mathbf{C}\Sigma_{n+1|n}\mathbf{C}^{\mathrm{T}} + \mathbf{R}\right)^{-1}$$

(Simplified) derivation based on mean and variance (1/3)

- Prediction step:

$$\hat{\mathbf{x}}_{n+1|n} = \mathrm{E}\left[\mathbf{x}_{n+1} \mid \mathbf{Z}_{1:n}\right] = \mathrm{E}\left[\mathbf{A}\mathbf{x}_n + \mathbf{w}_n \mid \mathbf{Z}_{1:n}\right] = \mathbf{A}\hat{\mathbf{x}}_{n|n}$$

$$\begin{aligned}
\Sigma_{n+1|n} &= \mathrm{Cov}\left[\mathbf{x}_{n+1} \mid \mathbf{Z}_{1:n}\right] \\
&= \mathrm{E}\left[\left(\mathbf{x}_{n+1} - \hat{\mathbf{x}}_{n+1|n}\right)\left(\mathbf{x}_{n+1} - \hat{\mathbf{x}}_{n+1|n}\right)^{\mathrm{T}} \mid \mathbf{Z}_{1:n}\right] \\
&= \mathrm{E}\left[\left\{\mathbf{A}\left(\mathbf{x}_n - \hat{\mathbf{x}}_{n|n}\right) + \mathbf{w}_n\right\}\left\{\mathbf{A}\left(\mathbf{x}_n - \hat{\mathbf{x}}_{n|n}\right) + \mathbf{w}_n\right\}^{\mathrm{T}} \mid \mathbf{Z}_{1:n}\right] \\
&= \mathbf{A}\Sigma_{n|n}\mathbf{A}^{\mathrm{T}} + \mathbf{Q}
\end{aligned}$$

$$\boxed{\begin{aligned}
\hat{\mathbf{x}}_{n+1|n} &= \mathbf{A}\hat{\mathbf{x}}_{n|n} \\
\Sigma_{n+1|n} &= \mathbf{A}\Sigma_{n|n}\mathbf{A}^{\mathrm{T}} + \mathbf{Q}
\end{aligned}}$$

(Simplified) derivation based on mean and variance (2/3)

- Filtering step:
  Assume that the posterior mean has a linear form:

$$\hat{\mathbf{x}}_{n+1|n+1} = \mathrm{E}\left[\mathbf{x}_{n+1} \mid \mathbf{Z}_{1:n+1}\right] = \hat{\mathbf{x}}_{n+1|n} + \mathbf{K}_{n+1}\left(\mathbf{z}_{n+1} - \mathbf{C}\hat{\mathbf{x}}_{n+1|n}\right).$$

  Then, the Kalman gain K is determined so as to minimize the trace of covariance matrix:

$$\mathrm{Cov}\left[\mathbf{x}_{n+1} \mid \mathbf{Z}_{1:n+1}\right].$$

  By noting

$$\mathbf{x}_{n+1} - \hat{\mathbf{x}}_{n+1|n+1} = \left(\mathbf{I} - \mathbf{K}_{n+1}\mathbf{C}\right)\left(\mathbf{x}_{n+1} - \hat{\mathbf{x}}_{n+1|n}\right) + \mathbf{K}_{n+1}\mathbf{v}_{n+1},$$

  we obtain

$$\mathrm{Cov}\left[\mathbf{x}_{n+1} \mid \mathbf{Z}_{1:n+1}\right] = \left(\mathbf{I} - \mathbf{K}_{n+1}\mathbf{C}\right)\Sigma_{n+1|n}\left(\mathbf{I} - \mathbf{K}_{n+1}\mathbf{C}\right)^{\mathrm{T}} + \mathbf{K}_{n+1}\mathbf{R}\mathbf{K}_{n+1}^{\mathrm{T}}$$

(Simplified) derivation based on mean and variance (3/3)

- Finding the Kalman gain:

$$
\begin{aligned}
0 &= \frac{\partial}{\partial \mathbf{K}_{n+1}} \operatorname{tr} \operatorname{Cov}\left[\mathbf{x}_{n+1} \mid \mathbf{Z}_{1:n+1}\right] \\
&= \frac{\partial}{\partial \mathbf{K}_{n+1}} \operatorname{tr}\left\{\left(\mathbf{I}-\mathbf{K}_{n+1}\mathbf{C}\right)\Sigma_{n+1|n}\left(\mathbf{I}-\mathbf{K}_{n+1}\mathbf{C}\right)^{\mathrm{T}}+\mathbf{K}_{n+1}\mathbf{R}\mathbf{K}_{n+1}^{\mathrm{T}}\right\} \\
&= -2\left(\mathbf{I}-\mathbf{K}_{n+1}\mathbf{C}\right)\Sigma_{n+1|n}\mathbf{C}^{\mathrm{T}}+2\mathbf{K}_{n+1}\mathbf{R}
\end{aligned}
$$

$$
\therefore \mathbf{K}_{n+1} = \Sigma_{n+1|n}\mathbf{C}^{\mathrm{T}}\left(\mathbf{C}\Sigma_{n+1|n}\mathbf{C}^{\mathrm{T}}+\mathbf{R}\right)^{-1}
$$

$$
\boxed{
\begin{aligned}
\hat{\mathbf{x}}_{n+1|n+1} &= \hat{\mathbf{x}}_{n+1|n}+\mathbf{K}_{n+1}\left(\mathbf{z}_{n+1}-\mathbf{C}\hat{\mathbf{x}}_{n+1|n}\right) \\
\Sigma_{n+1|n+1} &= \left(\mathbf{I}-\mathbf{K}_{n+1}\mathbf{C}\right)\Sigma_{n+1|n}
\end{aligned}
}
$$

(More rigorous) derivation based on probability densities (1/3)

- Prediction step:

$$
\begin{aligned}
p\left(\mathbf{x}_{n+1} \mid \mathbf{Z}_{1:n}\right) &= \int d\mathbf{x}_n \, p\left(\mathbf{x}_{n+1}, \mathbf{x}_n \mid \mathbf{Z}_{1:n+1}\right) \\
&= \int d\mathbf{x}_n \, p\left(\mathbf{x}_{n+1} \mid \mathbf{x}_n\right) p\left(\mathbf{x}_n \mid \mathbf{Z}_{1:n+1}\right) \\
&= \int d\mathbf{x}_n \frac{1}{(2\pi)^{n/2} |\mathbf{Q}|} \exp\left(-\frac{1}{2}\left(\mathbf{x}_{n+1} - \mathbf{A}\mathbf{x}_n\right)^{\mathrm{T}} \mathbf{Q}^{-1} \left(\mathbf{x}_{n+1} - \mathbf{A}\mathbf{x}_n\right)\right) \\
&\qquad \times \frac{1}{(2\pi)^{n/2} \left|\Sigma_{n|n}\right|} \exp\left(-\frac{1}{2}\left(\mathbf{x}_n - \hat{\mathbf{x}}_{n|n}\right)^{\mathrm{T}} \Sigma_{n|n}^{-1} \left(\mathbf{x}_n - \hat{\mathbf{x}}_{n|n}\right)\right) \\
&= N\left(\mathbf{x}_{n+1}; \hat{\mathbf{x}}_{n+1|n}, \Sigma_{n+1|n}\right)
\end{aligned}
$$

$$
\boxed{
\begin{aligned}
\hat{\mathbf{x}}_{n+1|n} &= \mathbf{A}\hat{\mathbf{x}}_{n|n} \\
\Sigma_{n+1|n} &= \mathbf{A}\Sigma_{n|n}\mathbf{A}^{\mathrm{T}} + \mathbf{Q}
\end{aligned}
}
$$

(More rigorous) derivation based on probability densities (2/3)

- Filtering step:

$$
\begin{aligned}
p\left(\mathbf{x}_{n+1} \mid \mathbf{Z}_{1:n+1}\right) &= p\left(\mathbf{x}_{n+1} \mid \mathbf{z}_{n+1}, \mathbf{Z}_{1:n}\right) \\
&= \frac{p\left(\mathbf{z}_{n+1} \mid \mathbf{x}_{n+1}\right) p\left(\mathbf{x}_{n+1} \mid \mathbf{Z}_{1:n}\right)}{p\left(\mathbf{z}_{n+1} \mid \mathbf{Z}_{1:n}\right)} \\
&\propto p\left(\mathbf{z}_{n+1} \mid \mathbf{x}_{n+1}\right) p\left(\mathbf{x}_{n+1} \mid \mathbf{Z}_{1:n}\right) \\
&\propto \exp\left(-\frac{1}{2}\left(\mathbf{z}_{n+1} - \mathbf{C}\mathbf{x}_{n+1|n}\right)^{\mathrm{T}} \mathbf{R}^{-1} \left(\mathbf{z}_{n+1} - \mathbf{C}\mathbf{x}_{n+1|n}\right)\right) \exp\left(-\frac{1}{2}\left(\mathbf{x}_{n+1} - \hat{\mathbf{x}}_{n+1|n}\right)^{\mathrm{T}} \Sigma_{n+1|n}^{-1} \left(\mathbf{x}_{n+1} - \hat{\mathbf{x}}_{n+1|n}\right)\right) \\
&= \exp\left(-\frac{1}{2}\left(\mathbf{x}_{n+1} - \hat{\mathbf{x}}_{n+1|n+1}\right)^{\mathrm{T}} \Sigma_{n+1|n+1}^{-1} \left(\mathbf{x}_{n+1} - \hat{\mathbf{x}}_{n+1|n+1}\right)\right)
\end{aligned}
$$

$$
\boxed{
\begin{aligned}
\hat{\mathbf{x}}_{n+1|n+1} &= \left(\Sigma_{n+1|n}^{-1} + \mathbf{C}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{C}\right)^{-1} \left(\Sigma_{n+1|n}^{-1}\hat{\mathbf{x}}_{n+1|n} + \mathbf{C}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{z}_{n+1}\right) \\
\Sigma_{n+1|n+1} &= \left(\Sigma_{n+1|n}^{-1} + \mathbf{C}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{C}\right)^{-1}
\end{aligned}
}
$$

- Filtering step:

$$\Sigma_{n+1|n+1} = \left(\Sigma_{n+1|n}^{-1} + \mathbf{C}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{C}\right)^{-1}$$

$$= \Sigma_{n+1|n} - \Sigma_{n+1|n}\mathbf{C}^{\mathrm{T}}\left(\mathbf{R} + \mathbf{C}\Sigma_{n+1|n}\mathbf{C}^{\mathrm{T}}\right)^{-1}\mathbf{C}\Sigma_{n+1|n}$$

$$= \left(\mathbf{I} - \mathbf{K}_{n+1}\mathbf{C}\right)\Sigma_{n+1|n}$$

$$\hat{\mathbf{x}}_{n+1|n+1} = \left(\Sigma_{n+1|n}^{-1} + \mathbf{C}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{C}\right)^{-1}\left(\Sigma_{n+1|n}^{-1}\hat{\mathbf{x}}_{n+1|n} + \mathbf{C}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{z}_{n+1}\right)$$

$$= \hat{\mathbf{x}}_{n+1|n} + \mathbf{K}_{n+1}\left(\mathbf{z}_{n} - \mathbf{C}\hat{\mathbf{x}}_{n+1|n}\right)$$

$$\boxed{\begin{aligned}\hat{\mathbf{x}}_{n+1|n+1} &= \hat{\mathbf{x}}_{n+1|n} + \mathbf{K}_{n+1}\left(\mathbf{z}_{n+1} - \mathbf{C}\hat{\mathbf{x}}_{n+1|n}\right)\\ \Sigma_{n+1|n+1} &= \left(\mathbf{I} - \mathbf{K}_{n+1}\mathbf{C}\right)\Sigma_{n+1|n}\end{aligned}}$$