430.457

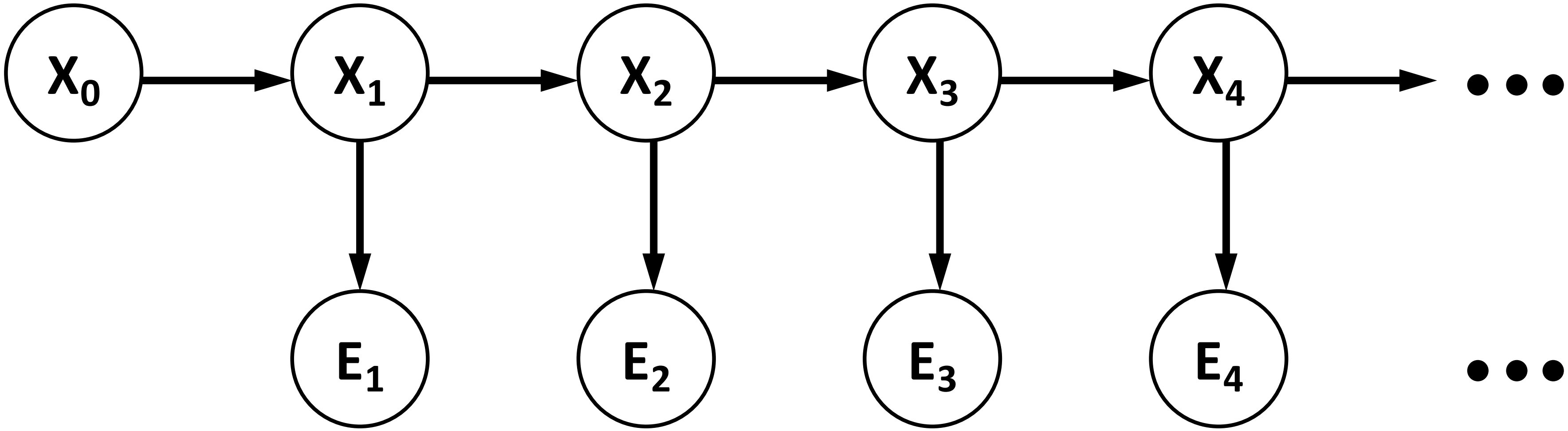
Introduction to Intelligent Systems

Prof. Songhwai Oh

ECE, SNU

KALMAN FILTERING

Kalman Filtering vs. Hidden Markov Model

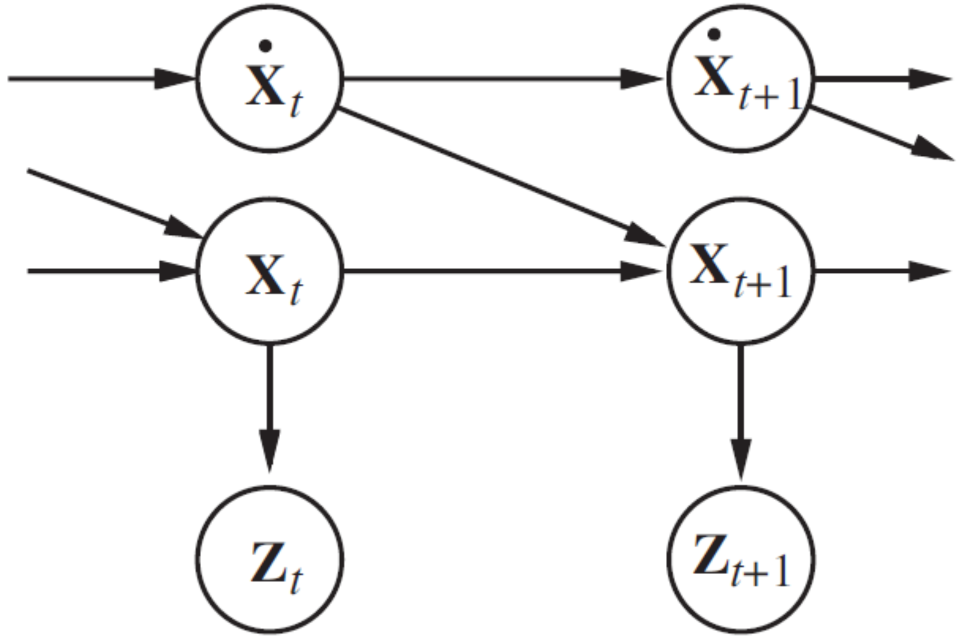


- **Hidden Markov model**: Discrete state space, $X_t \in \{1, \ldots, S\}$.
- **Kalman filtering**: Continuous state space, $X_t \in \mathbb{R}^d$. Linear-Gauss Model.

Kalman Filtering

- **Kalman filtering**: Continuous state space, $X_t \in \mathbb{R}^d$. Linear-Gauss Model.
  - Linear Gaussian transition model: $X_{t+\Delta} = X_t + \dot{X}_t\Delta + W_t$, where Δ is the time interval between observations and $W_t \sim \mathcal{N}(0, \sigma^2)$.

$$P(X_{t+\Delta} = x_{t+\Delta} | X_t = x_t, \dot{X}_t = \dot{x}_t) = \mathcal{N}(x_{t+\Delta}; x_t + \dot{x}_t\Delta, \sigma^2)$$



Properties of the Gaussian Distribution

1. If the current distribution $\mathbf{P}(\mathbf{X}_t \mid \mathbf{e}_{1:t})$ is Gaussian and the transition model $\mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{x}_t)$ is linear Gaussian, then the one-step predicted distribution given by

$$\mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t}) = \int_{\mathbf{x}_t} \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{x}_t) P(\mathbf{x}_t \mid \mathbf{e}_{1:t})\, d\mathbf{x}_t \tag{15.17}$$

is also a Gaussian distribution.

2. If the prediction $\mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t})$ is Gaussian and the sensor model $\mathbf{P}(\mathbf{e}_{t+1} \mid \mathbf{X}_{t+1})$ is linear Gaussian, then, after conditioning on the new evidence, the updated distribution

$$\mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t+1}) = \alpha\, \mathbf{P}(\mathbf{e}_{t+1} \mid \mathbf{X}_{t+1}) \mathbf{P}(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t}) \tag{15.18}$$

is also a Gaussian distribution.

- $Forward$ operator for Kalman filtering: $\mathbf{f}_{1:t}$ contains μ_t (mean) and $\mathbf{\Sigma}_t$ (covariance matrix).
- Initialization: $\mathbf{f}_{1:0} = P(X_0) = \mathcal{N}(\mu_0, \mathbf{\Sigma}_0)$.

Simple One-Dimensional Example

- Prior distribution: $P(X_0) = \mathcal{N}(\mu_0, \sigma_0^2)$

$$P(x_0) = \alpha\, e^{-\frac{1}{2}\left(\frac{(x_0-\mu_0)^2}{\sigma_0^2}\right)}$$

- Transition model: $X_{t+1} = X_t + W_t$, where $W_t \sim \mathcal{N}(0, \sigma_x^2)$

$$P(x_{t+1} \,|\, x_t) = \alpha\, e^{-\frac{1}{2}\left(\frac{(x_{t+1}-x_t)^2}{\sigma_x^2}\right)}$$

- Sensor model: $Z_t = X_t + V_t$, where $V_t \sim \mathcal{N}(0, \sigma_z^2)$

$$P(z_t \,|\, x_t) = \alpha\, e^{-\frac{1}{2}\left(\frac{(z_t-x_t)^2}{\sigma_z^2}\right)}$$

One Step Prediction

$$P(x_1) = \int_{-\infty}^{\infty} P(x_1 \,|\, x_0) P(x_0)\, dx_0 = \alpha \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{(x_1-x_0)^2}{\sigma_x^2}\right)} e^{-\frac{1}{2}\left(\frac{(x_0-\mu_0)^2}{\sigma_0^2}\right)} dx_0$$

$$= \alpha \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{\sigma_0^2(x_1-x_0)^2+\sigma_x^2(x_0-\mu_0)^2}{\sigma_0^2\sigma_x^2}\right)} dx_0 \,.$$

Completing the square:

$$\begin{aligned} ax_0^2 + bx_0 + c &= \left(ax_0^2 + bx_0 + \frac{b^2}{4a}\right) + \left(c - \frac{b^2}{4a}\right) \\ &= a\left(x_0^2 + 2x_0\frac{b}{2a} + \left(\frac{b}{2a}\right)^2\right) + \left(c - \frac{b^2}{4a}\right) \\ &= a\left(x_0 + \frac{b}{2a}\right)^2 + \left(c - \frac{b^2}{4a}\right) \end{aligned}$$

$$\begin{aligned} a &= \frac{\sigma_0^2 + \sigma_x^2}{\sigma_0^2\sigma_x^2} \\ b &= \frac{-2(\sigma_0^2 x_1 + \sigma_x^2 \mu_0)}{\sigma_0^2\sigma_x^2} \\ c &= \frac{\sigma_0^2 x_1^2 + \sigma_x^2 \mu_0^2}{\sigma_0^2\sigma_x^2} \end{aligned}$$

$$P(x_1) = \alpha\, e^{-\frac{1}{2}\left(c-\frac{b^2}{4a}\right)} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(a(x_0-\frac{-b}{2a})^2\right)} dx_0 \;= \alpha\, e^{-\frac{1}{2}\left(\frac{(x_1-\mu_0)^2}{\sigma_0^2+\sigma_x^2}\right)}$$

Hence, $P(X_1) = \mathcal{N}(\mu_0, (\sigma_0^2 + \sigma_x^2))$.

Measurement Update

$$P(x_1 \mid z_1) = \alpha P(z_1 \mid x_1) P(x_1)$$
$$= \alpha e^{-\frac{1}{2}\left(\frac{(z_1 - x_1)^2}{\sigma_z^2}\right)} e^{-\frac{1}{2}\left(\frac{(x_1 - \mu_0)^2}{\sigma_0^2 + \sigma_x^2}\right)}$$

after completing the square: $P(x_1 \mid z_1) = \alpha e^{-\frac{1}{2}\left(\frac{\left(x_1 - \frac{(\sigma_0^2 + \sigma_x^2) z_1 + \sigma_z^2 \mu_0}{\sigma_0^2 + \sigma_x^2 + \sigma_z^2}\right)^2}{(\sigma_0^2 + \sigma_x^2)\sigma_z^2 / (\sigma_0^2 + \sigma_x^2 + \sigma_z^2)}\right)}$

Hence, $P(X_{t+1}) = \mathcal{N}(\mu_{t+1}, \sigma_{t+1}^2)$, where

$$\mu_{t+1} = \frac{(\sigma_t^2 + \sigma_x^2) z_{t+1} + \sigma_z^2 \mu_t}{\sigma_t^2 + \sigma_x^2 + \sigma_z^2} \qquad \text{and} \qquad \sigma_{t+1}^2 = \frac{(\sigma_t^2 + \sigma_x^2)\sigma_z^2}{\sigma_t^2 + \sigma_x^2 + \sigma_z^2}$$

Notice that μ_{t+1} is a weighted mean between μ_t and z_{t+1}.

Numerical Example

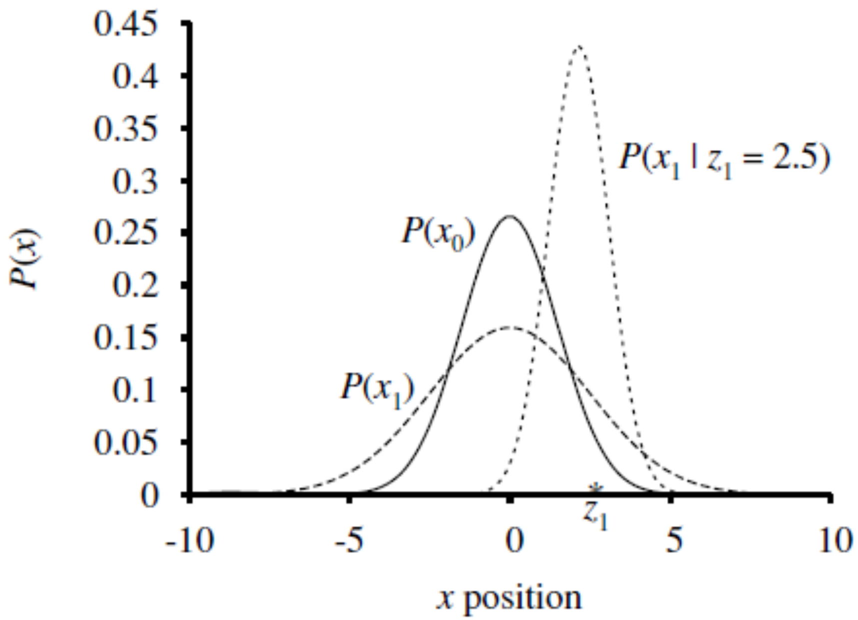


Figure 15.10 Stages in the Kalman filter update cycle for a random walk with a prior given by $\mu_0 = 0.0$ and $\sigma_0 = 1.0$, transition noise given by $\sigma_x = 2.0$, sensor noise given by $\sigma_z = 1.0$, and a first observation $z_1 = 2.5$ (marked on the x-axis). Notice how the prediction $P(x_1)$ is flattened out, relative to $P(x_0)$, by the transition noise. Notice also that the mean of the posterior distribution $P(x_1 \mid z_1)$ is slightly to the left of the observation z_1 because the mean is a weighted average of the prediction and the observation.

Kalman Filtering

- **Multivariate Gaussian distribution**

$$\mathcal{N}(\mathbf{x};\mu,\mathbf{\Sigma}) = \alpha \exp\left(-\frac{1}{2}(\mathbf{x}-\mu)^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\mu)\right)$$

- Prior distribution: $P(\mathbf{X}_0) = \mathcal{N}(\mu_0, \mathbf{\Sigma}_0)$
- Transition model: $\mathbf{X}_{t+1} = \mathbf{F}\mathbf{X}_t + \mathbf{W}_t$, where $\mathbf{W}_t \sim \mathcal{N}(\mathbf{0}, \Sigma_x)$

$$P(\mathbf{x}_{t+1}|\mathbf{x}_t) = \mathcal{N}(\mathbf{x}_{t+1}; \mathbf{F}\mathbf{x}_t, \mathbf{\Sigma}_x)$$

- Sensor model: $\mathbf{Z}_t = \mathbf{H}\mathbf{X}_t + \mathbf{V}_t$, where $\mathbf{V}_t \sim \mathcal{N}(\mathbf{0}, \Sigma_z)$

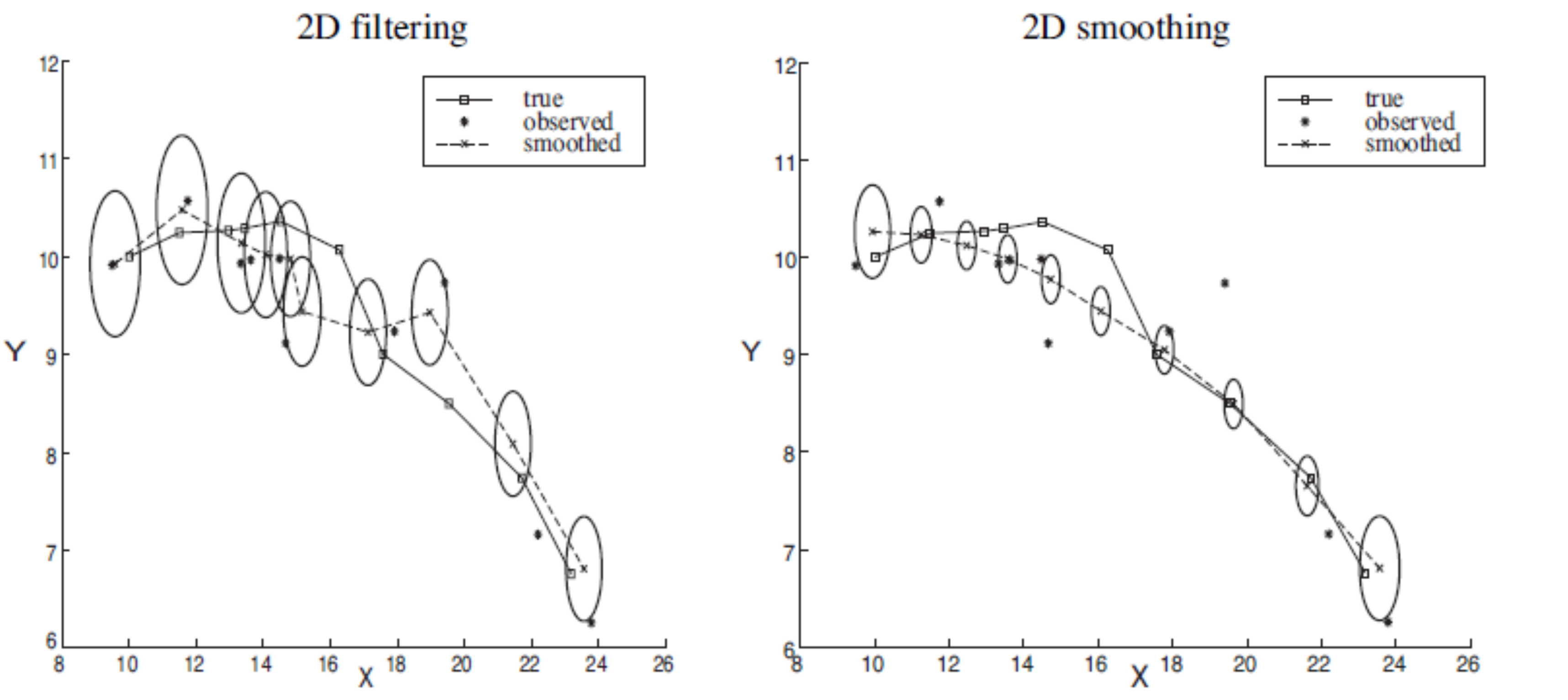
$$P(\mathbf{z}_t|\mathbf{x}_t) = \mathcal{N}(\mathbf{z}_t; \mathbf{H}\mathbf{x}_t, \mathbf{\Sigma}_z)$$

- Update equations:

$$\begin{aligned}\boldsymbol{\mu}_{t+1} &= \mathbf{F}\boldsymbol{\mu}_t + \mathbf{K}_{t+1}(\mathbf{z}_{t+1} - \mathbf{H}\mathbf{F}\boldsymbol{\mu}_t) \\ \mathbf{\Sigma}_{t+1} &= (\mathbf{I} - \mathbf{K}_{t+1}\mathbf{H})(\mathbf{F}\mathbf{\Sigma}_t\mathbf{F}^\top + \mathbf{\Sigma}_x)\end{aligned}$$

$$\mathbf{K}_{t+1} = (\mathbf{F}\mathbf{\Sigma}_t\mathbf{F}^\top + \mathbf{\Sigma}_x)\mathbf{H}^\top(\mathbf{H}(\mathbf{F}\mathbf{\Sigma}_t\mathbf{F}^\top + \mathbf{\Sigma}_x)\mathbf{H}^\top + \mathbf{\Sigma}_z)^{-1}$$ Kalman Gain

Numerical Example



(a)

(b)