ECE276A: Sensing & Estimation in Robotics
Lecture 6: Bayes and Kalman Filtering

Lecturer:

Nikolay Atanasov: natanasov@ucsd.edu

Teaching Assistants:

Siwei Guo: s9guo@eng.ucsd.edu

Anwesan Pal: a2pal@eng.ucsd.edu

UC San Diego

JACOBS SCHOOL OF ENGINEERING

Electrical and Computer Engineering

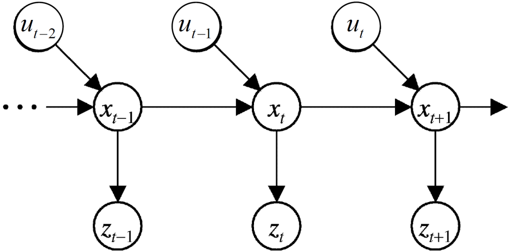
Filtering Examples

- Track the center $c_t \in \mathbb{R}^2$ and radius $r_t \in \mathbb{R}$ of a ball in images: http://www.pyimagesearch.com/2015/09/14/ball-tracking-with-opencv/
- Track the positions $p_t \in \mathbb{R}^{3n}$ of n people
- Track the position $p_t \in \mathbb{R}^3$ and orientation $R_t \in SO(3)$ of a camera: https://www.youtube.com/watch?v=CsJkci5lfco

Bayes Filter

- A **Bayes filter** is a probabilistic tool for estimating the state of a dynamical system that combines evidence from control inputs and system observations using **Markov assumptions** and **Bayes rule**:
  - **Total probability**: $p(x) = \int p(x, y)dy$
  - **Conditioning**: $p(x, y) = p(y \mid x)p(x)$
  - **Bayes rule**: $p(x \mid y, z) = \frac{p(y \mid x, z)p(x \mid z)}{\int p(y, s \mid z)ds} = \frac{p(y \mid x, z)p(z \mid x)p(x)}{p(y \mid z)p(z)}$
- Special cases of the Bayes filter:
  - Kalman filter
  - Particle filter
  - Forward algorithm for Hidden Markov models
  - Dynamic Bayesian networks
  - State evolution in Partially Observable Markov Decision Processes (POMDPs)

Filtering Problem



- State: x_t (**hidden**)
- Control input: u_t
- Observation: z_t
- **Markov Assumptions**
  - The state x_{t+1} only depends on the previous time input u_t and state x_t
  - The observation z_t only depends on the state x_t
- **Joint distribution**:

$$p(x_{0:T}, z_{0:T}, u_{0:T-1}) = \underbrace{p_{0|0}(x_0)}_{\text{prior}} \prod_{t=0}^{T} \underbrace{p_h(z_t \mid x_t)}_{\text{observation model}} \prod_{t=0}^{T} \underbrace{p_a(x_t \mid x_{t-1}, u_{t-1})}_{\text{motion model}}$$

- **Filtering**: keeps track of

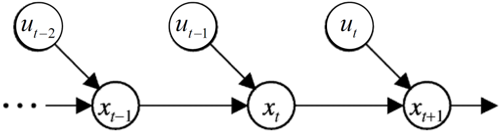
$$p_{t|t}(x_t) := p(x_t \mid z_{0:t}, u_{0:t-1})$$
$$p_{t+1|t}(x_{t+1}) := p(x_{t+1} \mid z_{0:t}, u_{0:t})$$

- **Smoothing**: keeps track of

$$p_{t|t}(x_{0:t}) := p(x_{0:t} \mid z_{0:t}, u_{0:t-1})$$
$$p_{t+1|t}(x_{0:t+1}) := p(x_{0:t+1} \mid z_{0:t}, u_{0:t})$$

First-Order Markov Motion Model

- State: x_t
- Control input: u_t
- Motion noise: w_t



- **Motion model**: $a(\cdot,\cdot,\cdot)$ or $p_a(\cdot \mid \cdot,\cdot)$

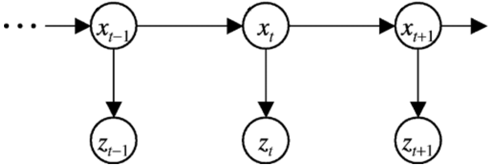
$$\boxed{x_{t+1} = a(x_t, u_t, w_t) \sim p_a(\cdot \mid x_t, u_t)}$$

- **Prediction step**: given a prior density $p_{t|t}$ over x_t and the control input u_t, the goal is to use the motion model above to compute the predicted density $p_{t+1|t}$ over x_{t+1}:

$$\boxed{p_{t+1|t}(x) = \int p_a(x \mid s, u_t) p_{t|t}(s) ds}$$

First-Order Markov Observation Model

- State: x_t
- Observation: z_t
- Observation noise: v_t



- **Observation model**: $h(\cdot, \cdot)$ or $p_h(\cdot \mid \cdot)$

$$\boxed{z_t = h(x_t, v_t) \sim p_h(\cdot \mid x_t)}$$

- **Update step**: given the predicted density $p_{t+1|t}$ over x_{t+1} and the measurement z_{t+1}, the goal is to use the observation model above to incorporate the measurement information and obtain the posterior density $p_{t+1|t+1}$ over x_{t+1}:

$$\boxed{p_{t+1|t+1}(x) = \frac{p_h(z_{t+1} \mid x) p_{t+1|t}(x)}{\int p_h(z_{t+1} \mid s) p_{t+1|t}(s) ds}}$$

Bayes Filter

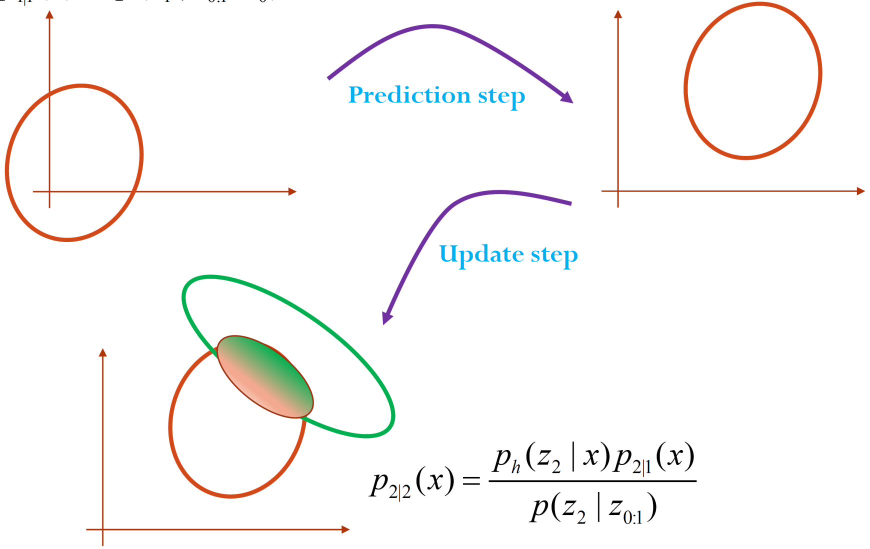
$$
\begin{aligned}
p_{t+1|t+1}(x_{t+1}) =& p(x_{t+1} \mid z_{0:t+1}, u_{0:t}) \\
\overset{\text{Bayes}}{=\!=\!=} & \frac{1}{\eta_{t+1}} p(z_{t+1} \mid x_{t+1}, z_{0:t}, u_{0:t}) p(x_{t+1} \mid z_{0:t}, u_{0:t}) \\
\overset{\text{Markov}}{=\!=\!=} & \frac{1}{\eta_{t+1}} p_h(z_{t+1} \mid x_{t+1}) p(x_{t+1} \mid z_{0:t}, u_{0:t}) \\
\overset{\text{Total prob.}}{=\!=\!=} & \frac{1}{\eta_{t+1}} p_h(z_{t+1} \mid x_{t+1}) \int p(x_{t+1}, x_t \mid z_{0:t}, u_{0:t}) dx_t \\
\overset{\text{Cond. prob.}}{=\!=\!=} & \frac{1}{\eta_{t+1}} p_h(z_{t+1} \mid x_{t+1}) \int p(x_{t+1} \mid z_{0:t}, u_{0:t}, x_t) p(x_t \mid z_{0:t}, u_{0:t}) dx_t \\
\overset{\text{Markov}}{=\!=\!=} & \frac{1}{\eta_{t+1}} p_h(z_{t+1} \mid x_{t+1}) \int p_a(x_{t+1} \mid x_t, u_t) p(x_t \mid z_{0:t}, u_{0:t-1}) dx_t \\
= & \boxed{\frac{1}{\eta_{t+1}} p_h(z_{t+1} \mid x_{t+1}) \int p_a(x_{t+1} \mid x_t, u_t) p_{t|t}(x_t) dx_t}
\end{aligned}
$$

- **Normalization constant**: $\eta_{t+1} := p(z_{t+1} \mid z_{0:t}, u_{0:t})$

Bayes Filter

$$p_{1|1}(x) := p(x_1 \mid z_{0:1}, u_0)$$

$$p_{2|1}(x) = \int p_a(x \mid s, u_1)\, p_{1|1}(s)\, ds$$



$$p_{2|2}(x) = \frac{p_h(z_2 \mid x)\, p_{2|1}(x)}{p(z_2 \mid z_{0:1})}$$

Kalman Filter

- A **Kalman filter** is a Bayes filter for which:
  - The prior pdf $p_{0|0}$ is Gaussian
  - The motion model is linear in the state and affected by Gaussian noise
  - The observation model is linear in the state and affected by Gaussian noise
  - The process noise w_t and measurement noise v_t are independent of each other, of the state x_t and across time

- **Prior**: $x_t \mid z_{0:t}, u_{0:t-1} \sim \mathcal{N}(\mu_{t|t}, \Sigma_{t|t})$

- **Motion Model**:

$$x_{t+1} = a(x_t, u_t, w_t) := Ax_t + Bu_t + w_t, \quad w_t \sim \mathcal{N}(0, W)$$

$$x_{t+1} \mid x_t, u_t \sim \mathcal{N}(Ax_t + Bu_t, W), \quad A \in \mathbb{R}^{d\times d}, B \in \mathbb{R}^{d\times d_u},\; W \in \mathbb{R}^{d\times d}$$

- **Observation Model**:

$$z_t = h(x_t, v_t) := Hx_t + v_t, \quad v_t \sim \mathcal{N}(0, V)$$

$$z_t \mid x_t \sim \mathcal{N}(Hx_t, V), \quad H \in \mathbb{R}^{d_z\times d},\; V \in \mathbb{R}^{d_z\times d_z}$$

Gaussian Distribution Properties

- **Stable distributon**: a linear combination $aX_1 + bX_2$ of two independent copies of a random variable X has the same distribution as $cX + d$ up to location d and scale $c > 0$ parameters

- The Gaussian distribution is stable, i.e., the space of Gaussian pdfs is **closed under convolution**:

$$\int \phi(x; As, W)\phi(s; \mu, \Sigma)ds = \phi\left(x; A\mu, A\Sigma A^T + W\right)$$

- The space of Gaussian pdfs is **closed under geometric averages** (up to scaling):

$$\prod_k \phi^{\alpha_k}(x; \mu_k, \Sigma_k) \propto \phi\left(x; \left(\sum_k \alpha_k \Sigma_k^{-1}\right)^{-1}\left(\sum_k \Sigma_k^{-1}\mu_k\right), \left(\sum_k \alpha_k \Sigma_k^{-1}\right)^{-1}\right)$$

Gaussian Marginals and Conditionals

- Consider a **joint Gaussian distribution**:

$$x := \begin{pmatrix} x_A \\ x_B \end{pmatrix} \sim \mathcal{N}\left(\underbrace{\begin{pmatrix} \mu_A \\ \mu_B \end{pmatrix}}_{\mu}, \underbrace{\begin{bmatrix} \Sigma_{AA} & \Sigma_{AB} \\ \Sigma_{AB}^T & \Sigma_{BB} \end{bmatrix}}_{\Sigma} \right)$$

- The **marginal distribution** is also Gaussian: $x_A \sim \mathcal{N}(\mu_A, \Sigma_{AA})$

- The **conditional distribution** is also Gaussian:

$$x_B \mid x_A \sim \mathcal{N}\left(\mu_B + \Sigma_{AB}^T \Sigma_{AA}^{-1} (x_A - \mu_A), \underbrace{\Sigma_{BB} - \Sigma_{AB}^T \Sigma_{AA}^{-1} \Sigma_{AB}}_{\text{Schur complement of } \Sigma_{AA}} \right)$$

Matrix Manipulation

- The main tools for proving the previous results and for deriving the Kalman filter are:
  - **Matrix inversion lemma**:

$$(A + BDC)^{-1} = A^{-1} - A^{-1}B\left(D^{-1} + CA^{-1}B\right)^{-1}CA^{-1}$$

  - **Matrix block inversion**:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} (A - BD^{-1}C)^{-1} & -(A - BD^{-1}C)^{-1}BD^{-1} \\ -D^{-1}C(A - BD^{-1}C)^{-1} & D^{-1} + D^{-1}C(A - BD^{-1}C)^{-1}BD^{-1} \end{bmatrix}$$

  - **Square completion**:

$$\frac{1}{2}x^TAx + b^Tx + c = \frac{1}{2}\left(x + A^{-1}b\right)^T A\left(x + A^{-1}b\right) - \frac{1}{2}b^TA^{-1}b + c$$

Information Space

- Natural parameterization of a Gaussian distribution: $x \sim \mathcal{G}(\nu, \Omega)$
- **Information mean**: $\nu = \Sigma^{-1}\mu$
- **Information matrix**: $\Omega := \begin{bmatrix} \Omega_{AA} & \Omega_{AB} \\ \Omega_{AB}^T & \Omega_{BB} \end{bmatrix} = \begin{bmatrix} \Sigma_{AA} & \Sigma_{AB} \\ \Sigma_{AB}^T & \Sigma_{BB} \end{bmatrix}^{-1}$

$$
\begin{aligned}
\Omega_{AA} &= (\Sigma_{AA} - \Sigma_{AB}\Sigma_{BB}^{-1}\Sigma_{AB}^T)^{-1} \\
\Omega_{AB} &= -\Omega_{AA}\Sigma_{AB}\Sigma_{BB}^{-1} \\
\Omega_{BB} &= \Sigma_{BB}^{-1} + \Sigma_{BB}^{-1}\Sigma_{AB}^T\Omega_{AA}\Sigma_{AB}\Sigma_{BB}^{-1} \\
&= (\Sigma_{BB} - \Sigma_{AB}^T\Sigma_{AA}^{-1}\Sigma_{AB})^{-1}
\end{aligned}
$$

Gaussian Marginal

- Let $\tilde{x}_A := x_A - \mu_A$ and $\tilde{x}_B := x_B - \mu_B$ and consider:

$$p(x_A) = \int \phi\left(\begin{pmatrix} x_A \\ x_B \end{pmatrix}; \begin{pmatrix} \mu_A \\ \mu_B \end{pmatrix}, \begin{bmatrix} \Sigma_{AA} & \Sigma_{AB} \\ \Sigma_{AB}^T & \Sigma_{BB} \end{bmatrix}\right) dx_B$$

$$= \frac{1}{(2\pi)^{\frac{n+m}{2}} |\Sigma|^{1/2}} \int \exp\left(-\frac{1}{2}\left(\tilde{x}_A^T \Omega_{AA} \tilde{x}_A + 2\tilde{x}_A^T \Omega_{AB} \tilde{x}_B + \tilde{x}_B^T \Omega_{BB} \tilde{x}_B\right)\right) dx_B$$

$$\overset{Sq.Comp.}{=\!=\!=} \kappa \int \exp\left(-\frac{1}{2}\left[(\tilde{x}_B + \Omega_{BB}^{-1}\Omega_{AB}^T \tilde{x}_A)^T \Omega_{BB} (\tilde{x}_B + \Omega_{BB}^{-1}\Omega_{AB}^T \tilde{x}_A) - \tilde{x}_A^T \Omega_{AB} \Omega_{BB}^{-1} \Omega_{AB}^T \tilde{x}_A + \tilde{x}_A^T \Omega_{AA} \tilde{x}_A\right]\right) dx_B$$

$$\underset{\Sigma_{AA}^{-1} = \Omega_{AA} - \Omega_{AB}\Omega_{BB}^{-1}\Omega_{AB}^T}{\overset{\int \phi(x;\mu,\Sigma)dx = 1}{=\!=\!=}} \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma_{AA}|^{1/2}} \exp\left(-\frac{1}{2}\tilde{x}_A^T \Sigma_{AA}^{-1} \tilde{x}_A\right)$$

$$\Rightarrow \boxed{x_A \sim \mathcal{N}(\mu_A, \Sigma_{AA})}$$

Gaussian Conditional

$$p(x_A \mid x_B) = \frac{p(x_A, x_B)}{p(x_B)} = \frac{\frac{1}{\sqrt{(2\pi)^{n+m}|\Sigma|}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)}}{\frac{1}{\sqrt{(2\pi)^{m}|\Sigma_{BB}|}} e^{-\frac{1}{2}(x_B-\mu_B)^T \Sigma_{BB}^{-1}(x_B-\mu_B)}}$$

$$= \frac{1}{\sqrt{(2\pi)^n |\Sigma| / |\Sigma_{BB}|}} e^{-\frac{1}{2}\left((x-\mu)^T \Sigma^{-1}(x-\mu) - (x_B-\mu_B)^T \Sigma_{BB}^{-1}(x_B-\mu_B)\right)}$$

- Consider the exponent:

$$\tilde{x}_A^T \Omega_{AA} \tilde{x}_A + \tilde{x}_A^T \Omega_{AB} \tilde{x}_B + \tilde{x}_B^T \Omega_{AB}^T \tilde{x}_A + \tilde{x}_B^T \Omega_{BB} \tilde{x}_B - \tilde{x}_B^T \Sigma_{BB}^{-1} \tilde{x}_B$$

$$= \tilde{x}_A^T \Omega_{AA} \tilde{x}_A - 2\tilde{x}_A^T \Omega_{AA} \Sigma_{AB} \Sigma_{BB}^{-1} \tilde{x}_B + \tilde{x}_B^T \Sigma_{BB}^{-1} \Sigma_{AB}^T \Omega_{AA} \Sigma_{AB} \Sigma_{BB}^{-1} \tilde{x}_B$$

$$= (\tilde{x}_A - \Sigma_{AB}\Sigma_{BB}^{-1}\tilde{x}_B)^T \Omega_{AA} (\tilde{x}_A - \Sigma_{AB}\Sigma_{BB}^{-1}\tilde{x}_B)$$

$$\Rightarrow \boxed{x_A \mid x_B \sim \mathcal{N}\left(\mu_A + \Sigma_{AB}\Sigma_{BB}^{-1}(x_B - \mu_B), \Sigma_{AA} - \Sigma_{AB}\Sigma_{BB}^{-1}\Sigma_{AB}^T\right)}$$

Kalman Filter

- A **Kalman filter** is a Bayes filter for which:
  - The prior pdf $p_{0|0}$ is Gaussian
  - The motion model is linear in the state and affected by Gaussian noise
  - The observation model is linear in the state and affected by Gaussian noise
  - The process noise w_t and measurement noise v_t are independent of each other, of the state x_t and across time

- **Prior**: $x_t \mid z_{0:t}, u_{0:t-1} \sim \mathcal{N}(\mu_{t|t}, \Sigma_{t|t})$

- **Motion Model**:

$$x_{t+1} = a(x_t, u_t, w_t) := Ax_t + Bu_t + w_t, \quad w_t \sim \mathcal{N}(0, W)$$

$$x_{t+1} \mid x_t, u_t \sim \mathcal{N}(Ax_t + Bu_t, W), \quad A \in \mathbb{R}^{d \times d}, B \in \mathbb{R}^{d \times d_u}, \; W \in \mathbb{R}^{d \times d}$$

- **Observation Model**:

$$z_t = h(x_t, v_t) := Hx_t + v_t, \quad v_t \sim \mathcal{N}(0, V)$$

$$z_t \mid x_t \sim \mathcal{N}(Hx_t, V), \quad H \in \mathbb{R}^{d_z \times d}, \; V \in \mathbb{R}^{d_z \times d_z}$$

Kalman Filter Prediction

$$p_{t+1|t}(x) = \int p_a(x \mid s, u_t) p_{t|t}(s) ds = \int \phi(x; As + Bu_t, W)\phi(s; \mu_{t|t}, \Sigma_{t|t}) ds$$

$$= \kappa_{t|t} \int \exp\left\{-\frac{1}{2}(x - As - Bu_t)^T W^{-1}(x - As - Bu_t)\right\} \times$$

$$\exp\left\{-\frac{1}{2}(s - \mu_{t|t})^T \Sigma_{t|t}^{-1}(s - \mu_{t|t})\right\} ds$$

$$= \kappa_{t|t} \int \exp\left\{-\frac{1}{2}\left(s^T(A^T W^{-1} A + \Sigma_{t|t}^{-1})s - 2(\Sigma_{t|t}^{-1}\mu_{t|t} + A^T W^{-1}(x - Bu_t))^T s + \dots\right)\right\} ds$$

$$\overset{Sq.Comp.}{\underset{Inv.Lemma}{=\!=\!=}} \phi(x; A\mu_{t|t} + Bu_t, A\Sigma_{t|t}A^T + W)$$

$$p_{t+1|t}(x) = \int \phi(x; As + Bu_t, W)\phi(s; \mu_{t|t}, \Sigma_{t|t}) ds$$
$$= \phi(x; A\mu_{t|t} + Bu_t, A\Sigma_{t|t}A^T + W)$$

Kalman Filter Prediction (easy version)

- Motion model with given prior:

$$x_{t+1} = Ax_t + Bu_t + w_t, \qquad w_t \sim \mathcal{N}(0, W), \quad x_t \sim \mathcal{N}(\mu_{t|t}, \Sigma_{t|t})$$

- Since w_t and x_t are independent and the Gaussian distribution is stable, we know that the distribution of x_{t+1} is Gaussian: $\mathcal{N}(\mu_{t+1|t}, \Sigma_{t+1|t})$

- We just need to compute its mean and covariance:

$$\mu_{t+1|t} = \mathbb{E}\left[Ax_t + Bu_t + w_t\right] = A\mathbb{E}[x_t] + Bu_t + \mathbb{E}[w_t] = A\mu_{t|t} + Bu_t$$

$$\begin{aligned}
\mathbb{E}\left[x_{t+1}x_{t+1}^T\right] &= \mathbb{E}\left[(Ax_t + Bu_t + w_t)(Ax_t + Bu_t + w_t)^T\right] \\
&= A\mathbb{E}\left[x_t x_t^T\right]A^T + A\mathbb{E}\left[x_t\right]u_t^T B^T + A\mathbb{E}\left[x_t w_t^T\right] \\
&\quad + Bu_t\mathbb{E}\left[x_t^T\right]A^T + Bu_t u_t^T B^T + Bu_t\mathbb{E}\left[w_t^T\right] \\
&\quad + \mathbb{E}\left[w_t x_t^T\right]A^T + \mathbb{E}\left[w_t\right]u_t^T B^T + \mathbb{E}\left[w_t w_t^T\right] \\
&= A\left(\Sigma_{t|t} + \mu_{t|t}\mu_{t|t}^T\right)A^T + A\mu_{t|t}u_t^T B^T + Bu_t\mu_{t|t}^T A^T + Bu_t u_t^T B^T + W
\end{aligned}$$

Kalman Filter Update

$$p_{t+1|t+1}(x) = \frac{p(z_{t+1} \mid x) p_{t+1|t}(x)}{p(z_{t+1} \mid z_{0:t}, u_{0:t})} = \frac{\phi(z_{t+1}; Hx, V)\phi(x; \mu_{t+1|t}, \Sigma_{t+1|t})}{\int \phi(z_{t+1}; Hs, V)\phi(s; \mu_{t+1|t}, \Sigma_{t+1|t}) ds}$$

$$= \frac{\kappa_{t+1}}{\eta_{t+1}} \exp\left\{-\frac{1}{2}(z_{t+1} - Hx)^T V^{-1}(z_{t+1} - Hx)\right\} \exp\left\{-\frac{1}{2}(x - \mu_{t+1|t})^T \Sigma_{t+1|t}^{-1}(x - \mu_{t+1|t})\right\}$$

$$= \frac{\kappa_{t+1}}{\eta_{t+1}} \exp\left\{-\frac{1}{2}\left(x^T(H^T V^{-1} H + \Sigma_{t+1|t}^{-1})x + (H^T V^{-1} z_{t+1} + \Sigma_{t+1|t}^{-1}\mu_{t+1|t})^T x + \ldots\right)\right\}$$

$$\overset{Sq.Comp.}{=\!=\!=} \phi\left(x; (H^T V^{-1} H + \Sigma_{t+1|t}^{-1})^{-1}(H^T V^{-1} z_{t+1} + \Sigma_{t+1|t}^{-1}\mu_{t+1|t}), (H^T V^{-1} H + \Sigma_{t+1|t}^{-1})^{-1}\right)$$

$$\overset{Inv.Lemma}{=\!=\!=} \phi\left(x; \mu_{t+1|t} + K_{t+1|t}(z_{t+1} - H\mu_{t+1|t}), (I - K_{t+1|t}H)\Sigma_{t+1|t}\right)$$

- **Kalman gain**: $\boxed{K_{t+1|t} := \Sigma_{t+1|t} H^T \left(H\Sigma_{t+1|t} H^T + V\right)^{-1}}$

$$\boxed{\begin{aligned} p_{t+1|t+1}(x) &= \frac{\phi(z_{t+1}; Hx, V)\phi(x; \mu_{t+1|t}, \Sigma_{t+1|t})}{\int \phi(z_{t+1}; Hs, V)\phi(s; \mu_{t+1|t}, \Sigma_{t+1|t}) ds} \\ &= \phi\left(x; \mu_{t+1|t} + K_{t+1|t}(z_{t+1} - H\mu_{t+1|t}), (I - K_{t+1|t}H)\Sigma_{t+1|t}\right) \end{aligned}}$$

Kalman Filter Update (easy version)

- Observation model with given prior:

$$z_{t+1} = Hx_{t+1} + v_{t+1}, \qquad v_{t+1} \sim \mathcal{N}(0, V), \quad x_{t+1} \sim \mathcal{N}(\mu_{t+1|t}, \Sigma_{t+1|t})$$

- The joint distribution of x_{t+1} and z_{t+1} is Gaussian:

$$\begin{pmatrix} x_{t+1} \\ z_{t+1} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_{t+1|t} \\ H\mu_{t+1|t} \end{pmatrix}, \begin{bmatrix} \Sigma_{t+1|t} & ? \\ ?^T & H\Sigma_{t+1|t}H^T + V \end{bmatrix} \right)$$

$$\begin{aligned} ? &= \mathbb{E}\left[\left(x_{t+1} - \mu_{t+1|t}\right) \left(z_{t+1} - H\mu_{t+1|t}\right)^T \right] \\ &= \mathbb{E}\left[\left(x_{t+1} - \mu_{t+1|t}\right) \left((x_{t+1} - \mu_{t+1|t})^T H^T + v_{t+1}^T \right) \right] = \Sigma_{t+1|t}H^T \end{aligned}$$

- The conditional distribution of $x_{t+1} \mid z_{t+1}$ is then also Gaussian:

$$\begin{aligned} x_{t+1} \mid z_{t+1} \sim \mathcal{N}\Big(&\mu_{t+1|t} + \Sigma_{t+1|t}H^T(H\Sigma_{t+1|t}H^T + V)^{-1}(z_{t+1} - H\mu_{t+1|t}), \\ &\Sigma_{t+1|t} - \Sigma_{t+1|t}H^T(H\Sigma_{t+1|t}H^T + V)^{-1}H\Sigma_{t+1|t} \Big) \end{aligned}$$

Kalman Filter (discrete time)

Prior: $x_t \mid z_{0:t}, u_{0:t-1} \sim \mathcal{N}(\mu_{t|t}, \Sigma_{t|t})$

Motion model: $x_{t+1} = Ax_t + Bu_t + w_t, \quad w_t \sim \mathcal{N}(0, W)$

Observation model: $z_t = Hx_t + v_t, \quad v_t \sim \mathcal{N}(0, V)$

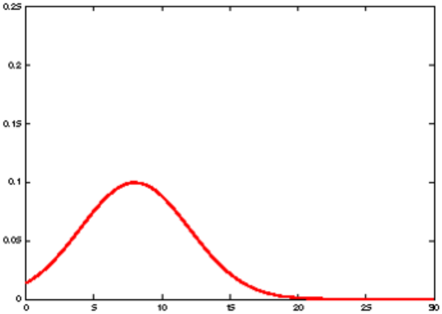
Prediction:
$$\mu_{t+1|t} = A\mu_{t|t} + Bu_t$$
$$\Sigma_{t+1|t} = A\Sigma_{t|t}A^T + W$$

Update:
$$\mu_{t+1|t+1} = \mu_{t+1|t} + K_{t+1|t}(z_{t+1} - H\mu_{t+1|t})$$
$$\Sigma_{t+1|t+1} = (I - K_{t+1|t}H)\Sigma_{t+1|t}$$

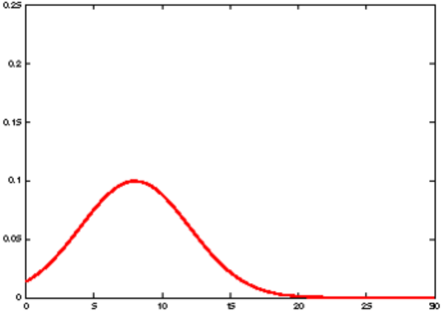
Kalman Gain: $K_{t+1|t} := \Sigma_{t+1|t}H^T \left(H\Sigma_{t+1|t}H^T + V\right)^{-1}$

Predict

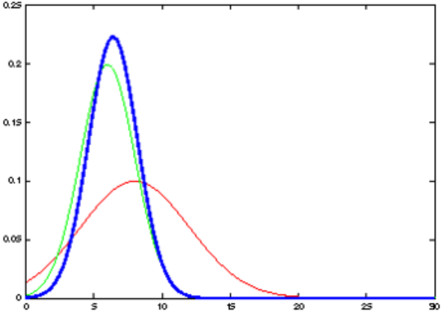
Update



Predict

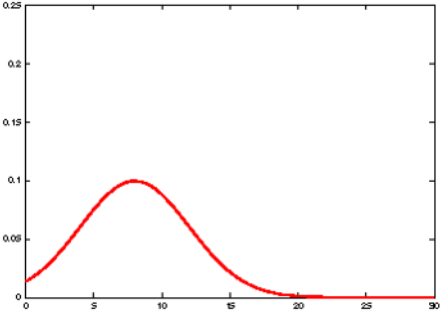


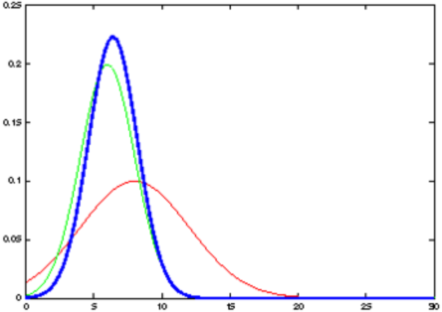
Update

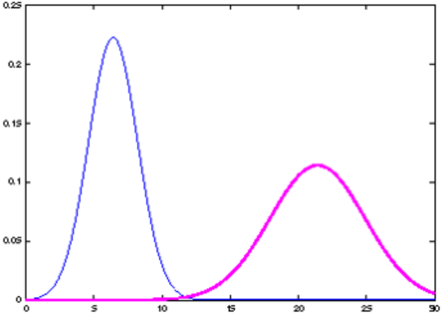


Predict

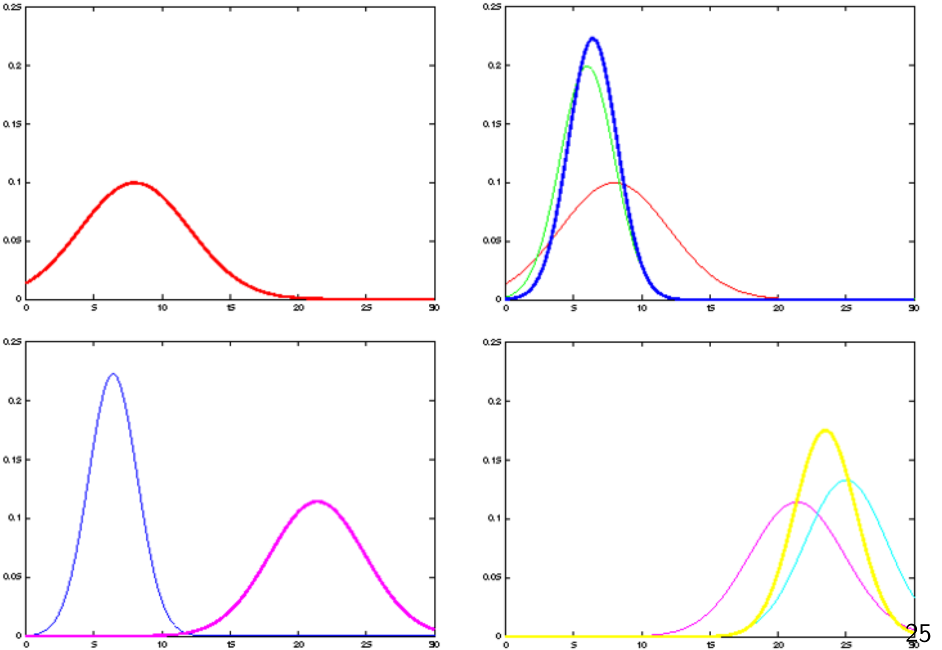
Update







Predict | Update



Kalman-Bucy Filter (continuous time)

Prior:	$x(0) \sim \mathcal{N}(\mu(0), \Sigma(0))$
Motion model:	$\dot{x}(t) = Ax(t) + Bu(t) + w(t)$
Observation model:	$z(t) = Hx(t) + v(t)$
Mean:	$\dot{\mu}(t) = A\mu(t) + Bu(t) + K(t)\,(z(t) - H\mu(t))$
Covariance:	$\dot{\Sigma}(t) = A\Sigma(t) + \Sigma(t)A^T + W - K(t)VK^T(t)$
Kalman Gain:	$K(t) = \Sigma(t)H^T V^{-1}$

Kalman Filter Comments

- **Efficient**: polynomial in measurement and state dim: $\boxed{O(d_z^{2.376} + d^2)}$

- **Optimal**: under linearity, Gaussianity, and independence assumptions with respect to the mean square error (MSE):
$\mathbb{E}\left[\|x_t - \mu_{t|t}\|_2^2\right] = \text{tr}(\Sigma_{t|t})$

- To deal with **unknown models** we can use EM to learn the dynamics model (A, B, W) and the measurement model (H, V)

- Given data $\{z_{0:T}, u_{0:T-1}\}$, apply EM with hidden variable x_T:
  - **E step**: Given initial parameter estimates $\omega^{(i)} := \{A^{(i)}, B^{(i)}, W^{(i)}, H^{(i)}, V^{(i)}\}$ calculate the likelihood of the hidden variable via the Kalman filter
  - **M step**: Optimize the parameters via MLE to obtain $\omega^{(i+1)}$ which explains the posterior distribution over x_T better

- Most robotic systems are **nonlinear**!