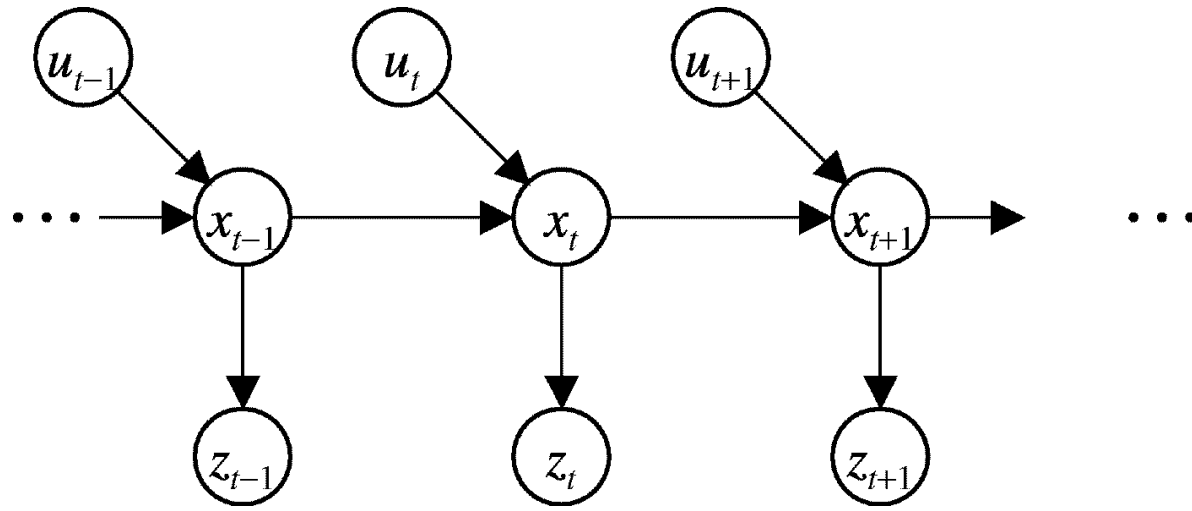


Probabilistic Reasoning Over Time

Bayes Filter Implementations

Gaussian filters

Markov Assumption



$$p(z_t | x_{0:t}, z_{1:t}, u_{1:t}) = p(z_t | x_t)$$

$$p(x_t | x_{1:t-1}, z_{1:t}, u_{1:t}) = p(x_t | x_{t-1}, u_t)$$

Underlying Assumptions

- Static world
- Independent noise
- Perfect model, no approximation errors

Application

- Filtering: Compute state given a series of evidences
- Prediction:
- Smoothing or hindsight
- Moste likely explanation

Application

- Filtering: Compute state given a series of evidences

$$p(x_t | e_{1:t})$$

- Prediction:

$$p(x_{t+k} | e_{1:t})$$

- Smoothing or hindsight

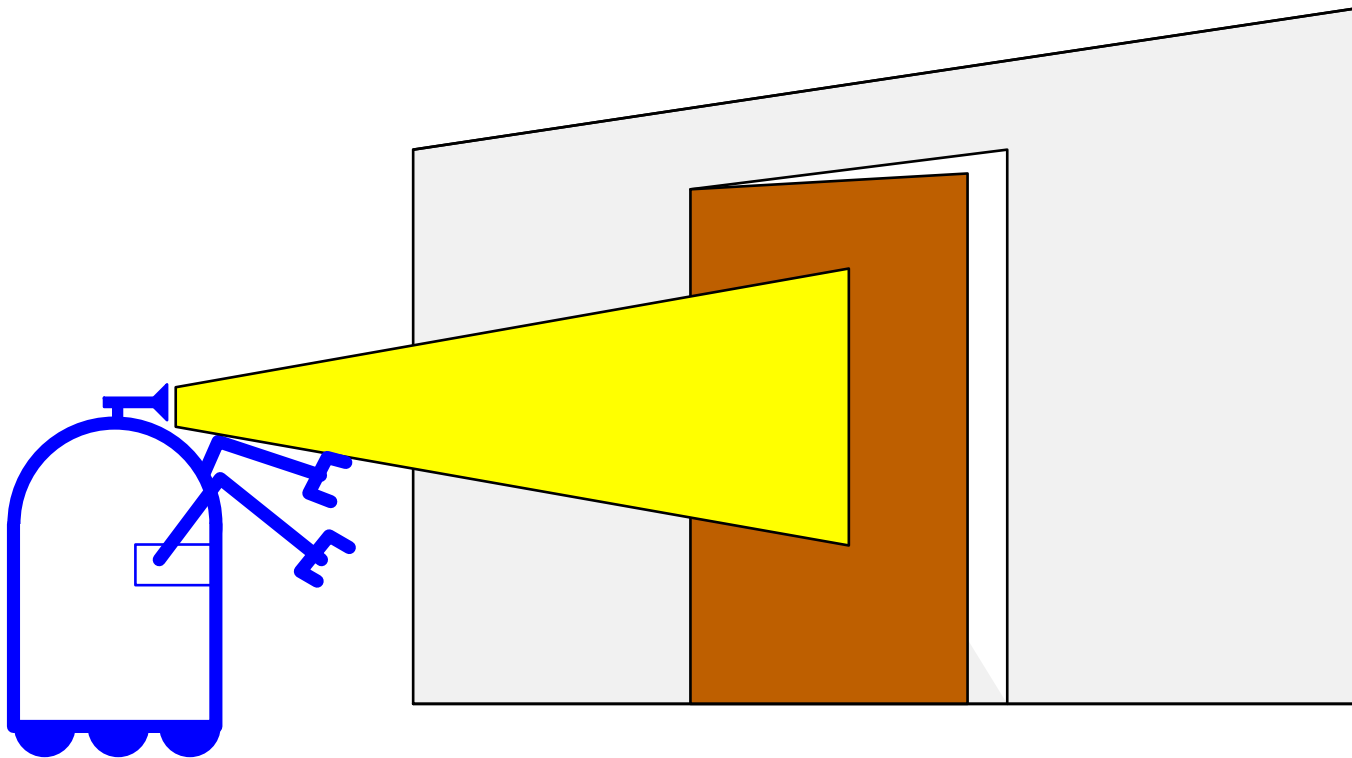
$$p(x_k | e_{1:t}) \quad k < t$$

- Moste likely explanation

$$p(x_{1:t} | e_{1:t})$$

Another Example: State Estimation

- Suppose a robot obtains measurement z
- What is $P(open/z)$?



Causal vs. Diagnostic Reasoning

- $P(open/z)$ is diagnostic.
- $P(z/open)$ is causal.
- Often causal knowledge is easier to obtain.
- Bayes rule allows **count frequencies!** knowledge:

$$P(open | z) = \frac{P(z | open)P(open)}{P(z)}$$

Example

- $P(z|open) = 0.6$ $P(z|\neg open) = 0.3$
- $P(open) = P(\neg open) = 0.5$

Example

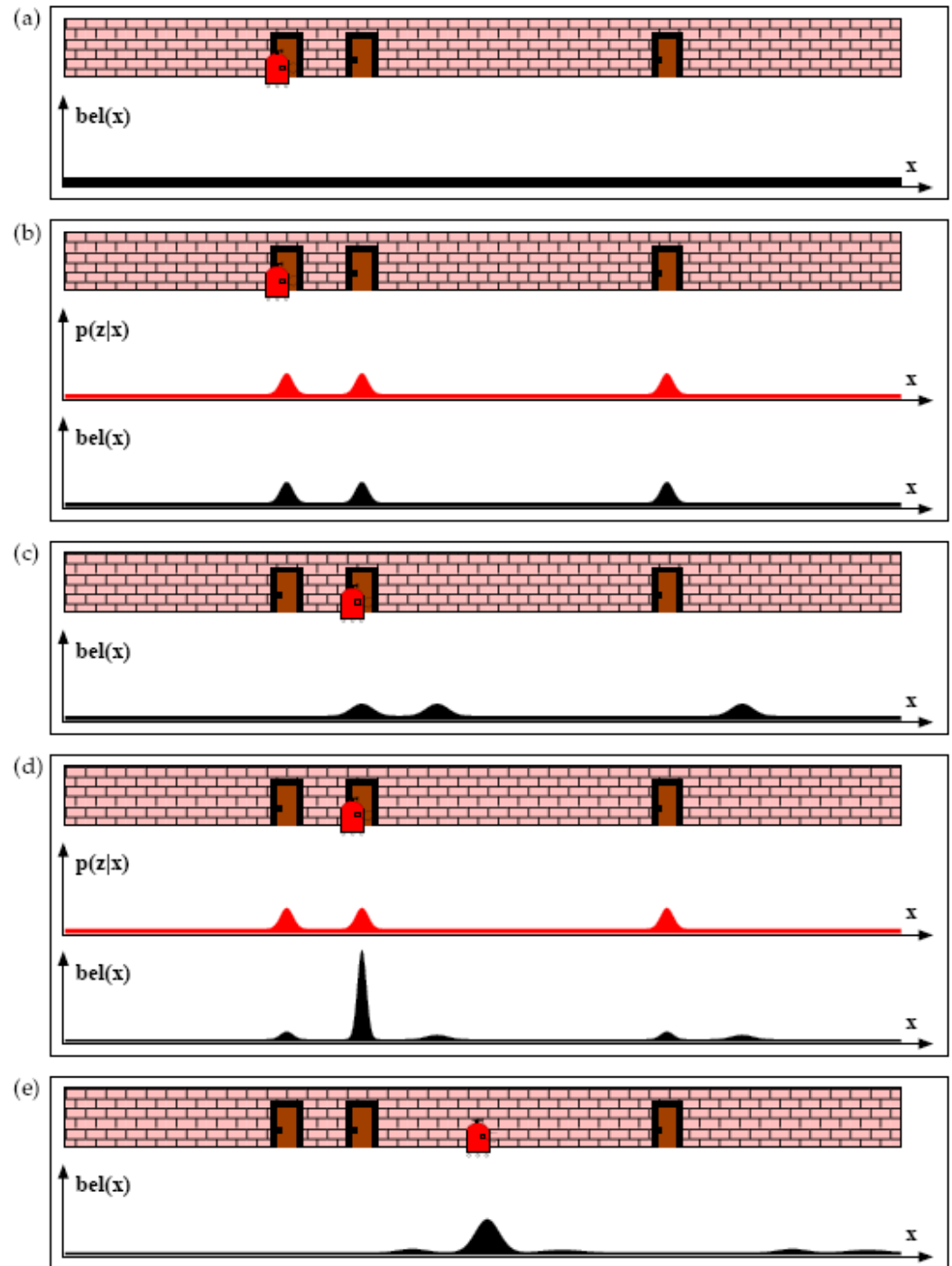
- $P(z|open) = 0.6$ $P(z|\neg open) = 0.3$
- $P(open) = P(\neg open) = 0.5$

$$P(open | z) = \frac{P(z | open)P(open)}{P(z | open)p(open) + P(z | \neg open)p(\neg open)}$$

$$P(open | z) = \frac{0.6 \cdot 0.5}{0.6 \cdot 0.5 + 0.3 \cdot 0.5} = \frac{2}{3} = 0.67$$

- z raises the probability that the door is open.

Combining Evidences



Combining Evidence

- Suppose our robot obtains another observation z_2 .
- How can we integrate this new information?
- More generally, how can we estimate $P(x | z_1 \dots z_n)$?

Recursive Bayesian Updating

$$P(x | z_1, \dots, z_n) = \frac{P(z_n | x, z_1, \dots, z_{n-1}) P(x | z_1, \dots, z_{n-1})}{P(z_n | z_1, \dots, z_{n-1})}$$

Recursive Bayesian Updating

$$P(x | z_1, \dots, z_n) = \frac{P(z_n | x, z_1, \dots, z_{n-1}) P(x | z_1, \dots, z_{n-1})}{P(z_n | z_1, \dots, z_{n-1})}$$

Markov assumption: z_n is independent of z_1, \dots, z_{n-1} if we know x .

$$\begin{aligned} P(x | z_1, \dots, z_n) &= \frac{P(z_n | x) P(x | z_1, \dots, z_{n-1})}{P(z_n | z_1, \dots, z_{n-1})} \\ &= \eta P(z_n | x) P(x | z_1, \dots, z_{n-1}) \\ &= \eta_{1\dots n} \prod_{i=1\dots n} P(z_i | x) P(x) \end{aligned}$$

$$P(A|B \wedge X) = \frac{P(B|A \wedge X)P(A \wedge X)}{P(B \wedge X)} = \frac{P(B|A \wedge X)P(A|X)P(X)}{P(B|X)P(X)} = \frac{P(B|A \wedge X)P(A|X)}{P(B|X)}$$

Example: Second Measurement

- $P(z_2|open) = 0.5$ $P(z_2|\neg open) = 0.6$
- $P(open|z_1) = 2/3$

Example: Second Measurement

- $P(z_2|open) = 0.5$ $P(z_2|\neg open) = 0.6$
- $P(open|z_1) = 2/3$

$$\begin{aligned} P(open | z_2, z_1) &= \frac{P(z_2 | open) P(open | z_1)}{P(z_2 | open) P(open | z_1) + P(z_2 | \neg open) P(\neg open | z_1)} \\ &= \frac{\frac{1}{2} \cdot \frac{2}{3}}{\frac{1}{2} \cdot \frac{2}{3} + \frac{3}{5} \cdot \frac{1}{3}} = \frac{5}{8} = 0.625 \end{aligned}$$

- z_2 lowers the probability that the door is open.

z = observation
 u = action
 x = state

Bayes Filters

$$\boxed{Bel(x_t)} = P(x_t | u_1, z_1, \dots, u_t, z_t)$$

Bayes $= \eta P(z_t | x_t, u_1, z_1, \dots, u_t) P(x_t | u_1, z_1, \dots, u_t)$

Markov $= \eta P(z_t | x_t) P(x_t | u_1, z_1, \dots, u_t)$

Total prob. $= \eta P(z_t | x_t) \int P(x_t | u_1, z_1, \dots, u_t, x_{t-1})$
 $P(x_{t-1} | u_1, z_1, \dots, u_t) dx_{t-1}$

Markov $= \eta P(z_t | x_t) \int P(x_t | u_t, x_{t-1}) P(x_{t-1} | u_1, z_1, \dots, u_t) dx_{t-1}$

Markov $= \eta P(z_t | x_t) \int P(x_t | u_t, x_{t-1}) P(x_{t-1} | u_1, z_1, \dots, z_{t-1}) dx_{t-1}$

$$\boxed{= \eta P(z_t | x_t) \int P(x_t | u_t, x_{t-1}) Bel(x_{t-1}) dx_{t-1}}$$

Bayes Filter Algorithm

1. Algorithm **Bayes_filter**($Bel(x), d$):
2. $\eta=0$
3. If d is a **perceptual** data item z then
4. For all x do
5. $Bel'(x) = P(z | x)Bel(x)$
6. $\eta = \eta + Bel'(x)$
7. For all x do
8. $Bel'(x) = \eta^{-1}Bel'(x)$
9. Else if d is an **action** data item u then
10. For all x do $Bel'(x) = \int P(x | u, x') Bel(x') dx'$
- 11.

$$Bel(x_t) = \eta P(z_t | x_t) \int P(x_t | u_t, x_{t-1}) Bel(x_{t-1}) dx_{t-1}$$

Bayes Filters are Familiar!

$$Bel(x_t) = \eta P(z_t | x_t) \int P(x_t | u_t, x_{t-1}) Bel(x_{t-1}) dx_{t-1}$$

- Kalman filters
- Particle filters
- Hidden Markov models
- Dynamic Bayesian networks
- Partially Observable Markov Decision Processes (POMDPs)

Bayes Filter Reminder

- Prediction

$$\overline{bel}(x_t) = \int p(x_t | u_t, x_{t-1}) bel(x_{t-1}) dx_{t-1}$$

- Correction

$$bel(x_t) = \eta p(z_t | x_t) \overline{bel}(x_t)$$

A Simple Example

- Assume you and your friend riding a boat
 - You are not good in navigation using stars
 - He is good in navigation using stars
- You do a location measurement at time t (only one dimensional)
 - Your estimate $x(t) = z(\text{me})$
 - Variance $\sigma^2(t) = \sigma^2(z)$

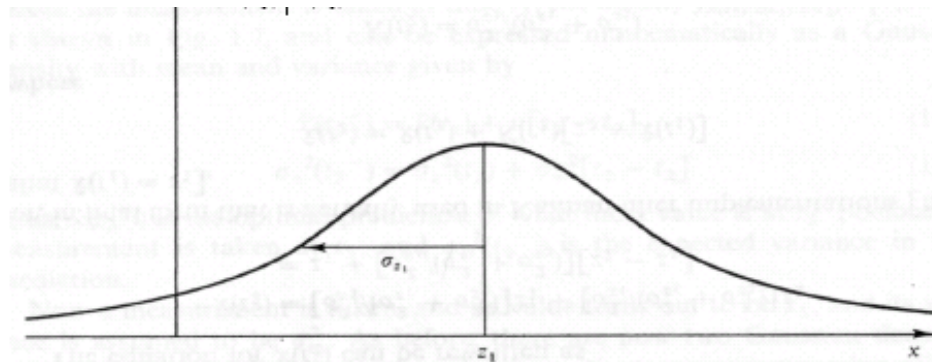
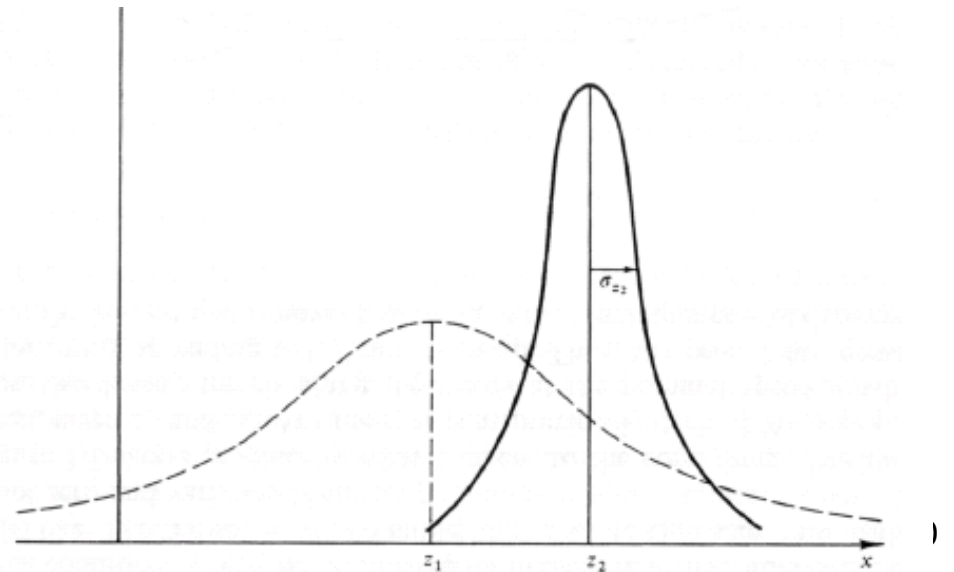


FIG. 1. 4 Conditional density of position based on measured value z_1 .

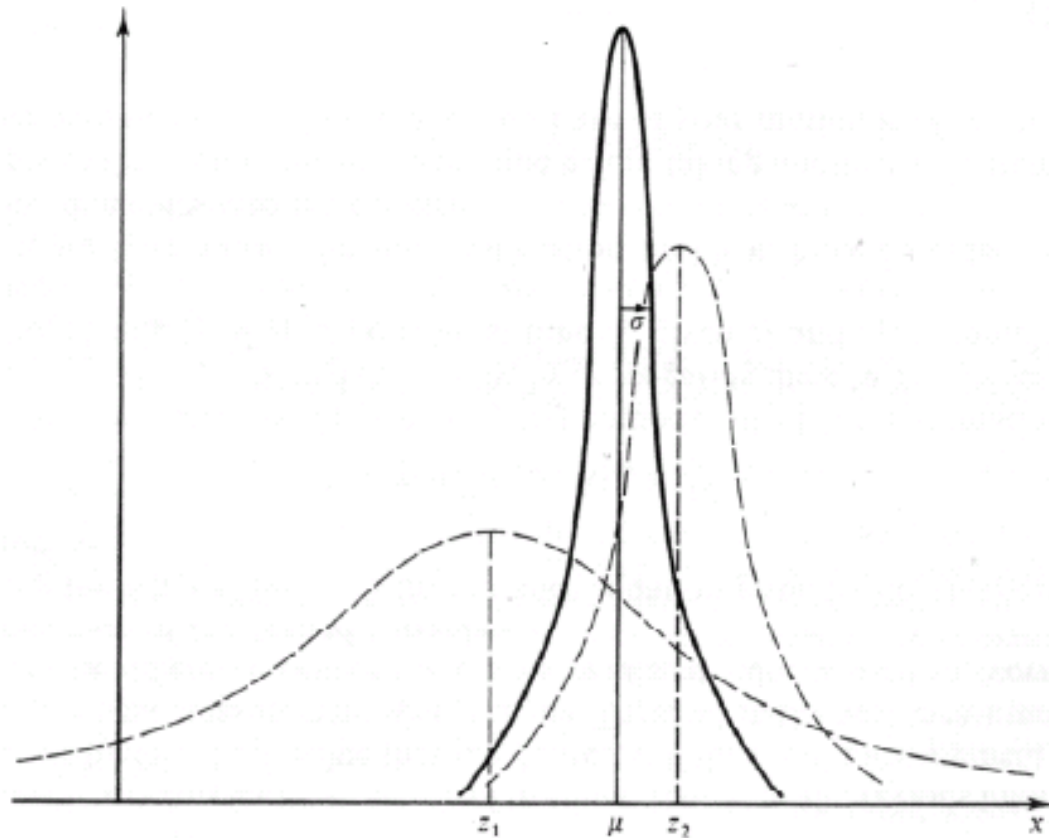
A Simple Example (2)

- Your friend makes another measurement at the same time
 - His estimate $x(t) = z(\text{my friend}) = z(m_f)$
 - Variance $\sigma^2(t) = \sigma^2(z(m_f))$



A Simple Example (2)

- Question: How to combine these measurements?

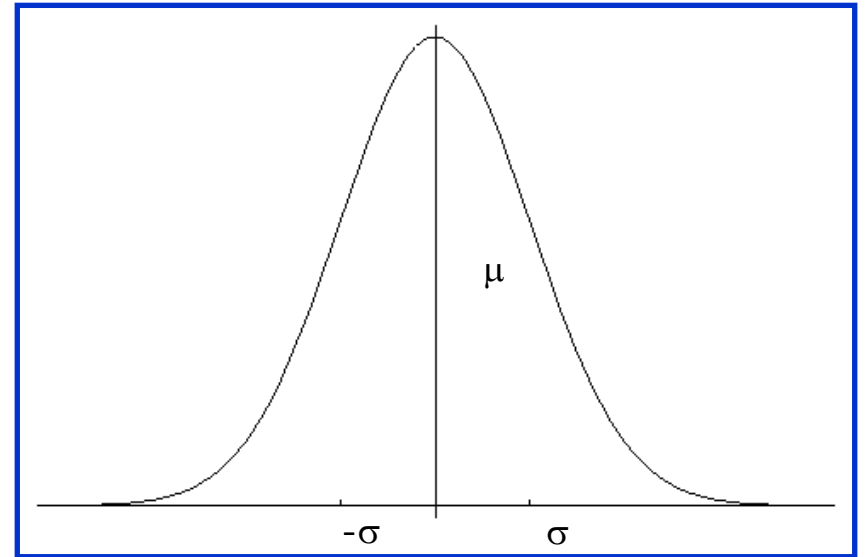


Gaussians

$$p(x) \sim N(\mu, \sigma^2):$$

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}}$$

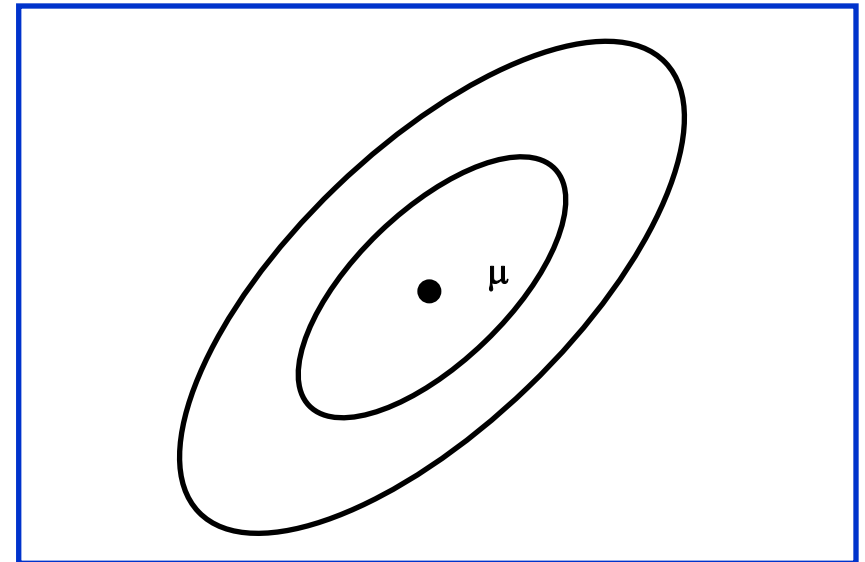
Univariate



$$p(\mathbf{x}) \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}):$$

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{d/2} |\boldsymbol{\Sigma}|^{1/2}} e^{-\frac{1}{2} (\mathbf{x}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})}$$

Multivariate



Discrete Kalman Filter

Estimates the state x of a discrete-time controlled process that is governed by the linear stochastic difference equation

$$x_t = A_t x_{t-1} + B_t u_t + \varepsilon_t$$

with a measurement

$$z_t = C_t x_t + \delta_t$$

Components of a Kalman Filter

$$x_t = A_t x_{t-1} + B_t u_t + \varepsilon_t$$

$$z_t = C_t x_t + \delta_t$$

 A_t

Matrix (nxn) that describes how the state evolves from $t-1$ to t without controls or noise.

 B_t

Matrix (n x l) that describes how the control u_t changes the state from $t-1$ to t .

 C_t

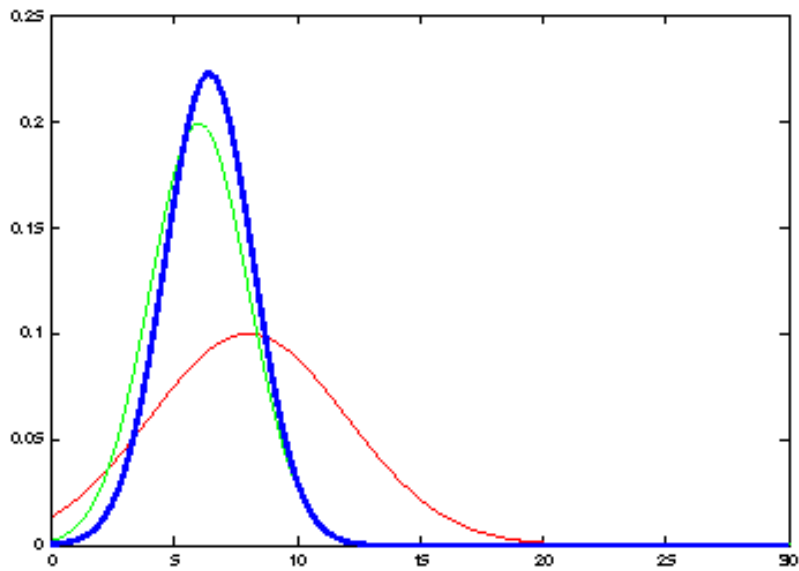
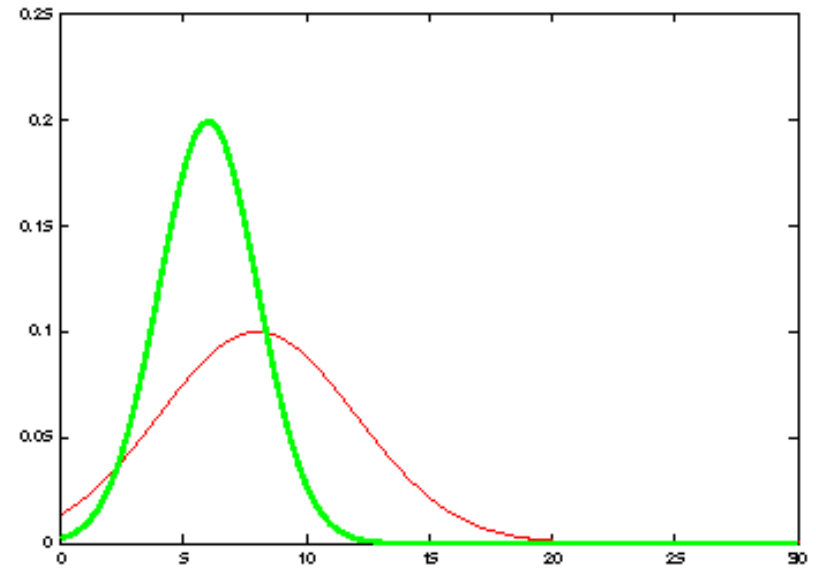
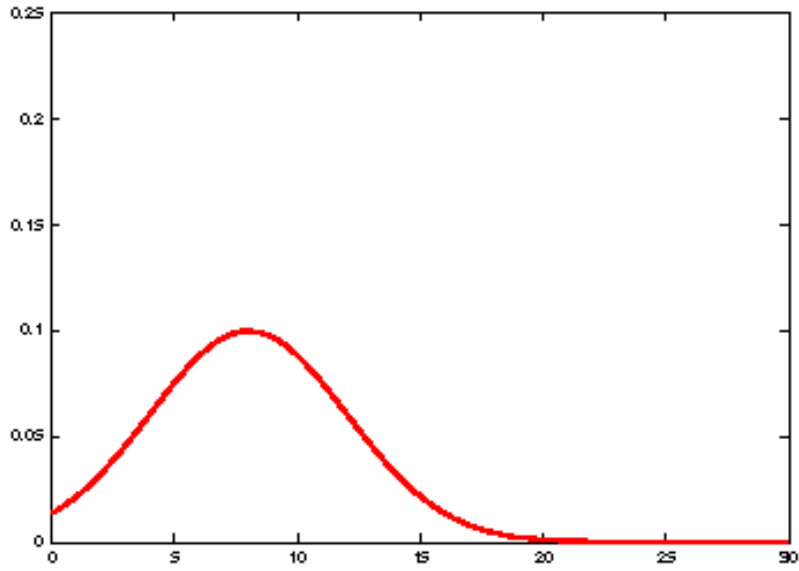
Matrix (k x n) that describes how to map the state x_t to an observation z_t .

 ε_t

Random variables representing the process and measurement noise that are assumed to be independent and normally distributed with covariance R_t and Q_t respectively.

 δ_t

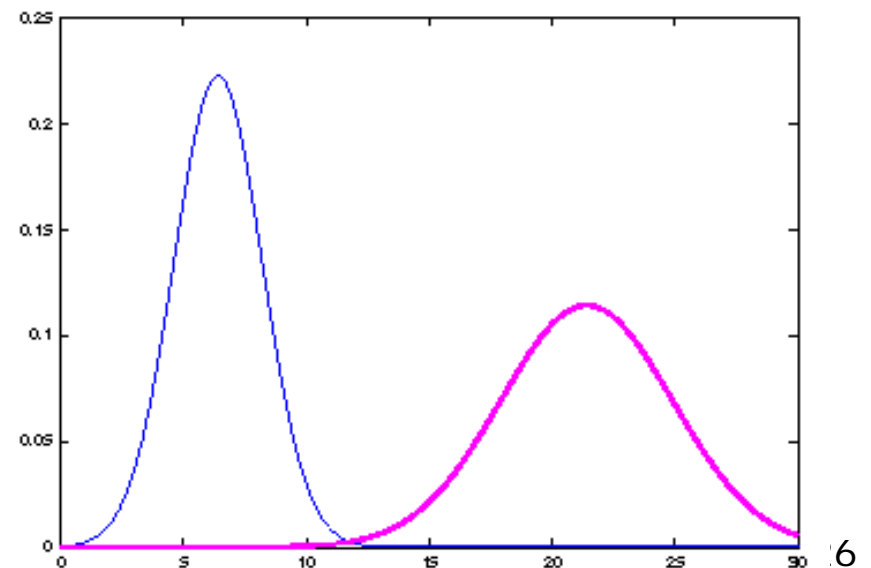
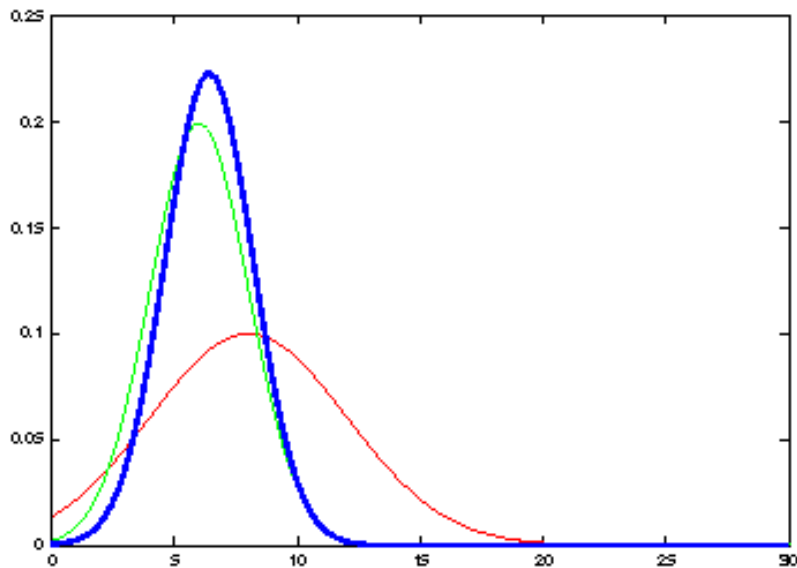
Kalman Filter Updates in 1D



Kalman Filter Updates in 1D: Prediction

$$\overline{bel}(x_t) = \begin{cases} \bar{\mu}_t = a_t \mu_{t-1} + b_t u_t \\ \bar{\sigma}_t^2 = a_t^2 \sigma_{t-1}^2 + \sigma_{act,t}^2 \end{cases}$$

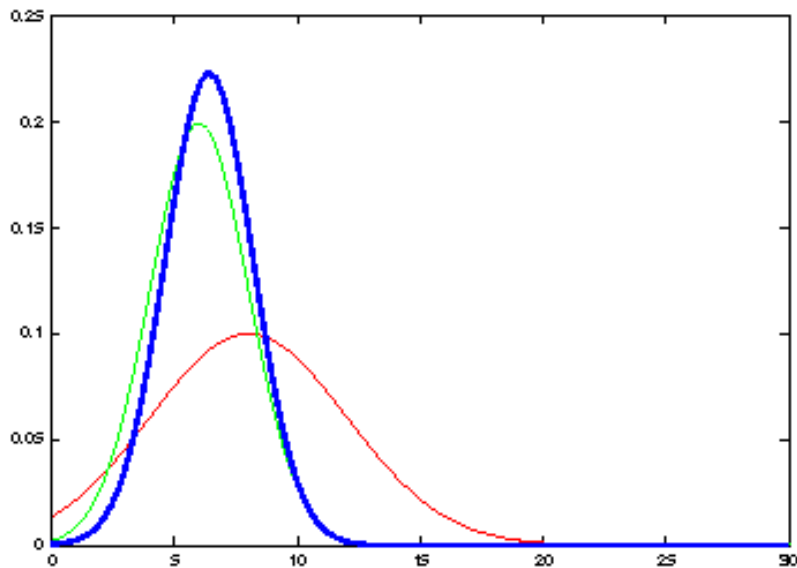
$$\overline{bel}(x_t) = \begin{cases} \bar{\mu}_t = A_t \mu_{t-1} + B_t u_t \\ \bar{\Sigma}_t = A_t \Sigma_{t-1} A_t^T + R_t \end{cases}$$



Kalman Filter Updates in 1D: Correction

$$bel(x_t) = \begin{cases} \mu_t = \bar{\mu}_t + K_t(z_t - \bar{\mu}_t) \\ \sigma_t^2 = (1 - K_t)\bar{\sigma}_t^2 \end{cases} \quad \text{with} \quad K_t = \frac{\bar{\sigma}_t^2}{\bar{\sigma}_t^2 + \bar{\sigma}_{obs,t}^2}$$

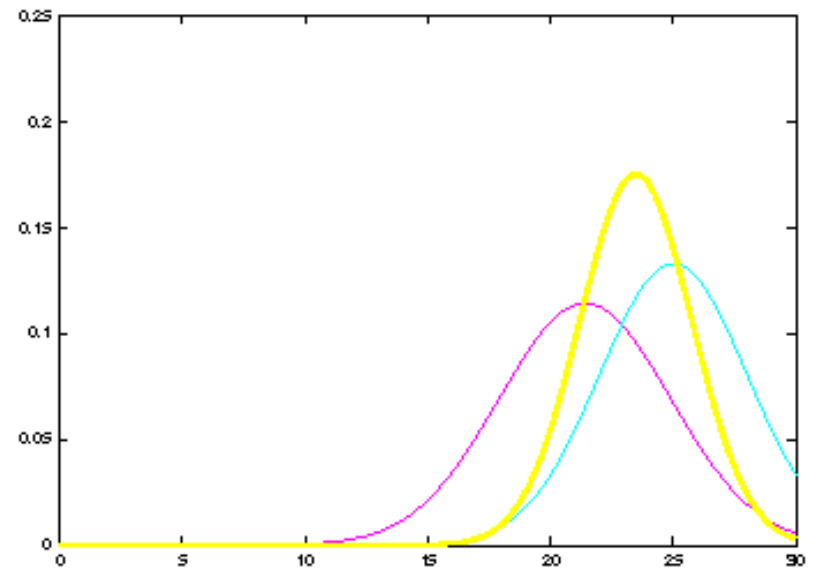
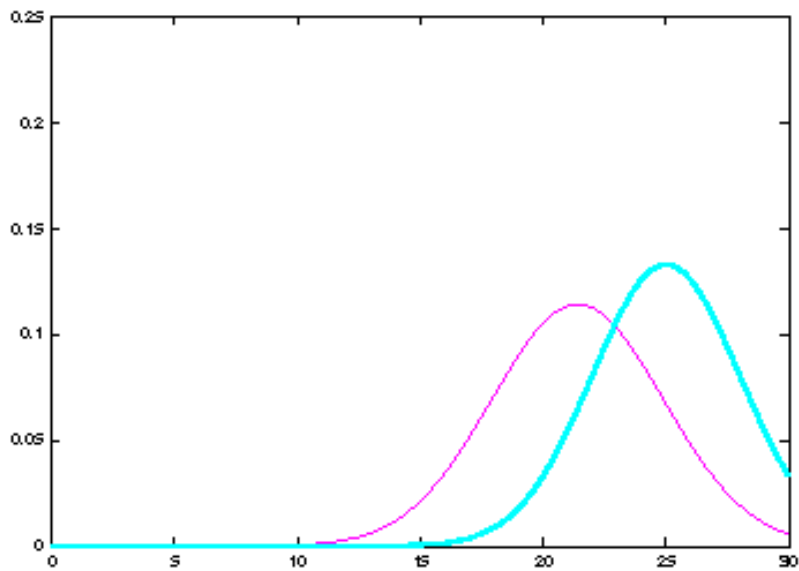
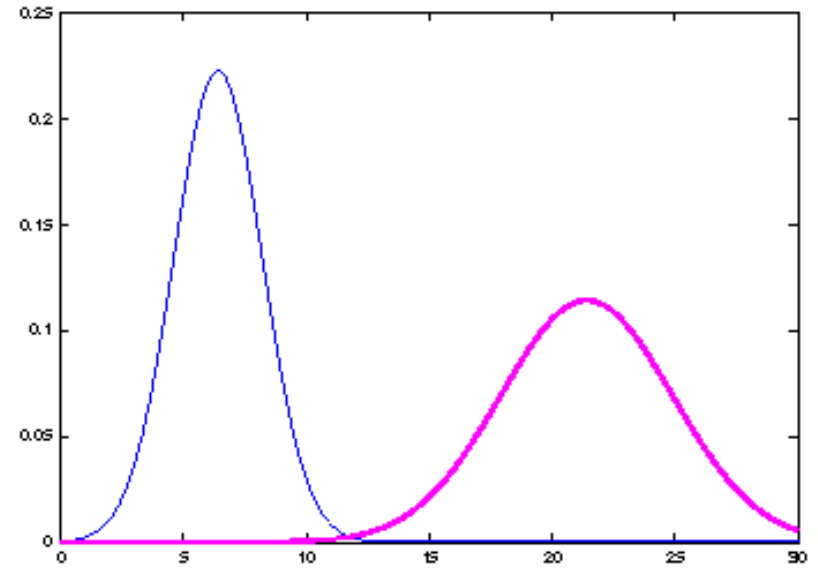
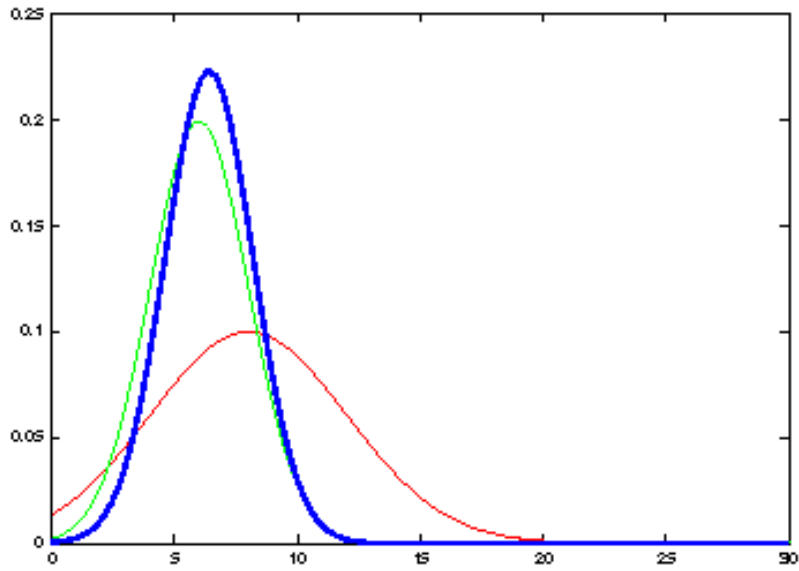
$$bel(x_t) = \begin{cases} \mu_t = \bar{\mu}_t + K_t(z_t - C_t\bar{\mu}_t) \\ \Sigma_t = (I - K_tC_t)\bar{\Sigma}_t \end{cases} \quad \text{with} \quad K_t = \bar{\Sigma}_t C_t^T (C_t \bar{\Sigma}_t C_t^T + Q_t)^{-1}$$



Estimated mean

Updated mean

Kalman Filter Updates



Linear Gaussian Systems: Initialization

- Initial belief is normally distributed:

$$bel(x_0) = N(x_0; \mu_0, \Sigma_0)$$

Linear Gaussian Systems: Dynamics

- Dynamics are linear function of state and control plus additive noise:

$$x_t = A_t x_{t-1} + B_t u_t + \varepsilon_t$$

$$p(x_t | u_t, x_{t-1}) = N(x_t; A_t x_{t-1} + B_t u_t, R_t)$$

$$\begin{array}{ccc} \overline{bel}(x_t) = \int p(x_t | u_t, x_{t-1}) & & bel(x_{t-1}) dx_{t-1} \\ \Downarrow & & \Downarrow \\ \sim N(x_t; A_t x_{t-1} + B_t u_t, R_t) & \sim & N(x_{t-1}; \mu_{t-1}, \Sigma_{t-1}) \end{array}$$

Linear Gaussian Systems: Observations

- Observations are linear function of state plus additive noise:

$$z_t = C_t x_t + \delta_t$$

$$p(z_t | x_t) = N(z_t; C_t x_t, Q_t)$$

$$\begin{array}{ccc} \text{bel}(x_t) = \eta & p(z_t | x_t) & \overline{\text{bel}}(x_t) \\ & \Downarrow & \Downarrow \\ & \sim N(z_t; C_t x_t, Q_t) & \sim N(x_t; \overline{\mu}_t, \overline{\Sigma}_t) \end{array}$$

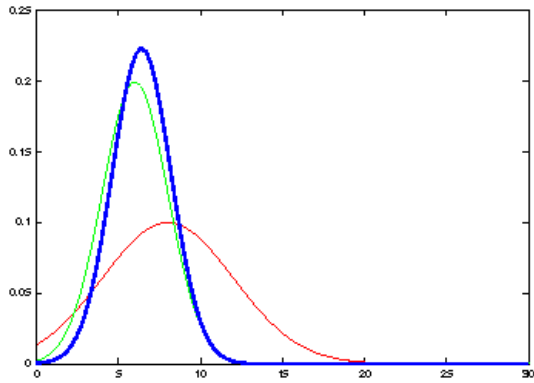
Linear Gaussian Systems: Observations

$$\begin{aligned}
 \text{bel}(x_t) &= \eta \, p(z_t | x_t) && \overline{\text{bel}}(x_t) \\
 &\Downarrow && \Downarrow \\
 &\sim N(z_t; C_t x_t, Q_t) && \sim N(x_t; \bar{\mu}_t, \bar{\Sigma}_t) \\
 &\Downarrow && \\
 \text{bel}(x_t) &= \eta \exp\left\{-\frac{1}{2}(z_t - C_t x_t)^T Q_t^{-1}(z_t - C_t x_t)\right\} \exp\left\{-\frac{1}{2}(x_t - \bar{\mu}_t)^T \bar{\Sigma}_t^{-1}(x_t - \bar{\mu}_t)\right\} \\
 \\
 \text{bel}(x_t) &= \begin{cases} \mu_t = \bar{\mu}_t + K_t(z_t - C_t \bar{\mu}_t) \\ \Sigma_t = (I - K_t C_t) \bar{\Sigma}_t \end{cases} && \text{with } K_t = \bar{\Sigma}_t C_t^T (C_t \bar{\Sigma}_t C_t^T + Q_t)^{-1}
 \end{aligned}$$

Kalman Filter Algorithm

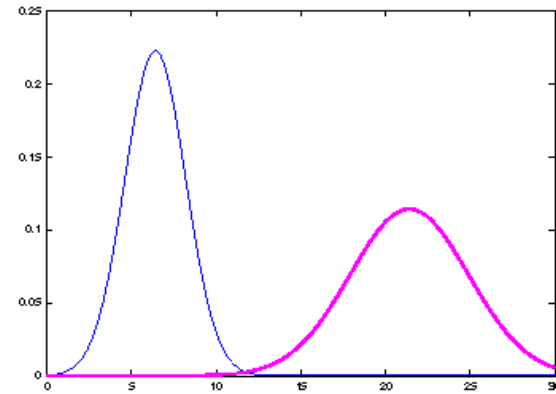
1. Algorithm **Kalman_filter**($\mu_{t-1}, \Sigma_{t-1}, u_t, z_t$):
2. Prediction:
3. $\bar{\mu}_t = A_t \mu_{t-1} + B_t u_t$
4. $\bar{\Sigma}_t = A_t \Sigma_{t-1} A_t^T + R_t$
5. Correction:
6. $K_t = \bar{\Sigma}_t C_t^T (C_t \bar{\Sigma}_t C_t^T + Q_t)^{-1}$
7. $\mu_t = \bar{\mu}_t + K_t (z_t - C_t \bar{\mu}_t)$
8. $\Sigma_t = (I - K_t C_t) \bar{\Sigma}_t$
9. Return μ_t, Σ_t

The Prediction-Correction-Cycle

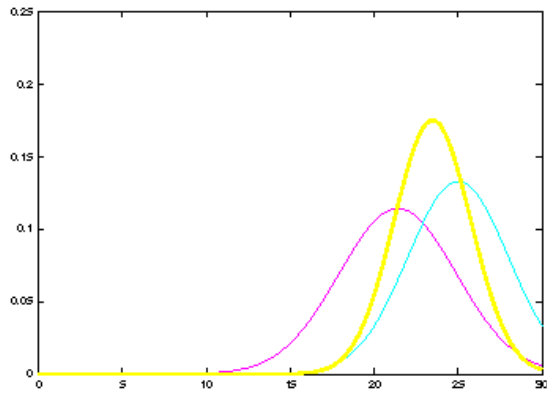


$$\overline{bel}(x_t) = \begin{cases} \bar{\mu}_t = a_t \mu_{t-1} + b_t u_t \\ \bar{\sigma}_t^2 = a_t^2 \sigma_{t-1}^2 + \sigma_{act,t}^2 \end{cases}$$

$$\overline{bel}(x_t) = \begin{cases} \bar{\mu}_t = A_t \mu_{t-1} + B_t u_t \\ \bar{\Sigma}_t = A_t \Sigma_{t-1} A_t^T + R_t \end{cases}$$

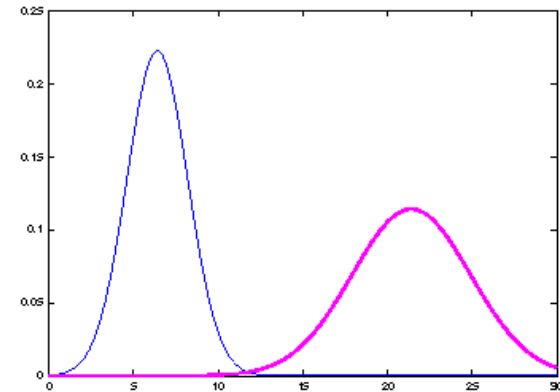


The Prediction-Correction-Cycle



$$bel(x_t) = \begin{cases} \mu_t = \bar{\mu}_t + K_t(z_t - \bar{\mu}_t) \\ \sigma_t^2 = (1 - K_t)\bar{\sigma}_t^2, K_t = \frac{\bar{\sigma}_t^2}{\bar{\sigma}_t^2 + \bar{\sigma}_{obs,t}^2} \end{cases}$$

$$bel(x_t) = \begin{cases} \mu_t = \bar{\mu}_t + K_t(z_t - C_t\bar{\mu}_t) \\ \Sigma_t = (I - K_tC_t)\bar{\Sigma}_t, K_t = \bar{\Sigma}_tC_t^T(C_t\bar{\Sigma}_tC_t^T + Q_t)^{-1} \end{cases}$$



The Prediction-Correction-Cycle



$$bel(x_t) = \begin{cases} \mu_t = \bar{\mu}_t + K_t(z_t - \bar{\mu}_t) \\ \sigma_t^2 = (1 - K_t)\bar{\sigma}_t^2 \end{cases}, K_t = \frac{\bar{\sigma}_t^2}{\bar{\sigma}_t^2 + \bar{\sigma}_{obs,t}^2}$$

$$\overline{bel}(x_t) = \begin{cases} \bar{\mu}_t = a_t \mu_{t-1} + b_t u_t \\ \bar{\sigma}_t^2 = a_t^2 \sigma_{t-1}^2 + \sigma_{act,t}^2 \end{cases}$$

$$bel(x_t) = \begin{cases} \mu_t = \bar{\mu}_t + K_t(z_t - C_t \bar{\mu}_t) \\ \Sigma_t = (I - K_t C_t) \bar{\Sigma}_t \end{cases}, K_t = \bar{\Sigma}_t C_t^T (C_t \bar{\Sigma}_t C_t^T + Q_t)^{-1}$$

$$\overline{bel}(x_t) = \begin{cases} \bar{\mu}_t = A_t \mu_{t-1} + B_t u_t \\ \bar{\Sigma}_t = A_t \Sigma_{t-1} A_t^T + R_t \end{cases}$$



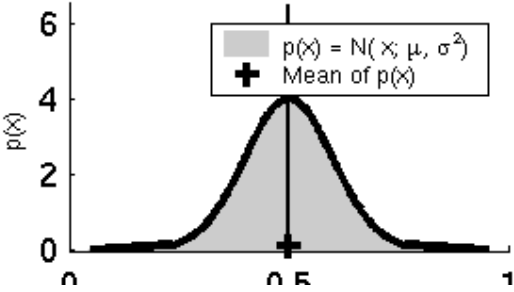
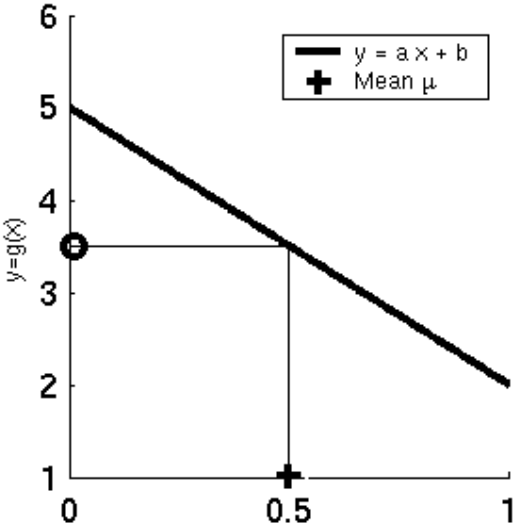
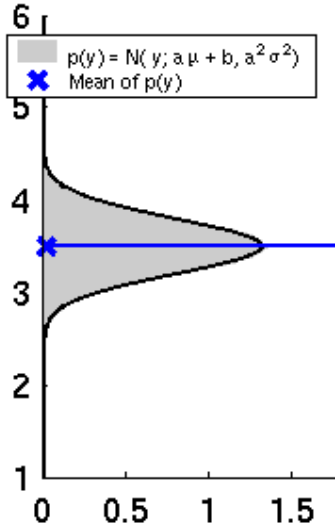
Kalman Filter Summary

- **Highly efficient:** Polynomial in measurement dimensionality k and state dimensionality n :

$$O(k^{2.376} + n^2)$$

- **Optimal for linear Gaussian systems!**
- Most robotics systems are **nonlinear!**

Linearity Assumption Revisited



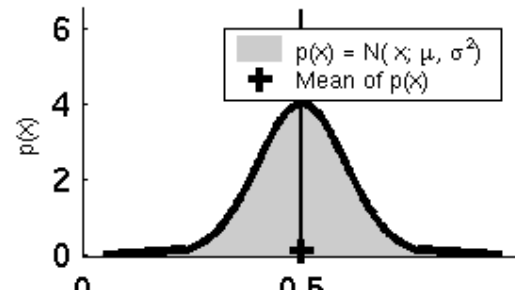
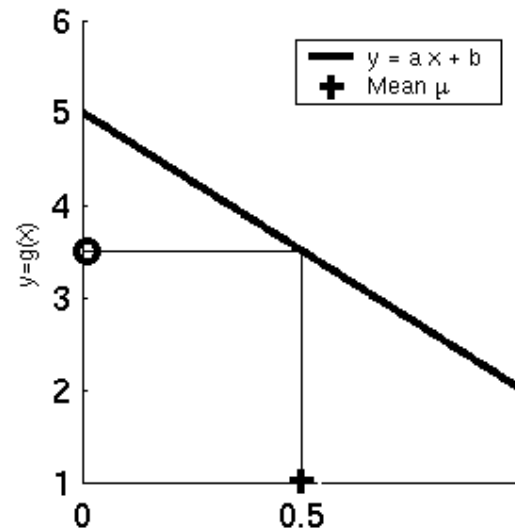
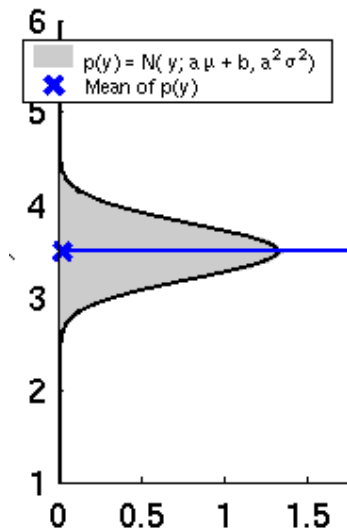
Nonlinear Dynamic Systems

- Most realistic robotic problems involve nonlinear functions

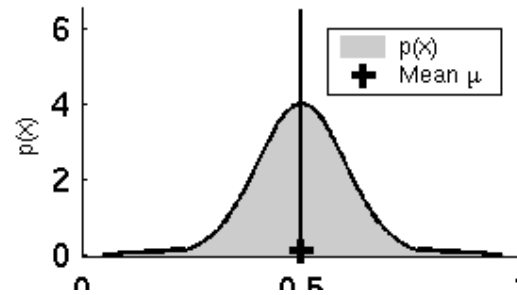
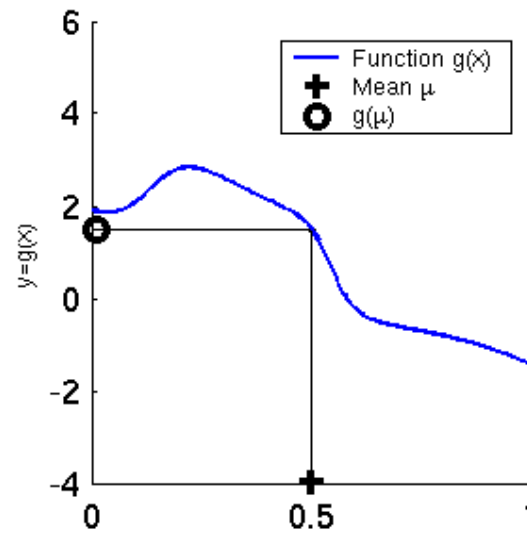
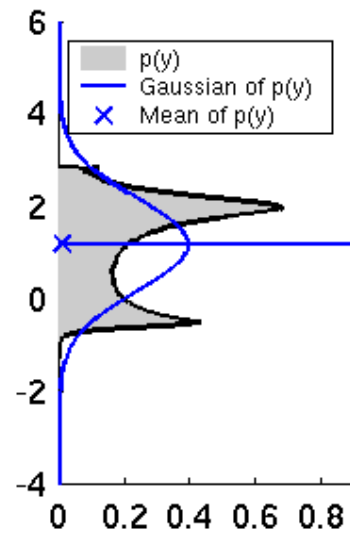
$$x_t = g(u_t, x_{t-1})$$

$$z_t = h(x_t)$$

Linearity Assumption Revisited

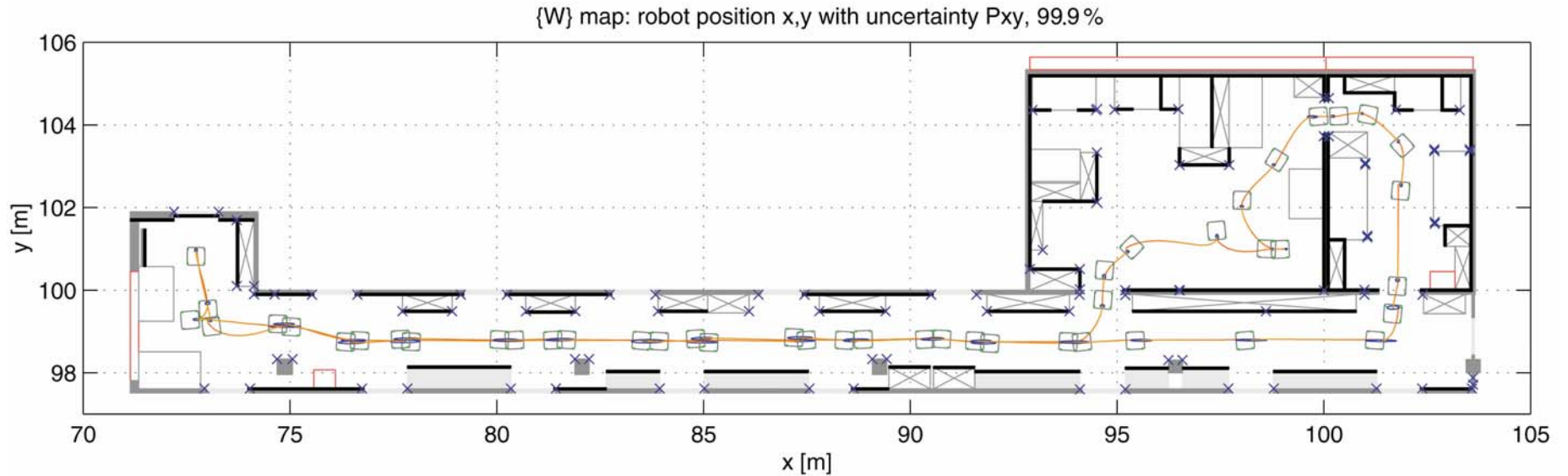


Non-linear Function



Kalman Filter-based System

- [Arras et al. 98]:
 - Laser range-finder and vision
 - High precision ($< 1\text{cm}$ accuracy)



[Courtesy of Kai Arras]

Multi-hypothesis Tracking

