

Principles of Robot Autonomy I

Parametric filtering
(Kalman Filter and Extended Kalman Filter)

Attendance Form



Agenda

- Agenda
 - Parametric filtering: Kalman Filter and Extended Kalman Filter
- Readings:
 - Chapter 14, sections 14.1 – 14.3 in D. Gammelli, J. Lorenzetti, K. Luo, G. Zardini, M. Pavone. *Principles of Robot Autonomy*. 2026.

Belief distribution: recap

- **Belief distribution**: reflects internal knowledge about the state
- A belief distribution assigns a probability to each possible hypothesis with regard to the true state
- Formally, belief distributions are posterior probabilities over state variables conditioned on the available data

$$bel(x_t) := p(x_t \mid z_{1:t}, u_{1:t})$$

- Similarly, the *prediction* distribution is defined as

$$\overline{bel}(x_t) := p(x_t \mid \textcolor{red}{z}_{1:t-1}, u_{1:t})$$

- Calculating $bel(x_t)$ from $\overline{bel}(x_t)$ is called correction or measurement update

Bayes filter algorithm: recap

- **Bayes' filter algorithm**: most general algorithm for calculating beliefs
- **Key assumption**: state is complete

- Recursive algorithm
 - Step 1 (prediction):
compute $\overline{bel}(x_t)$
 - Step 2 (measurement update):
compute $bel(x_t)$
- Algorithm initialized with $bel(x_0)$
(e.g., uniform or points mass)

Data: $bel(x_{t-1}), u_t, z_t$

Result: $bel(x_t)$

foreach x_t **do**

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

end

Return $bel(x_t)$

Update rule



Instantiating the Bayes' filter

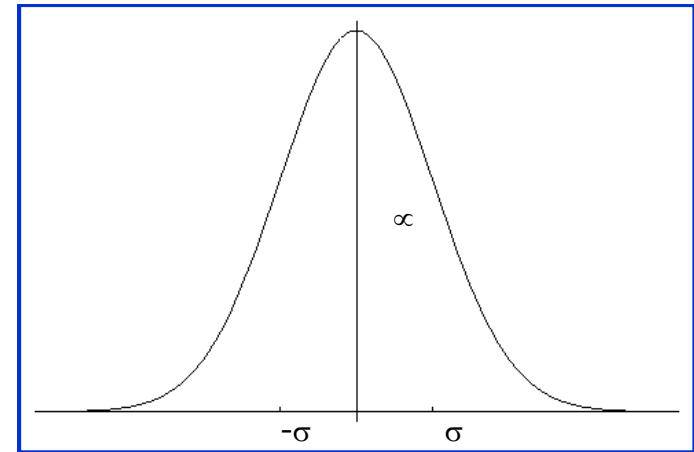
- Tractable implementations of Bayes' filter exploit structure and / or approximations; two main classes
 - Parametric filters: e.g., **KF**, **EKF**, UKF, etc.
 - Non parametric filters: e.g., histogram filter, particle filter, etc.

Gaussian distributions

- **Key idea:** belief represented as multivariate normal distribution

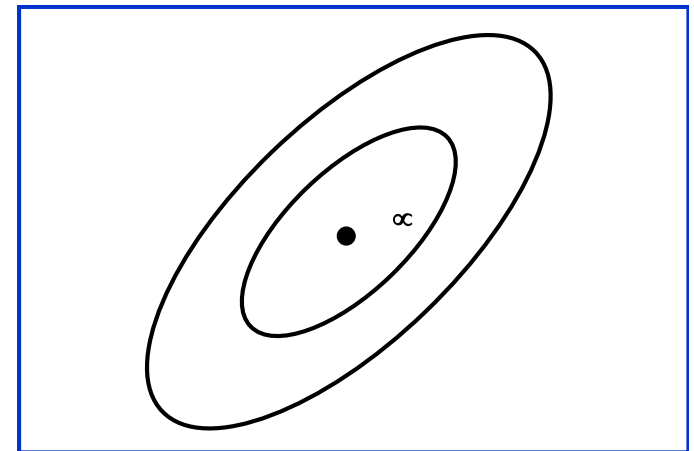
Univariate

$$p(x) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right) \\ \sim \mathcal{N}(x; \mu, \sigma^2)$$



Multivariate

$$p(x) = \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right) \\ \sim \mathcal{N}(\mu, \Sigma)$$



Key properties of Gaussian random variables

- If $X \sim \mathcal{N}(\mu, \Sigma)$ then

$$Y = AX + b \sim \mathcal{N}(A\mu + b, A\Sigma A^T)$$

- The sum of two independent Gaussian RVs

$$X_i \sim \mathcal{N}(\mu_i, \Sigma_i), \quad i = 1, 2$$

is Gaussian, specifically

$$X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \Sigma_1 + \Sigma_2)$$

- The product of Gaussian pdf is also Gaussian

Kalman filter (KF)

- Assumption #1: linear dynamics

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

- Independent process noise ϵ_t is $\mathcal{N}(0, R_t)$
- Assumption #1 implies that the probabilistic generative model is Gaussian

$$p(x_t | u_t, x_{t-1}) = \det(2\pi R_t)^{-\frac{1}{2}} \exp \left(-\frac{1}{2} (x_t - A_t x_{t-1} - B_t u_t)^T R_t^{-1} (x_t - A_t x_{t-1} - B_t u_t) \right)$$

Kalman filter (KF)

- Assumption #2: linear measurement model

$$z_t = C_t x_t + \delta_t$$

- Independent measurement noise δ_t is $\mathcal{N}(0, Q_t)$
- Assumption #2 implies that the measurement probability is Gaussian

$$p(z_t | x_t) = \det(2\pi Q_t)^{-\frac{1}{2}} \exp \left(-\frac{1}{2} (z_t - C_t x_t)^T Q_t^{-1} (z_t - C_t x_t) \right)$$

Kalman filter (KF)

- Assumption #3: the initial belief is Gaussian

$$bel(x_0) = p(x_0) = \det(2\pi\Sigma_0)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(x_0 - \mu_0)^T \Sigma_0^{-1}(x_0 - \mu_0)\right)$$

- **Key fact:** These three assumptions ensure that the posterior $bel(x_t)$ is Gaussian for all t , i.e., $bel(x_t) = \mathcal{N}(\mu_t, \Sigma_t)$
- Note:
 - KF implements belief computation for continuous states
 - Gaussians are unimodal -> commitment to single-hypothesis filtering

Kalman filter: algorithm

Prediction

Project state ahead

$$\bar{\mu}_t = A_t \mu_{t-1} + B_t u_t$$

Project covariance ahead

$$\bar{\Sigma}_t = A_t \Sigma_{t-1} A_t^T + R_t$$

Correction

Compute Kalman gain

$$K_t = \bar{\Sigma}_t C_t^T (C_t \bar{\Sigma}_t C_t^T + Q_t)^{-1}$$

Update estimate with new measurement

$$\mu_t = \bar{\mu}_t + K_t (z_t - C_t \bar{\mu}_t)$$

Update covariance

$$\Sigma_t = (I - K_t C_t) \bar{\Sigma}_t$$

Data: $\overbrace{(\mu_{t-1}, \Sigma_{t-1})}^{bel(x_{t-1})}, u_t, z_t$
Result: (μ_t, Σ_t)

Prediction:
 $\overbrace{(\bar{\mu}_t, \bar{\Sigma}_t)}^{bel(x_t)}$ $\left\{ \begin{array}{l} \bar{\mu}_t = A_t \mu_{t-1} + B_t u_t ; \\ \bar{\Sigma}_t = A_t \Sigma_{t-1} A_t^T + R_t ; \end{array} \right.$

Correction:
 $\overbrace{(\mu_t, \Sigma_t)}^{bel(x_t)}$ $\left\{ \begin{array}{l} K_t = \bar{\Sigma}_t C_t^T (C_t \bar{\Sigma}_t C_t^T + Q_t)^{-1} ; \\ \mu_t = \bar{\mu}_t + K_t (z_t - C_t \bar{\mu}_t) ; \\ \Sigma_t = (I - K_t C_t) \bar{\Sigma}_t ; \end{array} \right.$

Return $\underbrace{(\mu_t, \Sigma_t)}_{bel(x_t)}$

Kalman filter: derivation (sketch)

- Prediction

$$\overline{bel}(x_t) = \int \underbrace{p(x_t | x_{t-1}, u_t)}_{\mathcal{N}(A_t x_{t-1} + B_t u_t, R_t)} \cdot \underbrace{bel(x_{t-1})}_{\mathcal{N}(\mu_{t-1}, \Sigma_{t-1})} dx_{t-1}$$

- Recalling that $x_t = A_t x_{t-1} + B_t u_t + \epsilon_t$

$$\overline{bel}(x_t) = \mathcal{N}(\bar{\mu}_t, \bar{\Sigma}_t) \quad \text{with} \quad \begin{aligned} \bar{\mu}_t &= A_t \mu_{t-1} + B_t u_t \\ \bar{\Sigma}_t &= A_t \Sigma_{t-1} A_t^T + R_t \end{aligned}$$

Kalman filter: derivation (sketch)

- Correction

$$bel(x_t) = \eta \, p(z_t \mid x_t) \quad \cdot \quad \overline{bel(x_t)}$$

\downarrow \downarrow

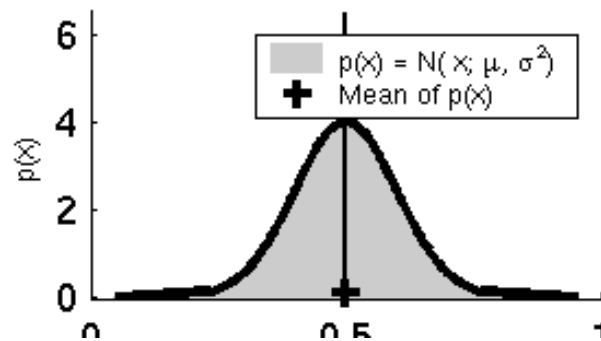
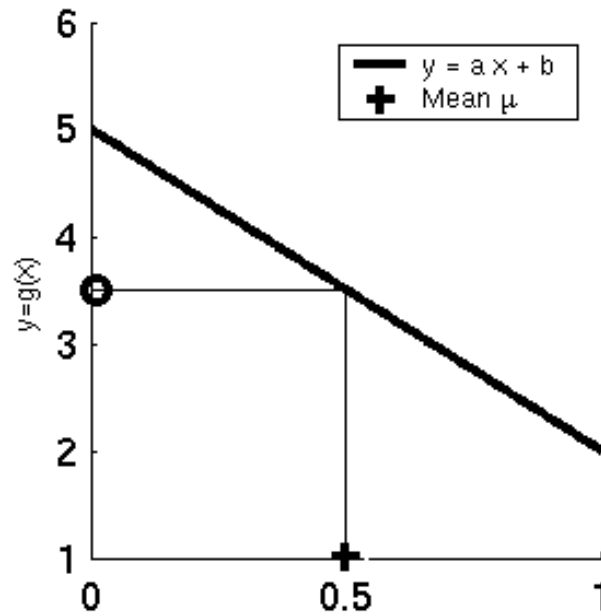
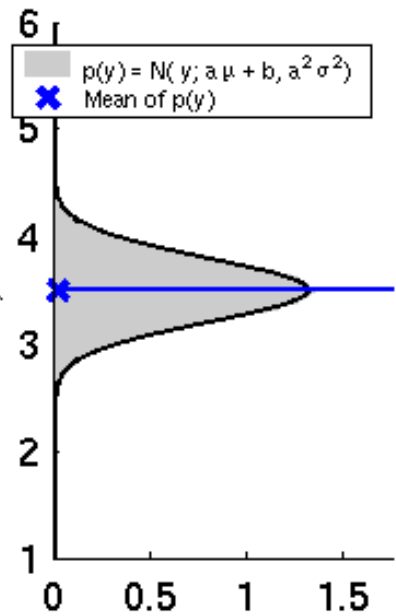
$$\mathcal{N}(C_t x_t, Q_t) \qquad \mathcal{N}(\bar{\mu}_t, \bar{\Sigma}_t)$$

- After some algebraic manipulations

$$bel(x_t) = \mathcal{N}(\mu_t, \Sigma_t) \qquad \text{with} \qquad \begin{aligned} K_t &= \bar{\Sigma}_t C_t^T (C_t \bar{\Sigma}_t C_t^T + Q_t)^{-1} \\ \mu_t &= \bar{\mu}_t + K_t (z_t - C_t \bar{\mu}_t) \\ \Sigma_t &= (I - K_t C_t) \bar{\Sigma}_t \end{aligned}$$

- Other derivations are possible; see, e.g., R. E. Kalman, A new approach to linear filtering and prediction problems. Journal of Basic Engineering, 82(1), 35-45, 1960.

Revisiting linearity assumption



- KF crucially exploits the property that a linear transformation of a Gaussian RV results in a Gaussian RV
- However, linearity assumptions are severely restrictive for robotics applications

Extended Kalman filter (EKF)

- **Goal:** relax the linearity assumption
- The dynamics are now given by

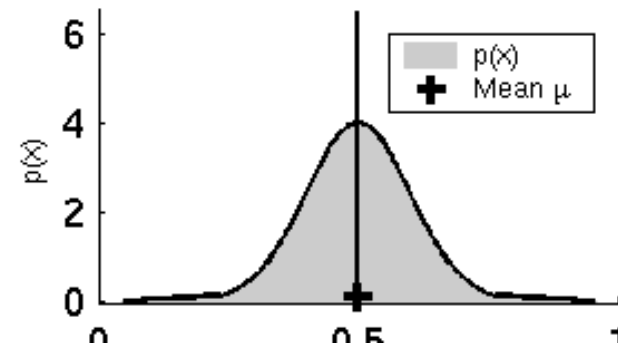
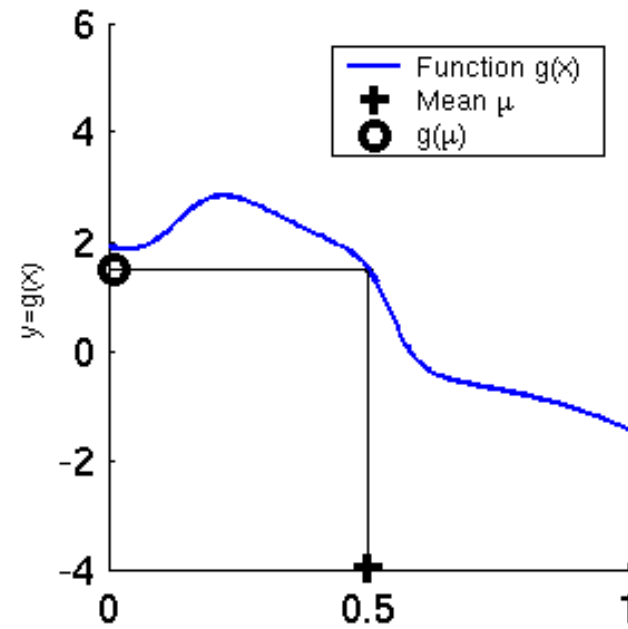
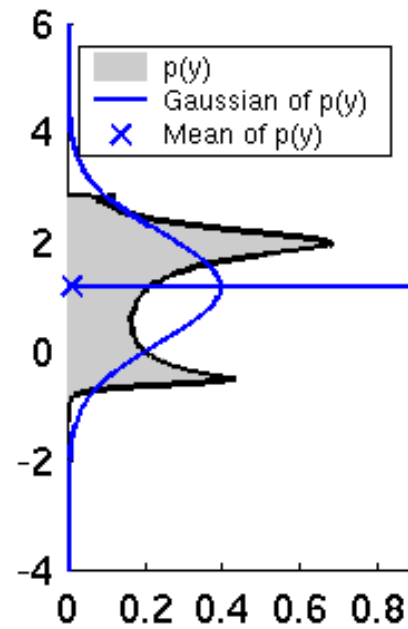
$$x_t = g(u_t, x_{t-1}) + \epsilon_t$$

- And the measurement model is now given by

$$z_t = h(x_t) + \delta_t$$

- Key idea: shift focus from computing exact posterior to efficiently compute a Gaussian approximation

Goal of EKF



EKF: key idea

- **Key idea:** linearize g and h around the most likely state and transform beliefs according to such linear approximations
- For the dynamics equation

$$g(u_t, x_{t-1}) \approx g(u_t, \mu_{t-1}) + \underbrace{J_g(u_t, \mu_{t-1})}_{\substack{\text{Jacobian of } g \\ := G_t}} (x_{t-1} - \mu_{t-1})$$

- Accordingly

$$p(x_t | u_t, x_{t-1}) = \det(2\pi R_t)^{-1/2}$$

$$\exp \left(-\frac{1}{2} [x_t - g(u_t, \mu_{t-1}) - G_t(x_{t-1} - \mu_{t-1})]^T R_t^{-1} [x_t - g(u_t, \mu_{t-1}) - G_t(x_{t-1} - \mu_{t-1})] \right)$$

EKF: key idea

- For the measurement model

$$h(x_t) \approx h(\bar{\mu}_t) + \underbrace{J_h(\bar{\mu}_t)}_{:=H_t}(x_t - \bar{\mu}_t)$$

- Accordingly,

$$p(z_t | x_t) = \det(2\pi Q_t)^{-1/2} \exp \left(-\frac{1}{2} [z_t - h(\bar{\mu}_t) - H_t(x_t - \bar{\mu}_t)] Q_t^{-1} [z_t - h(\bar{\mu}_t) - H_t(x_t - \bar{\mu}_t)] \right)$$

EKF: algorithm

- Main differences:
 1. Linear predictions are replaced by their nonlinear generalizations
 2. EKF uses Jacobians instead of linear system matrices
 3. Mathematical derivation of EKF parallels that of KF

Data: $(\mu_{t-1}, \Sigma_{t-1}), u_t, z_t$

Result: (μ_t, Σ_t)

$$\bar{\mu}_t = g(u_t, \mu_{t-1});$$

$$\bar{\Sigma}_t = G_t \Sigma_{t-1} G_t^T + R_t;$$

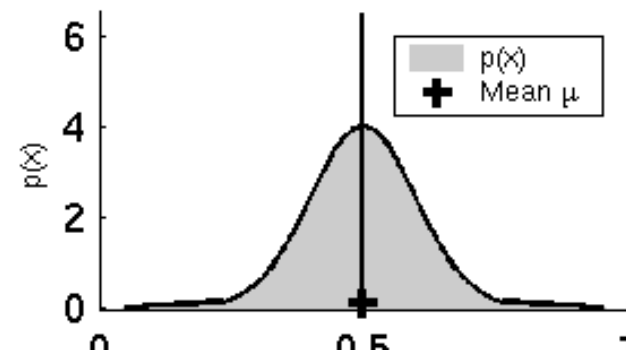
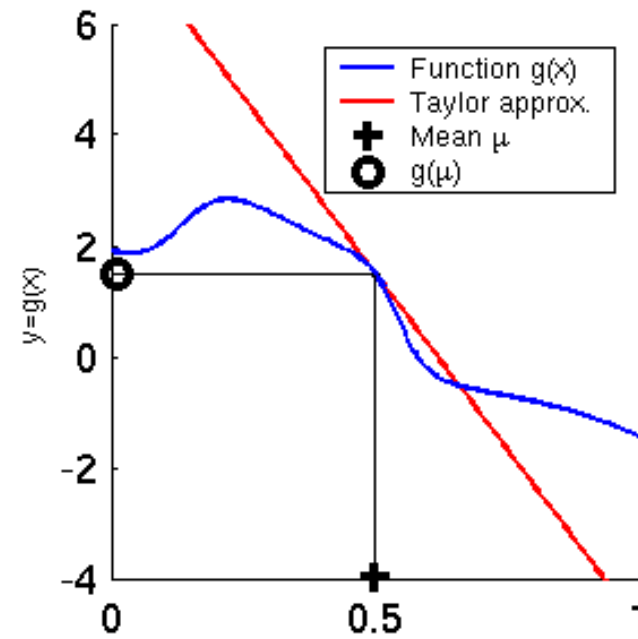
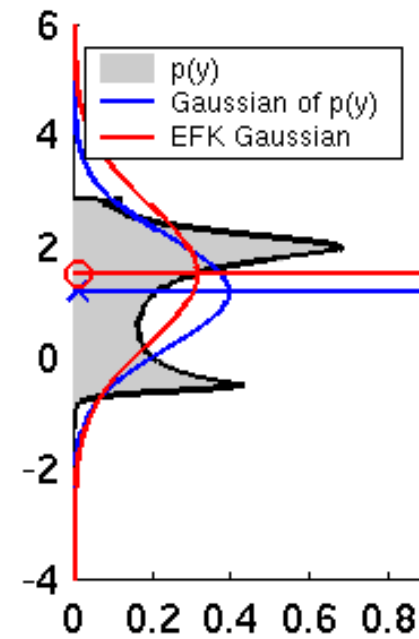
$$K_t = \bar{\Sigma}_t H_t^T (H_t \bar{\Sigma}_t H_t^T + Q_t)^{-1};$$

$$\mu_t = \bar{\mu}_t + K_t (z_t - h(\bar{\mu}_t));$$

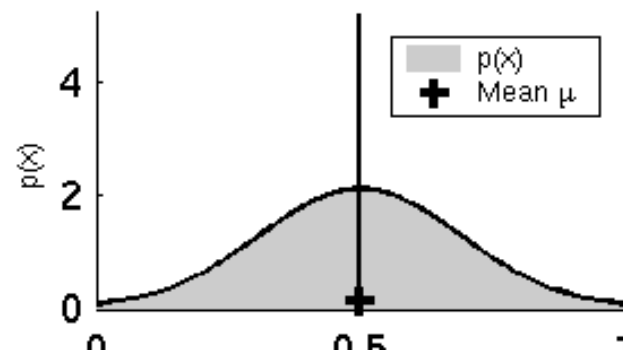
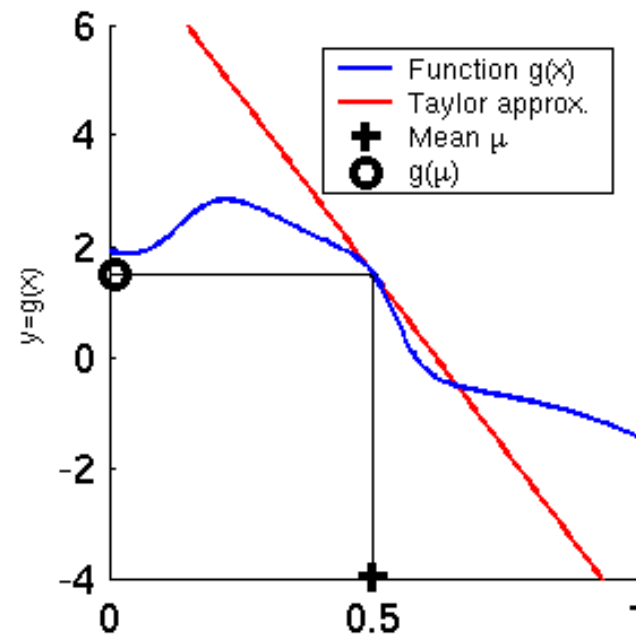
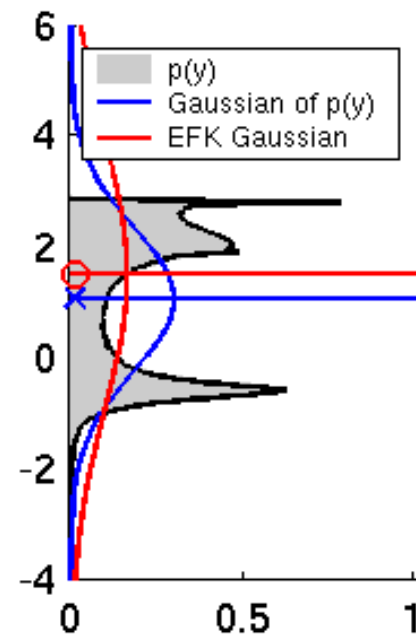
$$\Sigma_t = (I - K_t H_t) \bar{\Sigma}_t;$$

Return (μ_t, Σ_t)

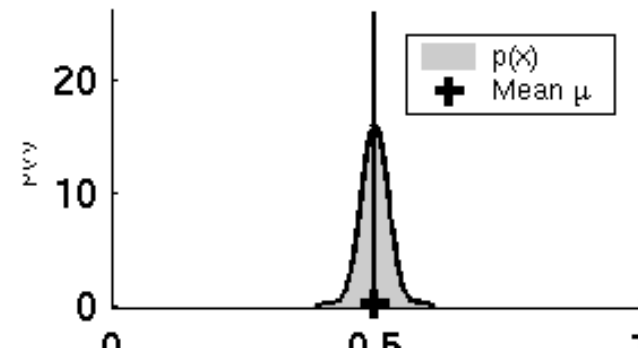
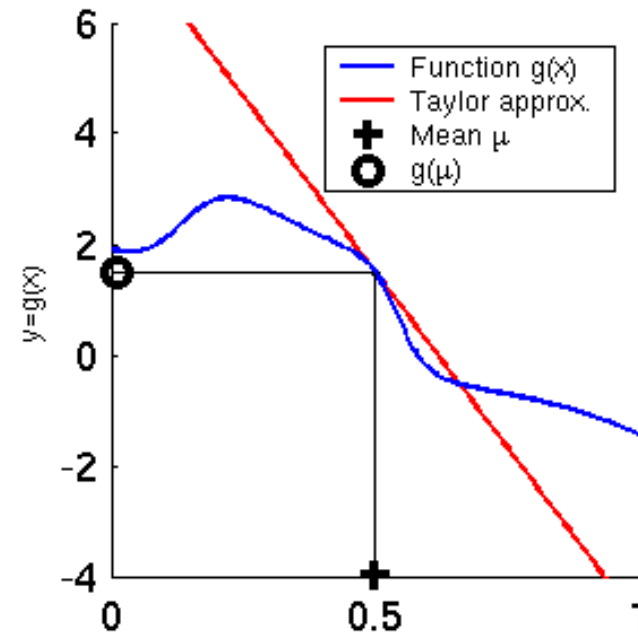
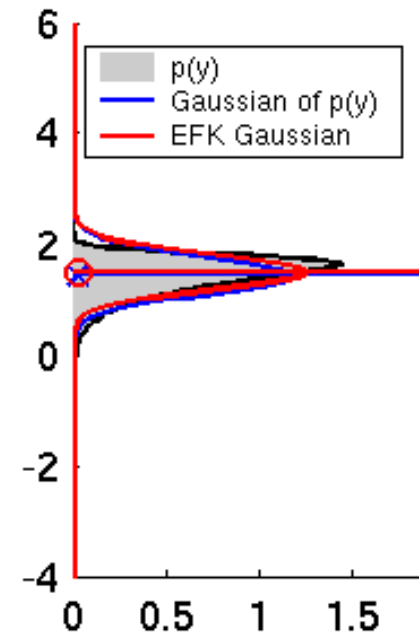
EKF: examples



EKF: examples



EKF: examples



Next time

