Kalman Filter

Robot Localization and Mapping
16-833

Michael Kaess

February 22, 2021

Slides courtesy of Ryan Eustice and probabilistic-robotics.org

Discrete Kalman Filter

Estimates the (n x 1) state $\mathbf{x}_t$ of a discrete-time controlled process that is governed by the linear stochastic difference equation

$$\mathbf{x}_t = A_t \mathbf{x}_{t-1} + B_t \mathbf{u}_t + \varepsilon_t$$

Observed through (k x 1) measurements $\mathbf{z}_t$

$$\mathbf{z}_t = C_t \mathbf{x}_t + \delta_t$$

Components of a Kalman Filter

A_t	Matrix (n x n) that describes how the state evolves from t-1 to t without controls or noise.
B_t	Matrix (n x m) that describes how the control u_t changes the state from t-1 to t .
C_t	Matrix (k x n) that describes a projection of state x_t to an observation z_t .
ε_t δ_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.

Reminder: Bayes Filters

z = observation
u = action
x = state

$$\boxed{Bel(x_t)} = p(x_t \mid u_1, z_1, \ldots, u_t, z_t)$$

Bayes $$= \eta\, p(z_t \mid x_t, u_1, z_1, \ldots, u_t)\, p(x_t \mid u_1, z_1, \ldots, u_t)$$

Markov $$= \eta\, p(z_t \mid x_t)\, p(x_t \mid u_1, z_1, \ldots, u_t)$$

Total prob. $$= \eta\, p(z_t \mid x_t) \int p(x_t \mid u_1, z_1, \ldots, u_t, x_{t-1})\, p(x_{t-1} \mid u_1, z_1, \ldots, u_t)\, dx_{t-1}$$

Markov $$= \eta\, p(z_t \mid x_t) \int p(x_t \mid u_t, x_{t-1})\, p(x_{t-1} \mid u_1, z_1, \ldots, u_t)\, dx_{t-1}$$

Markov $$= \eta\, p(z_t \mid x_t) \int p(x_t \mid u_t, x_{t-1})\, p(x_{t-1} \mid u_1, z_1, \ldots, z_{t-1})\, dx_{t-1}$$

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

Linear Gaussian Systems: Initialization

- Initial belief is normally distributed:

$$bel(\mathbf{x}_0) = N\left(\mathbf{x}_0; \mu_0, \Sigma_0\right)$$

Linear Gaussian Systems: Dynamics

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

$$\mathbf{x}_t = A_t \mathbf{x}_{t-1} + B_t \mathbf{u}_t + \varepsilon_t$$

$$p(\mathbf{x}_t \mid \mathbf{u}_t, \mathbf{x}_{t-1}) = N\left(\mathbf{x}_t; A_t \mathbf{x}_{t-1} + B_t \mathbf{u}_t, R_t\right)$$

$$\overline{bel}(\mathbf{x}_t) = \int p(\mathbf{x}_t \mid \mathbf{u}_t, \mathbf{x}_{t-1}) \qquad bel(\mathbf{x}_{t-1}) d\mathbf{x}_{t-1}$$

$$\Downarrow \qquad \qquad \Downarrow$$

$$\sim N\left(\mathbf{x}_t; A_t \mathbf{x}_{t-1} + B_t \mathbf{u}_t, R_t\right) \qquad \sim N\left(\mathbf{x}_{t-1}; \mu_{t-1}, \Sigma_{t-1}\right)$$

Linear Gaussian Systems: Dynamics

$$\overline{bel}(\mathbf{x}_t) = \int p(\mathbf{x}_t \mid \mathbf{u}_t, \mathbf{x}_{t-1}) \qquad bel(\mathbf{x}_{t-1}) d\mathbf{x}_{t-1}$$

$$\Downarrow \qquad\qquad \Downarrow$$

$$\sim N\left(\mathbf{x}_t; A_t \mathbf{x}_{t-1} + B_t \mathbf{u}_t, R_t\right) \sim N\left(\mathbf{x}_{t-1}; \mu_{t-1}, \Sigma_{t-1}\right)$$

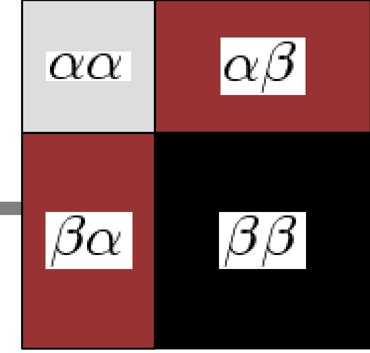
$$\Downarrow$$

$$\overline{bel}(\mathbf{x}_t) = \eta \int \exp\left\{-\frac{1}{2}(\mathbf{x}_t - A_t \mathbf{x}_{t-1} - B_t \mathbf{u}_t)^T R_t^{-1} (\mathbf{x}_t - A_t \mathbf{x}_{t-1} - B_t \mathbf{u}_t)\right\}$$

$$\exp\left\{-\frac{1}{2}(\mathbf{x}_{t-1} - \mu_{t-1})^T \Sigma_{t-1}^{-1} (\mathbf{x}_{t-1} - \mu_{t-1})\right\} d\mathbf{x}_{t-1}$$

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

Reminder: Gaussian Parameterizations



	Covariance Form	Information Form
Marginalization $p(\boldsymbol{\alpha}) = \int p(\boldsymbol{\alpha}, \boldsymbol{\beta}) d\boldsymbol{\beta}$	$\boldsymbol{\mu} = \boldsymbol{\mu}_\alpha$ $\Sigma = \Sigma_{\alpha\alpha}$ (sub-block)	$\boldsymbol{\eta} = \boldsymbol{\eta}_\alpha - \Lambda_{\alpha\beta}\Lambda_{\beta\beta}^{-1}\boldsymbol{\eta}_\beta$ $\Lambda = \Lambda_{\alpha\alpha} - \Lambda_{\alpha\beta}\Lambda_{\beta\beta}^{-1}\Lambda_{\beta\alpha}$ (Schur complement)
Conditioning $p(\boldsymbol{\alpha}\|\boldsymbol{\beta}) = \frac{p(\boldsymbol{\alpha}, \boldsymbol{\beta})}{p(\boldsymbol{\beta})}$	$\boldsymbol{\mu}' = \boldsymbol{\mu}_\alpha + \Sigma_{\alpha\beta}\Sigma_{\beta\beta}^{-1}(\boldsymbol{\beta} - \boldsymbol{\mu}_\beta)$ $\Sigma' = \Sigma_{\alpha\alpha} - \Sigma_{\alpha\beta}\Sigma_{\beta\beta}^{-1}\Sigma_{\beta\alpha}$ (Schur complement)	$\boldsymbol{\eta}' = \boldsymbol{\eta}_\alpha - \Lambda_{\alpha\beta}\boldsymbol{\beta}$ $\Lambda' = \Lambda_{\alpha\alpha}$ (sub-block)

Linear Gaussian Systems: Observations

- Observations are linear function of state plus additive noise:

$$\mathbf{z}_t = C_t \mathbf{x}_t + \delta_t$$

$$p(\mathbf{z}_t \mid \mathbf{x}_t) = N\left(\mathbf{z}_t ; C_t \mathbf{x}_t , Q_t\right)$$

$$\begin{array}{ll} bel(\mathbf{x}_t) = \eta\, p(\mathbf{z}_t \mid \mathbf{x}_t) & \overline{bel}(\mathbf{x}_t) \\ \qquad\qquad\qquad \Downarrow & \qquad \Downarrow \\ \qquad \sim N\left(\mathbf{z}_t ; C_t \mathbf{x}_t , Q_t\right) & \sim N\left(\mathbf{x}_t ; \bar{\mu}_t , \bar{\Sigma}_t\right) \end{array}$$

Linear Gaussian Systems: Observations

$$bel(\mathbf{x}_t) = \eta \quad p(\mathbf{z}_t \mid \mathbf{x}_t) \qquad \overline{bel}(\mathbf{x}_t)$$

$$\Downarrow \qquad\qquad \Downarrow$$

$$\sim N(\mathbf{z}_t; C_t\mathbf{x}_t, Q_t) \qquad \sim N(\mathbf{x}_t; \bar{\mu}_t, \bar{\Sigma}_t)$$

$$\Downarrow$$

$$bel(\mathbf{x}_t) = \eta \exp\left\{-\frac{1}{2}(\mathbf{z}_t - C_t\mathbf{x}_t)^T Q_t^{-1}(\mathbf{z}_t - C_t\mathbf{x}_t)\right\} \exp\left\{-\frac{1}{2}(\mathbf{x}_t - \bar{\mu}_t)^T \bar{\Sigma}_t^{-1}(\mathbf{x}_t - \bar{\mu}_t)\right\}$$

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

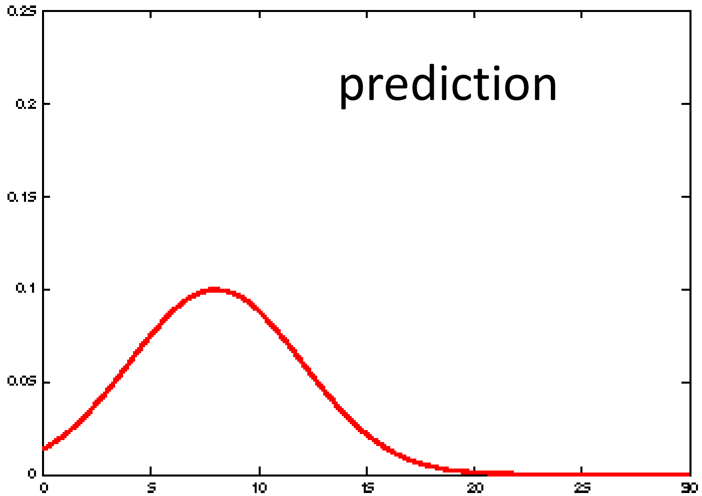
Kalman Filter Algorithm

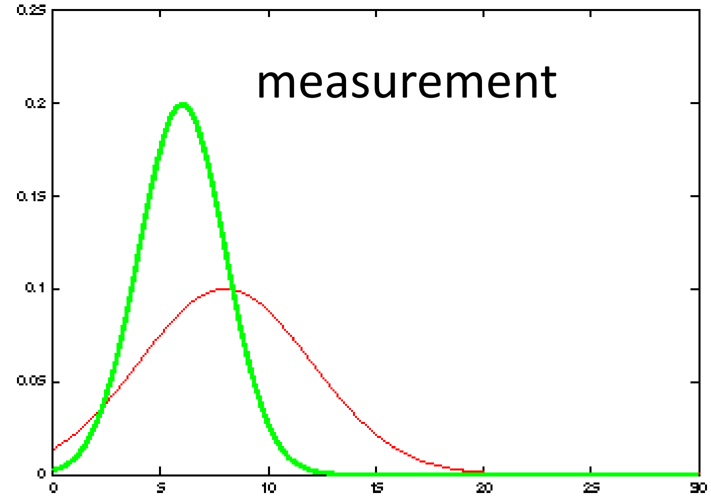
$$\textbf{Kalman\_filter}(\boldsymbol{\mu}_{t-1}, \Sigma_{t-1}, \mathbf{u}_t, \mathbf{z}_t)\textbf{:}$$

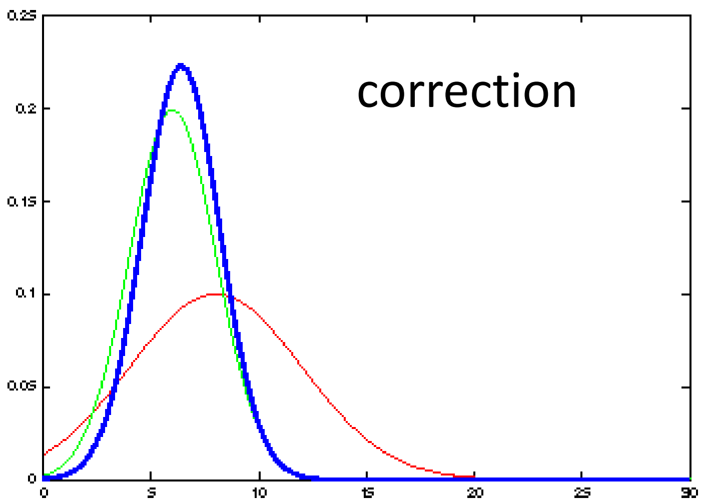
$$\bar{\boldsymbol{\mu}}_t = A_t\ \boldsymbol{\mu}_{t-1} + B_t\ \mathbf{u}_t$$
$$\bar{\Sigma}_t = A_t\ \Sigma_{t-1}\ A_t^\top + R_t$$

$$K_t = \bar{\Sigma}_t\ C_t^\top (C_t\ \bar{\Sigma}_t\ C_t^\top + Q_t)^{-1}$$
$$\boldsymbol{\mu}_t = \bar{\boldsymbol{\mu}}_t + K_t(\mathbf{z}_t - C_t\ \bar{\boldsymbol{\mu}}_t)$$
$$\Sigma_t = (I - K_t\ C_t)\ \bar{\Sigma}_t$$
$$\textit{return}\ \boldsymbol{\mu}_t, \Sigma_t$$

1D Kalman Filter Example (1)

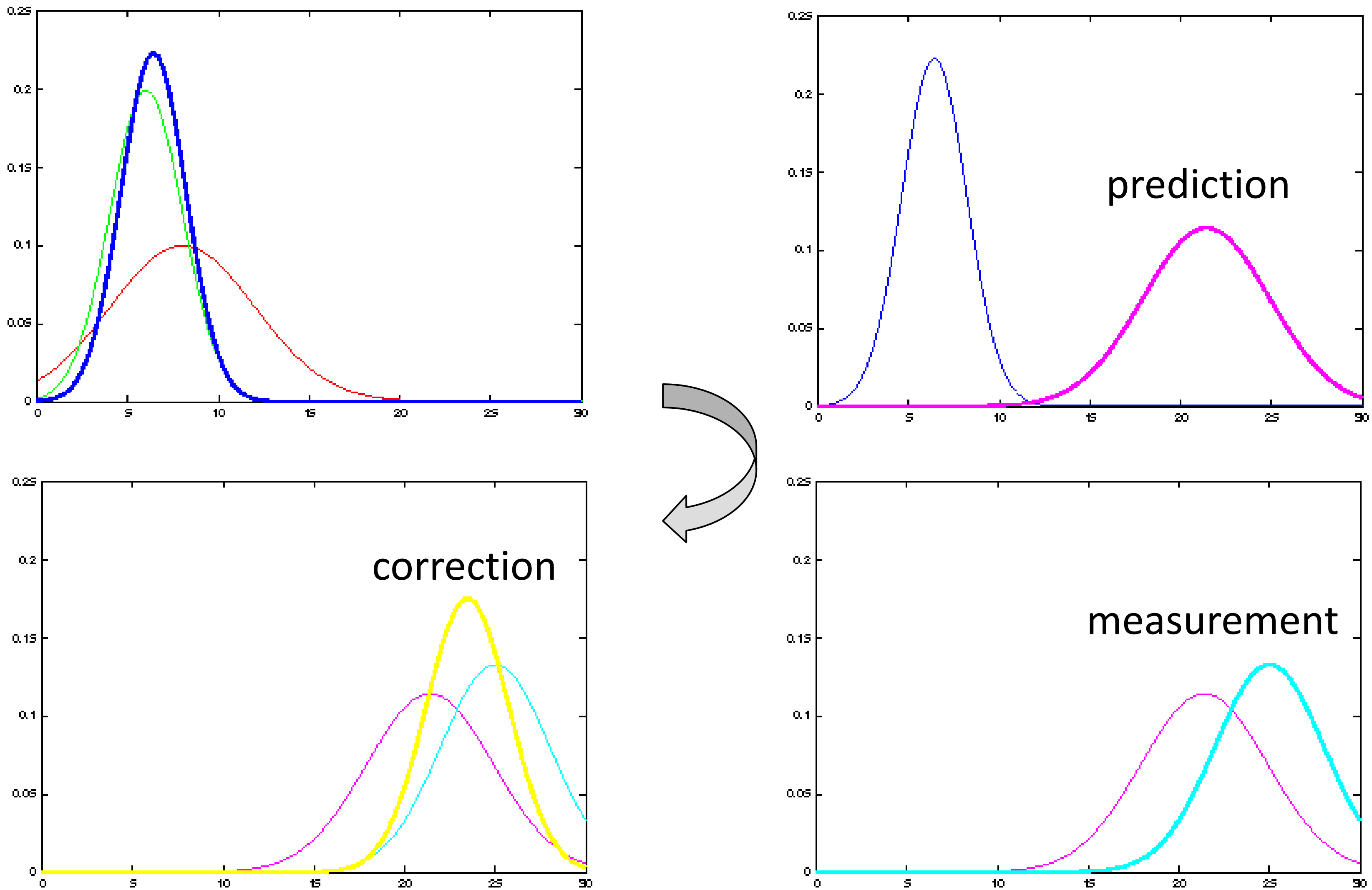






It's a weighted mean!

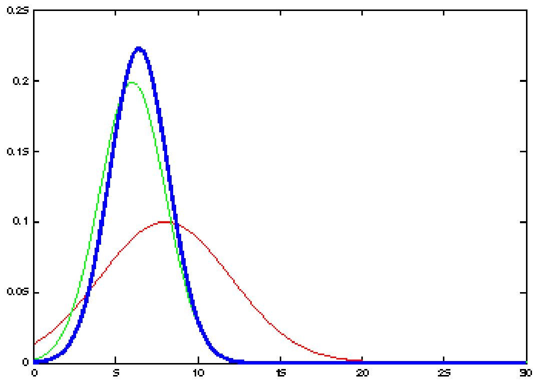
1D Kalman Filter Example (2)



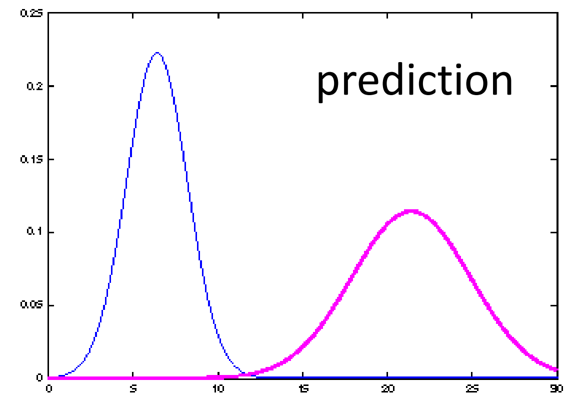
The Prediction-Correction-Cycle

Prediction

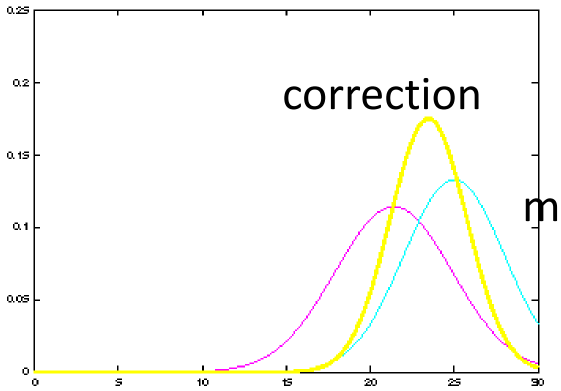
$$\overline{bel}(x_t) = \begin{cases} \bar{\mu}_t = a_t \mu_{t-1} + b_t u_t \\ \bar{\sigma}_t^2 = a_t^2 \sigma_t^2 + \sigma_{\varepsilon_t}^2 \end{cases}$$

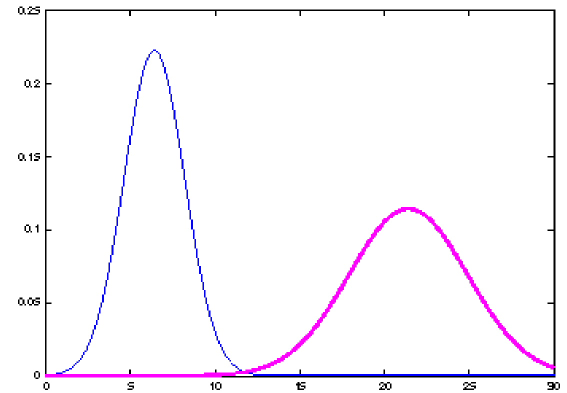


$$\overline{bel}(\mathbf{x}_t) = \begin{cases} \bar{\mu}_t = A_t \mu_{t-1} + B_t \mathbf{u}_t \\ \bar{\Sigma}_t = A_t \Sigma_{t-1} A_t^T + R_t \end{cases}$$



The Prediction-Correction-Cycle





$$bel(x_t) = \begin{cases} \mu_t = \bar{\mu}_t + k_t(z_t - c_t\bar{\mu}_t) \\ \sigma_t^2 = (1 - k_t c_t)\bar{\sigma}_t^2 \end{cases}, k_t = \frac{c_t \bar{\sigma}_t^2}{c_t^2 \bar{\sigma}_t^2 + \sigma_{\delta_t}^2}$$

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

Correction

The Prediction-Correction-Cycle

Prediction

$$bel(x_t) = \begin{cases} \mu_t = \bar{\mu}_t + k_t(z_t - c_t\bar{\mu}_t) \\ \sigma_t^2 = (1 - k_t c_t)\bar{\sigma}_t^2 \end{cases}, k_t = \frac{c_t \bar{\sigma}_t^2}{c_t^2 \bar{\sigma}_t^2 + \sigma_{\delta_t}^2}$$

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

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

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

Correction

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**!
  - No other estimator can do better
- Most robotics systems are nonlinear!