Kalman Filters

Gaurav Pandey

August 26, 2016

Kalman vs Bayes Filter

- Bayes Filter can track the entire posterior of the state for any probability distributions. For a small discrete state-space (like the robot and door problem) it is easy to estimate the full posterior. However for most of the real world problems the state-space is generally **continuous and too large**.
- Kalman filters are the first tractable implementation of the Bayes filter for a continuous space that assumes a **Linear Gaussian System.**
- In case of gaussian distributions the full posterior can be charecterized by just two parameters **mean & covarioance**

$$p(x) \quad = \quad \det\left(2\pi\Sigma\right)^{-\frac{1}{2}} \ \exp\left\{-\tfrac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right\}$$

Linear Gaussian System

- Linear Action Model (Measurement Model) with additive Gaussian noise.

Action / Process Model $$x_t = A_t x_{t-1} + B_t u_t + \varepsilon_t \longrightarrow N(0, R_t)$$

Gaussian Random Vectors

Measurement Model $$z_t = C_t x_t + \delta_t \longrightarrow N(0, Q_t)$$

$$x_t = \begin{pmatrix} x_{1,t} \\ x_{2,t} \\ \vdots \\ x_{n,t} \end{pmatrix} \qquad u_t = \begin{pmatrix} u_{1,t} \\ u_{2,t} \\ \vdots \\ u_{m,t} \end{pmatrix}$$

Linear Gaussian System

- State Transition Probability is Gaussian

$$
\begin{aligned}
& p(x_t \mid u_t, x_{t-1}) \\
&\quad = \quad \det\left(2\pi R_t\right)^{-\frac{1}{2}} \; \exp\left\{-\tfrac{1}{2}(x_t - A_t x_{t-1} - B_t u_t)^T R_t^{-1} (x_t - A_t x_{t-1} - B_t u_t)\right\}
\end{aligned}
$$

- Measurement Probability is also Gaussian

$$
p(z_t \mid x_t) \quad = \quad \det\left(2\pi Q_t\right)^{-\frac{1}{2}} \; \exp\left\{-\tfrac{1}{2}(z_t - C_t\, x_t)^T \; Q_t^{-1} \; (z_t - C_t\, x_t)\right\}
$$

Initial Belief

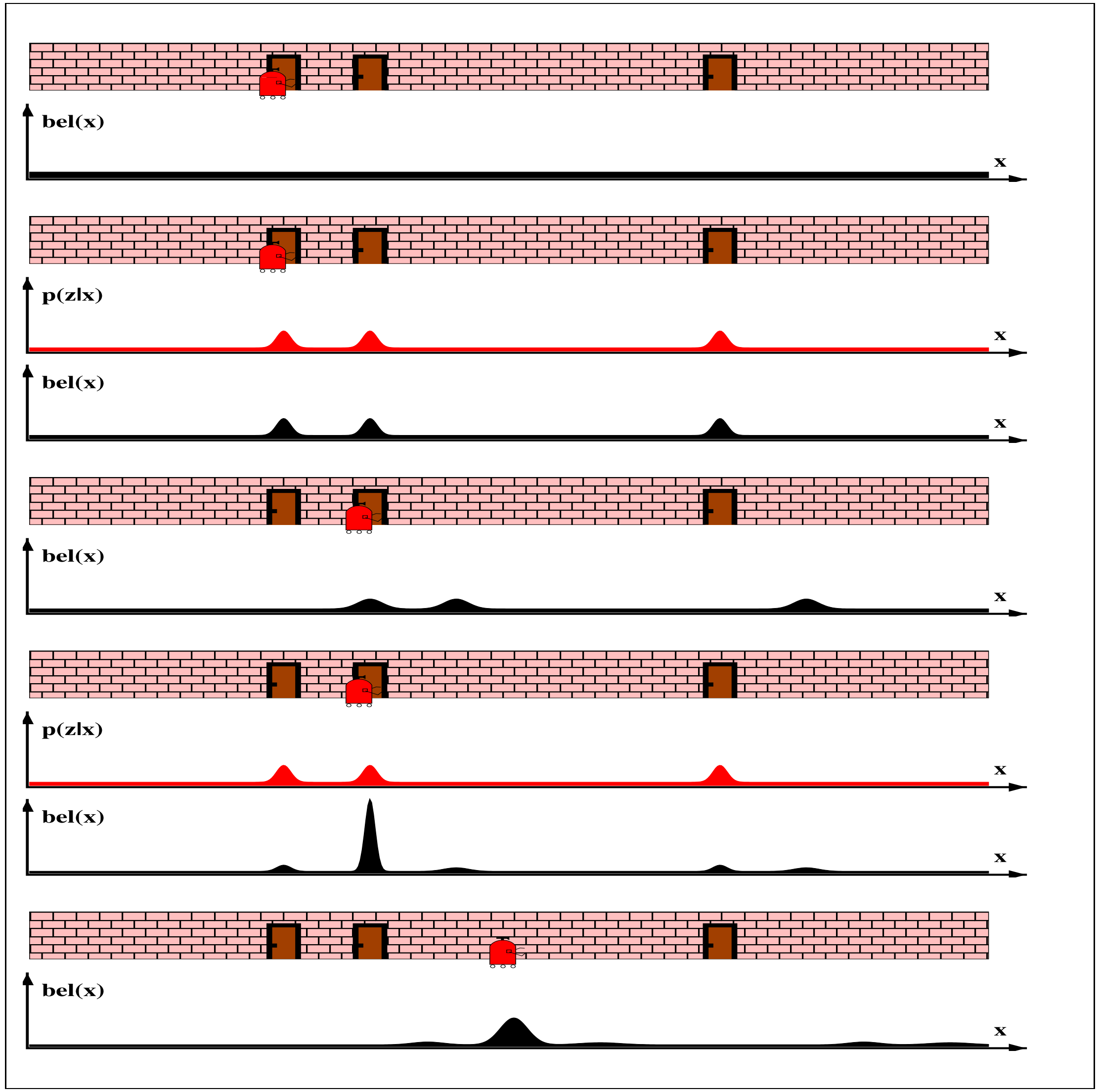
- Initial Belief is Considered to be Gaussian

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

- Recall Bayes Filter Algorithm

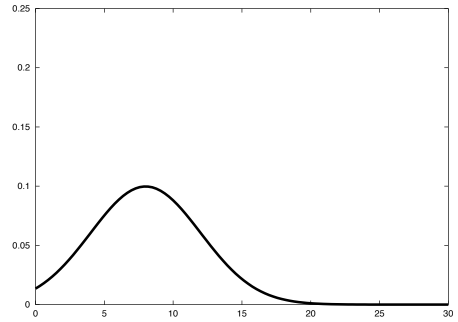
Algorithm Bayes_filter($bel(x_{t-1}), u_t, z_t$**):**
    for all x_t do
        $\overline{bel}(x_t) = \int p(x_t \mid u_t, x_{t-1})\ bel(x_{t-1})\ dx$
        $bel(x_t) = \eta\ p(z_t \mid x_t)\ \overline{bel}(x_t)$
    endfor
    return $bel(x_t)$

Recall

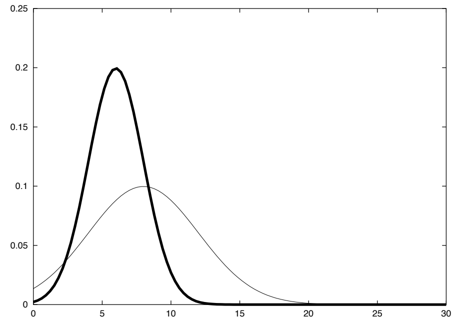


Kalman Filter in Action: Belief is Gaussian

Initial Belief



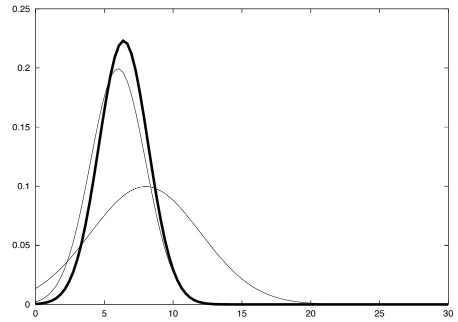
(a)



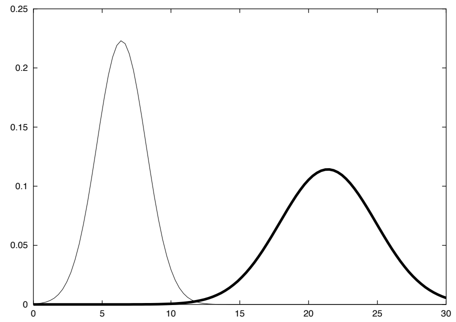
(b)

Measurement

Correction



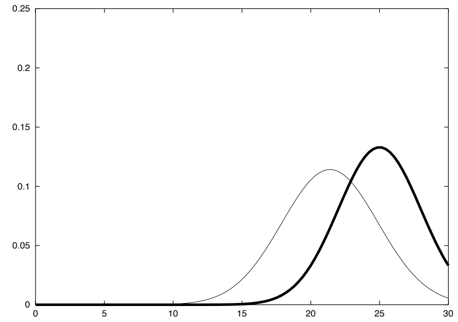
(c)



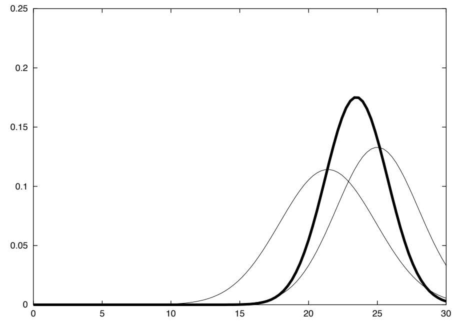
(d)

Prediction

Measurement



(e)



(f)

Correction

Kalman Filter vs Bayes Filter

Algorithm Kalman_filter($\mu_{t-1}, \Sigma_{t-1}, u_t, z_t$**):**

$$\bar{\mu}_t = A_t\, \mu_{t-1} + B_t\, u_t$$
$$\bar{\Sigma}_t = A_t\, \Sigma_{t-1}\, A_t^T + R_t$$
$$K_t = \bar{\Sigma}_t\, C_t^T (C_t\, \bar{\Sigma}_t\, C_t^T + Q_t)^{-1}$$
$$\mu_t = \bar{\mu}_t + K_t(z_t - C_t\, \bar{\mu}_t)$$
$$\Sigma_t = (I - K_t\, C_t)\, \bar{\Sigma}_t$$

return μ_t, Σ_t

Algorithm Bayes_filter($bel(x_{t-1}), u_t, z_t$**):**

for all x_t do

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

endfor

return $bel(x_t)$

Note: Predicted belief is Gaussian and posterior belief is also Gaussian

Prediction Step of Kalman Filter

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

$$\begin{aligned}\overline{bel}(x_t) \quad = \quad & \eta \int \exp\left\{-\tfrac{1}{2}\,(x_t - A_t\, x_{t-1} - B_t\, u_t)^T\, R_t^{-1}\,(x_t - A_t\, x_{t-1} - B_t\, u_t)\right\} \\ & \exp\left\{-\tfrac{1}{2}\,(x_{t-1} - \mu_{t-1})^T \Sigma_{t-1}^{-1} (x_{t-1} - \mu_{t-1})\right\}\; dx_{t-1}\end{aligned}$$

$$\overline{bel}(x_t) = \eta \int \exp\{-L_t\}\; dx_{t-1}$$

$$\boxed{\begin{aligned} L_t \quad = \quad & \tfrac{1}{2}\,(x_t - A_t\, x_{t-1} - B_t\, u_t)^T\, R_t^{-1}\,(x_t - A_t\, x_{t-1} - B_t\, u_t) \\ & + \tfrac{1}{2}\,(x_{t-1} - \mu_{t-1})^T\, \Sigma_{t-1}^{-1}\,(x_{t-1} - \mu_{t-1}). \end{aligned}}$$

Quadratic Function

Prediction

- We can decompose L_t :

$$L_t = L_t(x_{t-1}, x_t) + L_t(x_t)$$

$$\begin{aligned}
\overline{bel}(x_t) &= \eta \int \exp\{-L_t\}\ dx_{t-1} \\
&= \eta \int \exp\{-L_t(x_{t-1}, x_t) - L_t(x_t)\}\ dx_{t-1} \\
&= \eta\ \exp\{-L_t(x_t)\} \int \underbrace{\exp\{-L_t(x_{t-1}, x_t)\}\ dx_{t-1}}_{\substack{\text{Can be written as a Gaussian distribution} \\ \text{which can be integrated to a constant value !}}}
\end{aligned}$$

Derivative of Quadratic Form in Gaussian Distribution

- First derivative of the quadratic expression inside exponential gives us the mean of the distribution.
- Second derivative gives us the covariance.

$$C = \frac{1}{2}(x-\mu)^\top \Sigma^{-1}(x-\mu) = \frac{1}{2} x \Sigma^{-1} x - 2x^\top \Sigma^{-1}\mu + \mu^\top \mu$$

$$\frac{\partial C}{\partial x} = \Sigma^{-1} x - \Sigma^{-1}\mu$$

$$\frac{\partial^2 C}{\partial x^2} = \Sigma^{-1}$$

Prediction

We are seeking a function $L_t(x_{t-1}, x_t)$ quadratic in x_{t-1}

$$\frac{\partial L_t}{\partial x_{t-1}} \quad = \quad -A_t^T\ R_t^{-1}\ (x_t - A_t\ x_{t-1} - B_t\ u_t) + \Sigma_{t-1}^{-1}\ (x_{t-1} - \mu_{t-1})$$

$$\frac{\partial^2 L_t}{\partial x_{t-1}^2} \quad = \quad A_t^T\ R_t^{-1}\ A_t + \Sigma_{t-1}^{-1} \quad =: \quad \Psi_t^{-1}$$

Ψ_t defines the curvature of $L_t(x_{t-1}, x_t)$. Setting the first derivative of L_t to 0 gives us the mean:

$$A_t^T\ R_t^{-1}\ (x_t - A_t\ x_{t-1} - B_t\ u_t) \quad = \quad \Sigma_{t-1}^{-1}\ (x_{t-1} - \mu_{t-1})$$

Prediction

Thus, we now have a quadratic function $L_t(x_{t-1}, x_t)$, defined as follows:

$$\begin{aligned} L_t(x_{t-1}, x_t) \quad = \quad & \tfrac{1}{2}(x_{t-1} - \Psi_t\ [A_t^T\ R_t^{-1}\ (x_t - B_t\ u_t) + \Sigma_{t-1}^{-1}\ \mu_{t-1}])^T\ \Psi^{-1} \\ & (x_{t-1} - \Psi_t\ [A_t^T\ R_t^{-1}\ (x_t - B_t\ u_t) + \Sigma_{t-1}^{-1}\ \mu_{t-1}]) \end{aligned}$$

$$\int \det(2\pi\Psi)^{-\frac{1}{2}}\ \exp\{-L_t(x_{t-1}, x_t)\}\ dx_{t-1} \quad = \quad 1$$

$$\int \exp\{-L_t(x_{t-1}, x_t)\}\ dx_{t-1} \quad = \quad \det(2\pi\Psi)^{\frac{1}{2}} \longrightarrow$$

Constant
(independent of x_t)

$$\begin{aligned} \overline{bel}(x_t) \quad &= \quad \eta\ \exp\{-L_t(x_t)\} \int \exp\{-L_t(x_{t-1}, x_t)\}\ dx_{t-1} \\ &= \quad \eta\ \exp\{-L_t(x_t)\} \end{aligned}$$

Prediction

$$L_t(x_t) \quad = \quad L_t - L_t(x_{t-1}, x_t)$$

$$\begin{aligned} L_t(x_t) \quad = \quad & +\tfrac{1}{2}\ (x_t - B_t\ u_t)^T\ R_t^{-1}\ (x_t - B_t\ u_t)\ +\tfrac{1}{2}\ \mu_{t-1}^T\ \Sigma_{t-1}^{-1}\ \mu_{t-1} \\ & -\tfrac{1}{2}\ [A_t^T\ R_t^{-1}\ (x_t - B_t\ u_t) + \Sigma_{t-1}^{-1}\ \mu_{t-1}]^T\ (A_t^T\ R_t^{-1}\ A_t + \Sigma_{t-1}^{-1})^{-1} \\ & \qquad [A_t^T\ R_t^{-1}\ (x_t - B_t\ u_t) + \Sigma_{t-1}^{-1}\ \mu_{t-1}] \end{aligned}$$

- Note that this function is independent of x_{t-1}
- Now take the first and second derivative of this function to obtain the mean and covariance of the predicted belief
- After some matrix manipulations we get a very compact expression for the mean and covariance of the predicted belief !!

Prediction

- Predicted mean

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

- Predicted Covariance

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

Measurement Update

- Posterior Belief

$$bel(x_t) \quad = \quad \eta \quad \underbrace{p(z_t \mid x_t)}_{\sim\mathcal{N}(z_t; C_t x_t, Q_t)} \quad \underbrace{\overline{bel}(x_t)}_{\sim\mathcal{N}(x_t; \bar{\mu}_t, \bar{\Sigma}_t)}$$

$$bel(x_t) \quad = \quad \eta \; \exp\{-J_t\}$$

$$J_t \quad = \quad \frac{1}{2} \, (z_t - C_t x_t)^T \; Q_t^{-1} \; (z_t - C_t x_t) + \frac{1}{2} \; (x_t - \bar{\mu}_t)^T \; \bar{\Sigma}_t^{-1} \; (x_t - \bar{\mu}_t)$$

Measurement Update

$$
\begin{aligned}
J_t &= \tfrac{1}{2}\,(z_t - C_t x_t)^T\, Q_t^{-1}\,(z_t - C_t x_t) + \tfrac{1}{2}\,(x_t - \bar{\mu}_t)^T\,\bar{\Sigma}_t^{-1}\,(x_t - \bar{\mu}_t) \\
\frac{\partial J}{\partial x_t} &= -C_t^T\, Q_t^{-1}\,(z_t - C_t x_t) + \bar{\Sigma}_t^{-1}\,(x_t - \bar{\mu}_t) \\
\frac{\partial^2 J}{\partial x_t^2} &= C_t^T\, Q_t^{-1}\, C_t + \bar{\Sigma}_t^{-1} \\
\Sigma_t &= (C_t^T\, Q_t^{-1}\, C_t + \bar{\Sigma}_t^{-1})^{-1}
\end{aligned}
$$

The mean of $bel(x_t)$ is the minimum of this quadratic function, which we now calculate by setting the first derivative of J_t to zero (and substituting μ_t for x_t):

$$
C_t^T\, Q_t^{-1}\,(z_t - C_t \mu_t) = \bar{\Sigma}_t^{-1}\,(\mu_t - \bar{\mu}_t)
$$

Measurement Update

- Mean of Posterior belief:

$$C_t^T \, Q_t^{-1} \, (z_t - C_t \mu_t) \quad = \quad \bar{\Sigma}_t^{-1} \, (\mu_t - \bar{\mu}_t)$$

$$\begin{aligned} & C_t^T \, Q_t^{-1} \, (z_t - C_t \mu_t) \\ & \quad = \quad C_t^T \, Q_t^{-1} \, (z_t - C_t \, \mu_t + C_t \, \bar{\mu}_t - C_t \, \bar{\mu}_t) \\ & \quad = \quad C_t^T \, Q_t^{-1} \, (z_t - C_t \bar{\mu}_t) - C_t^T \, Q_t^{-1} \, C_t \, (\mu_t - \bar{\mu}_t) \end{aligned}$$

$$C_t^T \, Q_t^{-1} \, (z_t - C_t \bar{\mu}_t) \quad = \quad \underbrace{(C_t^T \, Q_t^{-1} \, C_t + \bar{\Sigma}_t^{-1})}_{= \Sigma_t^{-1}} \, (\mu_t - \bar{\mu}_t)$$

$$\Sigma_t \, C_t^T \, Q_t^{-1} \, (z_t - C_t \, \bar{\mu}_t) \quad = \quad \mu_t - \bar{\mu}_t$$

$$\boxed{\mu_t \quad = \quad \bar{\mu}_t + K_t \, (z_t - C_t \, \bar{\mu}_t)} \qquad \boxed{K_t \quad = \quad \Sigma_t \, C_t^T \, Q_t^{-1}}$$

Kalman Gain

- Kalman gain is a function of state covariance and the state covariance is obtained after twice inversion of the predicted state covariance matrix. In case of large state space (robot pose and landmarks !) this is computationally very expensive

$$K_t = \Sigma_t \, C_t^T \, Q_t^{-1}$$

$$\Sigma_t = \underbrace{(C_t^T \, Q_t^{-1} \, C_t + \bar{\Sigma}_t^{-1})^{-1}}_{\text{[n x n] Matrix}}$$

Kalman Gain Simplified

$$
\begin{aligned}
K_t &= \Sigma_t\, C_t^T\, Q_t^{-1} \\
&= \Sigma_t\, C_t^T\, Q_t^{-1}\, \underbrace{(C_t\, \bar{\Sigma}_t\, C_t^T + Q_t)\, (C_t\, \bar{\Sigma}_t\, C_t^T + Q_t)^{-1}}_{=I} \\
&= \Sigma_t\, (C_t^T\, Q_t^{-1}\, C_t\, \bar{\Sigma}_t\, C_t^T + C_t^T\, \underbrace{Q_t^{-1}\, Q_t}_{=I})\, (C_t\, \bar{\Sigma}_t\, C_t^T + Q_t)^{-1} \\
&= \Sigma_t\, (C_t^T\, Q_t^{-1}\, C_t\, \bar{\Sigma}_t\, C_t^T + C_t^T)\, (C_t\, \bar{\Sigma}_t\, C_t^T + Q_t)^{-1} \\
&= \Sigma_t\, (C_t^T\, Q_t^{-1}\, C_t\, \bar{\Sigma}_t\, C_t^T + \underbrace{\bar{\Sigma}_t^{-1}\, \bar{\Sigma}_t}_{=I}\, C_t^T)\, (C_t\, \bar{\Sigma}_t\, C_t^T + Q_t)^{-1} \\
&= \Sigma_t\, \underbrace{(C_t^T\, Q_t^{-1}\, C_t + \bar{\Sigma}_t^{-1})}_{=\Sigma_t^{-1}}\, \bar{\Sigma}_t\, C_t^T\, (C_t\, \bar{\Sigma}_t\, C_t^T + Q_t)^{-1} \\
&= \underbrace{\Sigma_t\, \Sigma_t^{-1}}_{=I}\, \bar{\Sigma}_t\, C_t^T\, (C_t\, \bar{\Sigma}_t\, C_t^T + Q_t)^{-1} \\
&= \bar{\Sigma}_t\, C_t^T\, (C_t\, \bar{\Sigma}_t\, C_t^T + Q_t)^{-1}
\end{aligned}
$$

Posterior State Covariance

- Covariance:

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

- Matrix Inversion Lemma

$$(R + P \, Q \, P^T)^{-1} \quad = \quad R^{-1} - R^{-1} \, P \, (Q^{-1} + P^T \, R^{-1} \, P)^{-1} \, P^T \, R^{-1}$$

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

$$\begin{aligned} \Sigma_t &= \bar{\Sigma}_t - \bar{\Sigma}_t \, C_t^T \, (Q_t + C_t \, \bar{\Sigma}_t \, C_t^T)^{-1} \, C_t \, \bar{\Sigma}_t \\ &= [I - \underbrace{\bar{\Sigma}_t \, C_t^T \, (Q_t + C_t \, \bar{\Sigma}_t \, C_t^T)^{-1}}_{K_t} \, C_t] \, \bar{\Sigma}_t \end{aligned}$$

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

Kalman Filter Algorithm

Algorithm Kalman_filter($\mu_{t-1}, \Sigma_{t-1}, u_t, z_t$**):**

$$\bar{\mu}_t = A_t \, \mu_{t-1} + B_t \, u_t$$
$$\bar{\Sigma}_t = A_t \, \Sigma_{t-1} \, A_t^T + R_t$$

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

return μ_t, Σ_t

Kalman Filetr Example: Falling body problem

Consider an object falling under a constant gravitational field. Let y(t) denote the height of the object, then

$$\ddot{y}(t) = -g$$

$$\Rightarrow \dot{y}(t) = \dot{y}(t_0) - g(t - t_0)$$

$$\Rightarrow y(t) = y(t_0) + \dot{y}(t_0)(t - t_0) - \frac{g}{2}(t - t_0)^2$$

As a discrete time system with time increment of $t\text{-}t_0\text{=}1$

$$y(k+1) = y(k) + \dot{y}(k) - \frac{g}{2}$$

Kalman Filter Example: Falling body problem

- State:

$$\mathbf{x}(\mathrm{k}) \equiv [\mathrm{y}(\mathrm{k}) \quad \dot{\mathrm{y}}(\mathrm{k})]'$$

- Process Model:

$$\mathbf{x}(\mathrm{k}+1) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \mathbf{x}(\mathrm{k}) + \begin{bmatrix} 0.5 \\ 1 \end{bmatrix} (-g)$$

$$= \mathrm{A}\ \mathrm{x}(\mathrm{k}) + \mathrm{B}\ \mathrm{u}$$

Kalman Filter Example: Falling body problem

- Measurement: Assuming that we can measure the height of the ball directly.

$$\mathbf{z}(k) = [1 \;\; 0]\, \mathbf{x}(k) + w(k)$$

Measurement Noise

$$= C\, x(k) + w(k)$$