Introduction to Kalman Filter

Po-Chen Wu

Media IC and System Lab

Graduate Institute of Electronics Engineering

National Taiwan University

Outline

- Introduction to Kalman Filter
- Conceptual Overview
- The Theory of Kalman Filter
- Simple Example
- Extended Kalman Filter

Takeaways

- What is a Kalman Filter?
- Why do we need Kalman Filters?

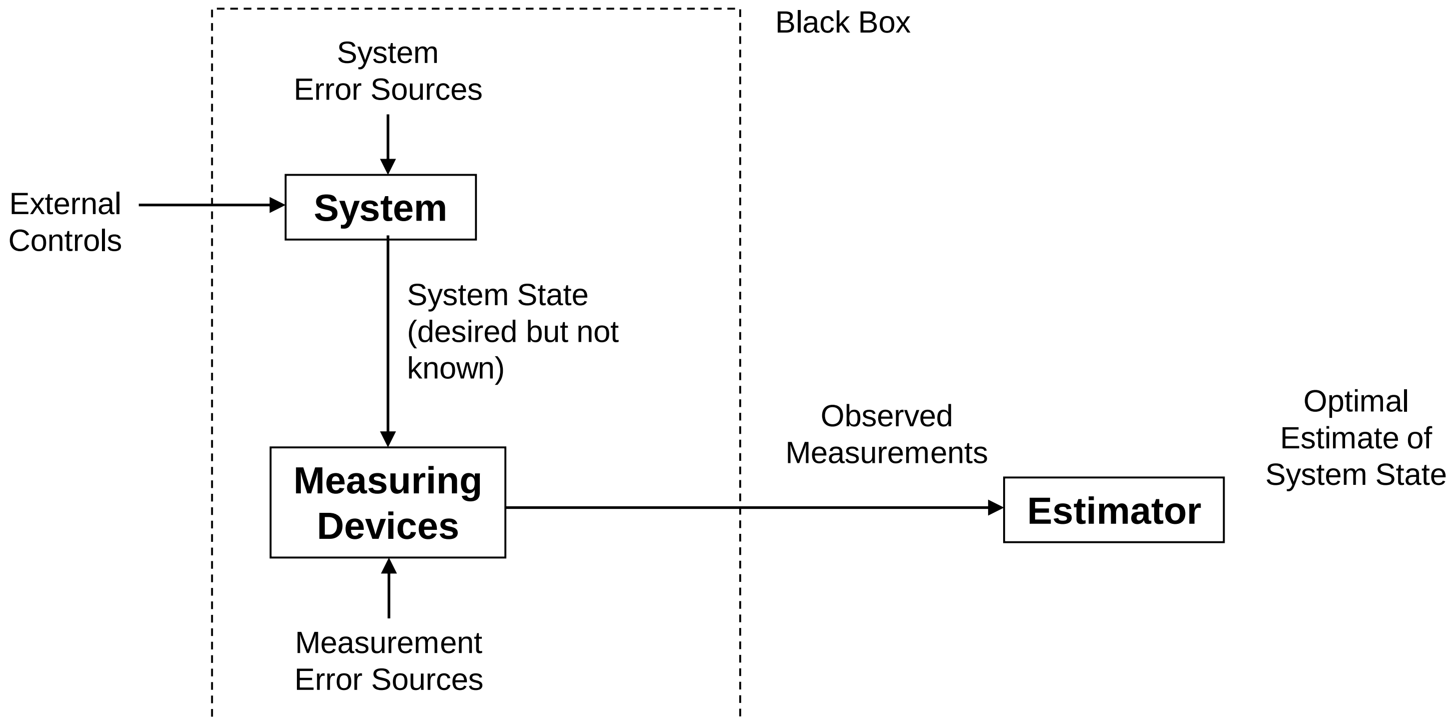
Outline

- Introduction to Kalman Filter
- Conceptual Overview
- The Theory of Kalman Filter
- Simple Example
- Extended Kalman Filter

Introduction

- Recursive data processing algorithm
  - Generates optimal estimate of desired quantities given the set of measurements
  - Optimal: For linear system and white Gaussian errors, Kalman filter is "best" estimate based on all previous measurements.
  - Recursive: doesn't need to store all previous measurements and reprocess all data each time step.

The Problem

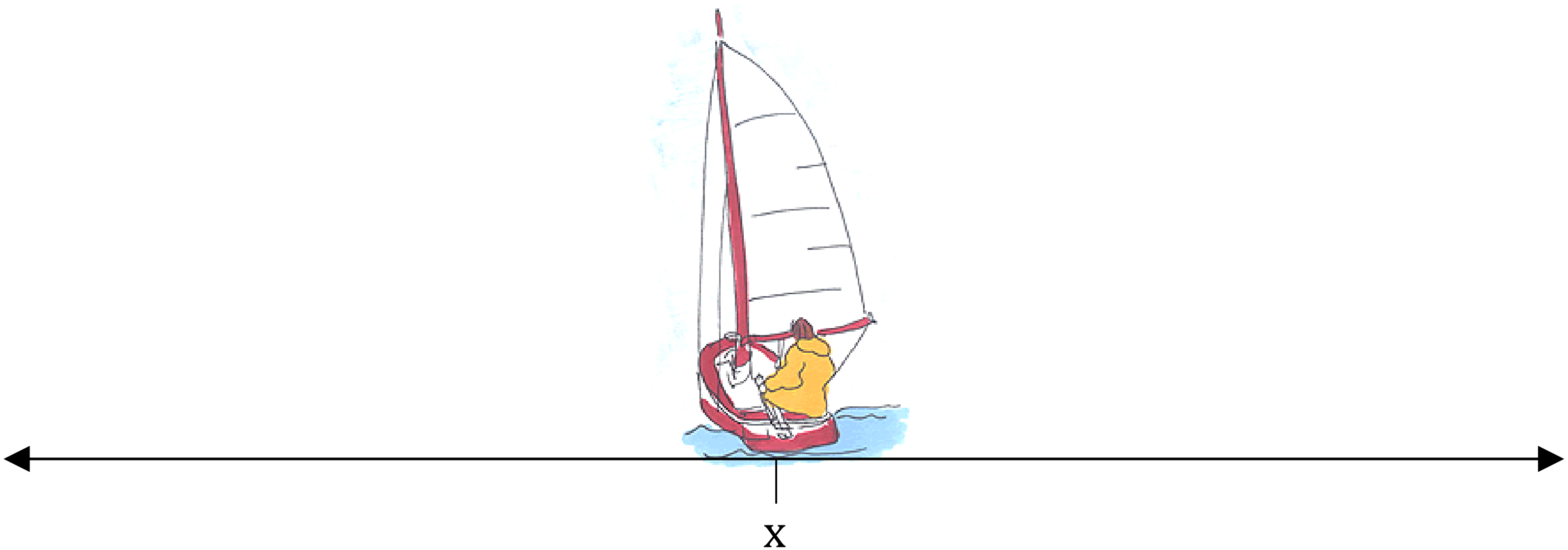


- System state cannot be measured directly.
- Need to estimate “optimally” from measurements.

Outline

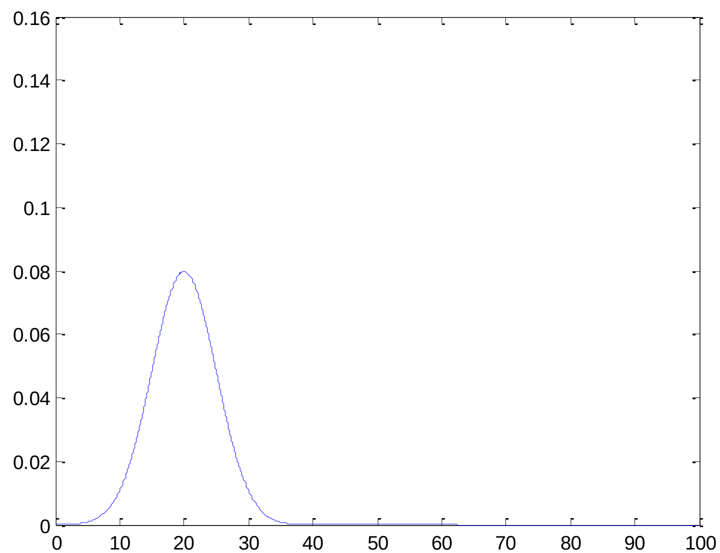
- Introduction to Kalman Filter
- Conceptual Overview
- The Theory of Kalman Filter
- Simple Example
- Extended Kalman Filter

Conceptual Overview (1/9)



- Lost on the 1-dimensional line
- Position – $x(t)$
- Assume Gaussian distributed measurements

Conceptual Overview (2/9)

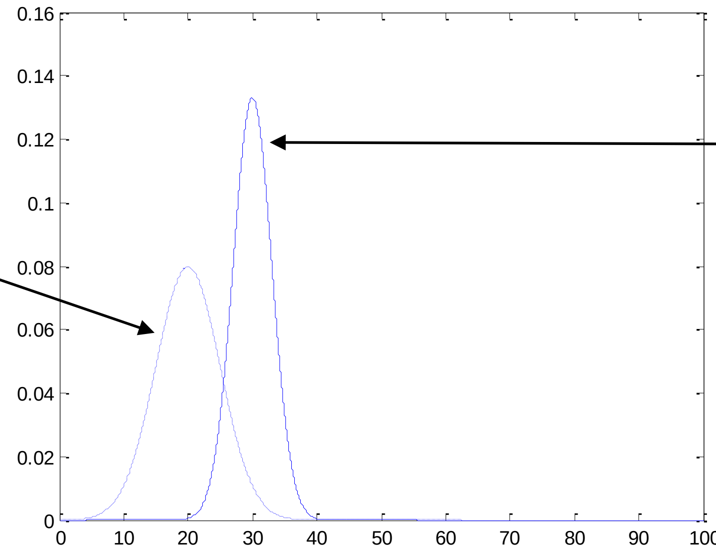


- Sextant Measurement at t_1: Mean = z_1 and Variance = σ_{z_1}
- Optimal estimate of position is: $\hat{x}(t_1) = z_1$
- Variance of error in estimate: $\sigma_x^2(t_1) = \sigma_{z_1}^2$
- Boat in same position at time t_2 - Predicted position is z_1

Conceptual Overview (3/9)

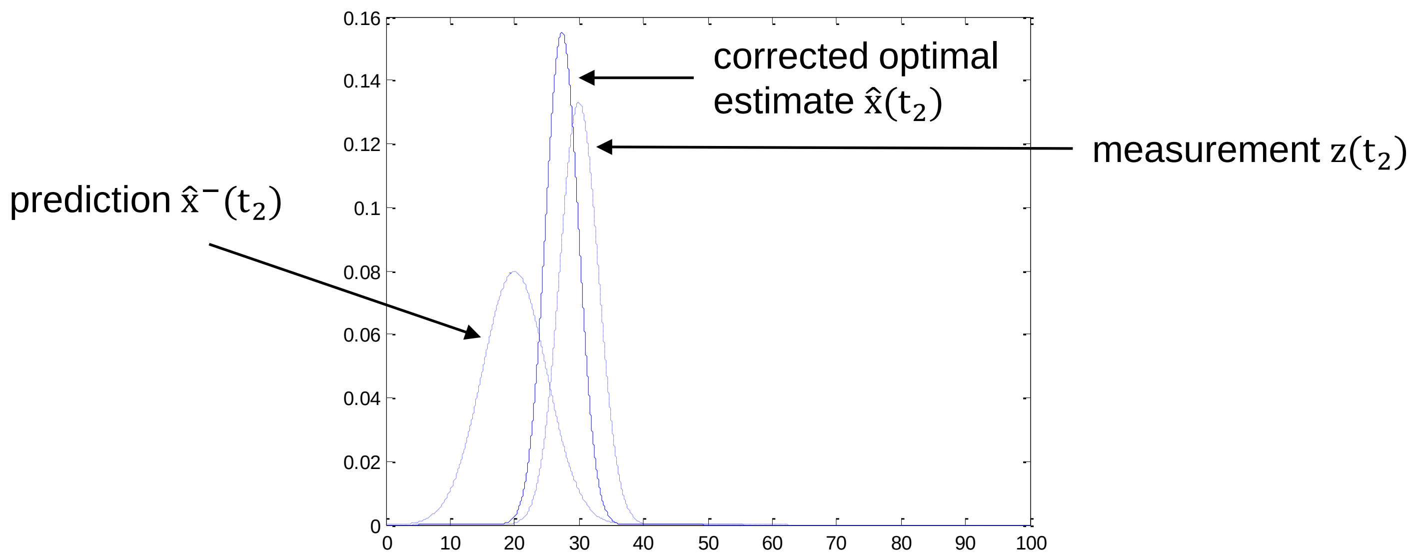
prediction $\hat{x}^{-}(t_2)$

measurement $z(t_2)$



- So we have the prediction $\hat{y}^{-}(t_2)$
- GPS Measurement at t_2: Mean = z_2 and Variance = σ_{z_2}
- Need to correct the prediction due to measurement to get $\hat{x}(t_2)$
- Closer to more trusted measurement – linear interpolation?

Conceptual Overview (4/9)



- Corrected mean is the new optimal estimate of position
- New variance is smaller than either of the previous two variances

Conceptual Overview (5/9)

- Lessons so far:

Make prediction based on previous data: $\hat{x}^-$, σ^-

↓

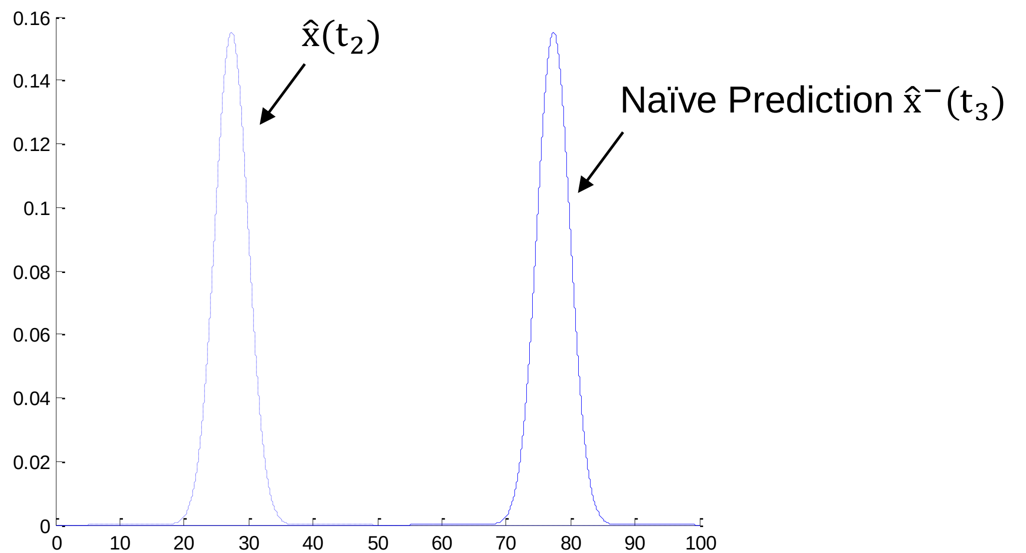
Take measurement: z_k, σ_z

↓

Optimal estimate ($\hat{y}$) = Prediction + (Kalman Gain) * (Measurement - Prediction)

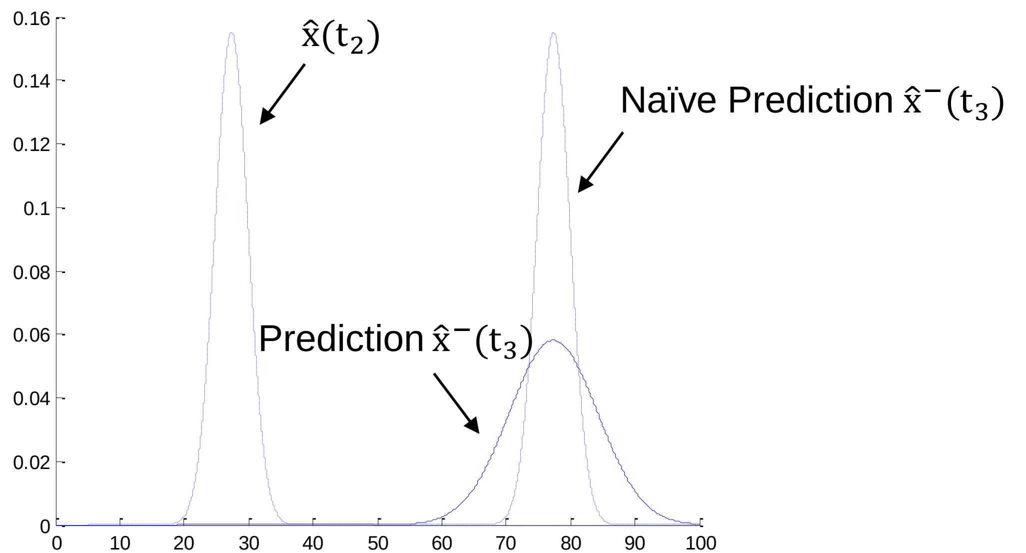
Variance of estimate = Variance of prediction * (1 – Kalman Gain)

Conceptual Overview (6/9)



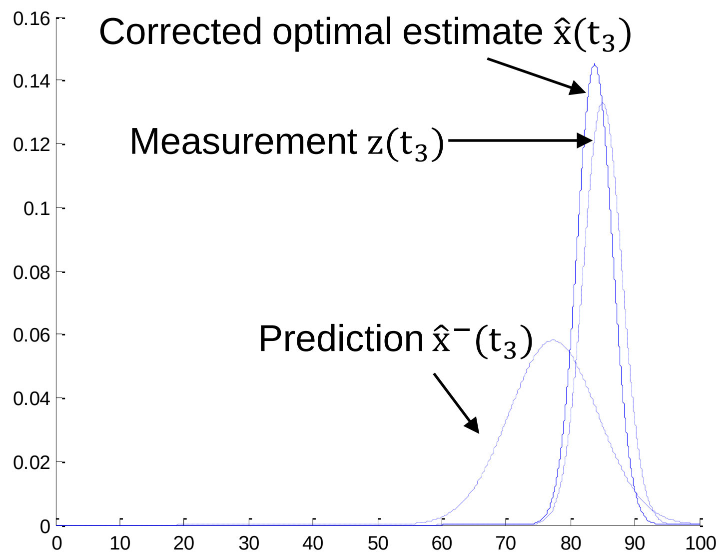
- At time t_3, boat moves with velocity $\frac{dx}{dt} = u$
- Naïve approach: Shift probability to the right to predict
- This would work if we knew the velocity exactly (perfect model)

Conceptual Overview (7/9)



- Better to assume imperfect model by adding Gaussian noise
- $\frac{dx}{dt} = u + w$
- Distribution for prediction moves and spreads out

Conceptual Overview (8/9)



- Now we take a measurement at t_3
- Need to once again correct the prediction
- Same as before

Conceptual Overview (9/9)

- Lessons learnt from conceptual overview:
  - Initial conditions ($\hat{x}_{k-1}$ and σ_{k-1})
  - Prediction ($\hat{x}_k^-$, σ_k^-)
    - Use initial conditions and model (eg. constant velocity) to make prediction
  - Measurement (z_k)
    - Take measurement
  - Correction ($\hat{x}_k$, σ_k)
    - Use measurement to correct prediction by 'blending' prediction and residual – always a case of merging only two Gaussians
    - Optimal estimate with smaller variance

Outline

- Introduction to Kalman Filter
- Conceptual Overview
- The Theory of Kalman Filter
- Simple Example
- Extended Kalman Filter

Theoretical Basis

- Process to be estimated:

$x_k = Ax_{k-1} + Bu_k + w_k$ Process Noise (w) with covariance Q

$z_k = Hx_k + v_k$ Measurement Noise (v) with covariance R

- Kalman Filter

Predicted: $\hat{x}_k^-$ is estimate based on measurements at previous time-steps

$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_k$ Predicted (a priori) state estimate

$P_k^- = AP_{k-1}A^T + Q$ Predicted (a priori) estimate covariance

Corrected: $\hat{x}_k$ has additional information – the measurement at time k

$\hat{x}_k = \hat{x}_k^- + K(z_k - H\hat{x}_k^-)$

$P_k = (1 - KH)P_k^-$

$$K = P_k^- H^T (HP_k^- H^T + R)^{-1}$$

Optimal Kalman gain

Blending Factor

- If we are sure about measurements:
  - Measurement error covariance (R) decreases to zero
  - K increases and weights residual more heavily than prediction

- If we are sure about prediction
  - Prediction error covariance P_k^- decreases to zero
  - K decreases and weights prediction more heavily than residual

Theoretical Basis

Initial estimates for $\hat{x}_{k-1}$ and P_{k-1}

Prediction (Time Update)

(1) Project the state ahead

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_k$$

(2) Project the error covariance ahead

$$P_k^- = AP_{k-1}A^T + Q$$

Correction (Measurement Update)

(1) Compute the Kalman Gain

$$K = P_k^- H^T (HP_k^- H^T + R)^{-1}$$

(2) Update estimate with measurement z_k

$$\hat{x}_k = \hat{x}_k^- + K(z_k - H\hat{x}_k^-)$$

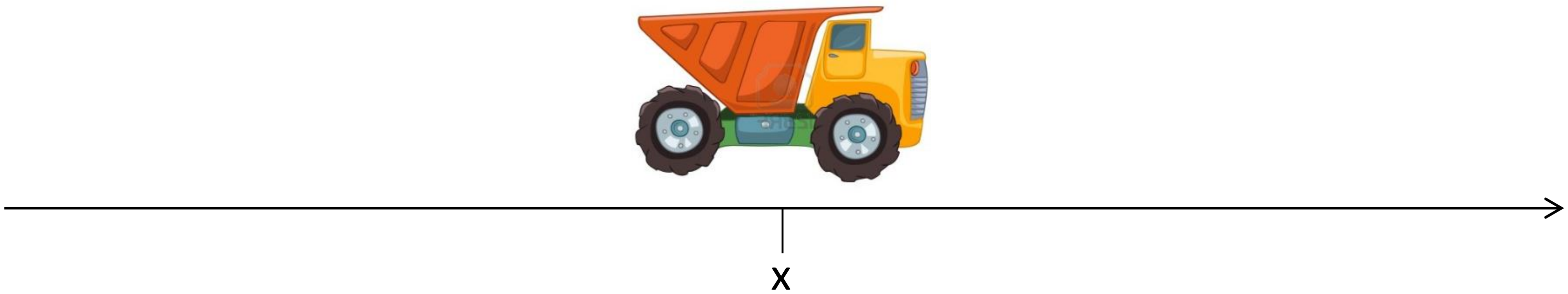
(3) Update Error Covariance

$$P_k = (1 - KH)P_k^-$$

Outline

- Introduction to Kalman Filter
- Conceptual Overview
- The Theory of Kalman Filter
- Simple Example
- Extended Kalman Filter

Simple Example (1/4)



- Consider a truck on perfectly frictionless, infinitely long straight roads.
  - Initially the truck is stationary at position 0, but it is buffeted this way and that by random acceleration.
  - We measure the position of the truck every Δt seconds, but these measurements are imprecise; we want to maintain a model of where the truck is and what its velocity is.
  - We show here how we derive the model from which we create our Kalman filter.

Simple Example (2/4)

- The position and velocity of the truck are described by the linear state space:

$$\mathbf{x}_k = \begin{bmatrix} x \\ \dot{x} \end{bmatrix}$$

- We assume that between the (k−1) and k timestep the truck undergoes a constant acceleration of a_k that is normally distributed, with mean 0 and standard deviation σ_a.

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{G}a_k$$

$$\mathbf{A} = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix}, \mathbf{G} = \begin{bmatrix} \frac{\Delta t^2}{2} \\ \Delta t \end{bmatrix}$$

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{w}_k$$

$$\mathbf{w}_k \sim N(0, \mathbf{Q})$$

$$\mathbf{Q} = \mathbf{G}\mathbf{G}^{\mathrm{T}}\sigma_a^2 = \begin{bmatrix} \frac{\Delta t^4}{4} & \frac{\Delta t^3}{2} \\ \frac{\Delta t^3}{2} & \Delta t^2 \end{bmatrix}$$

Simple Example (3/4)

- At each time step, a noisy measurement of the true position of the truck is made.
- Let us suppose the measurement noise $\mathbf{v}_k$ is also normally distributed, with mean 0 and standard deviation σ_z.

$$\mathbf{z}_k = \mathbf{H}\mathbf{x}_k + \mathbf{v}_k$$

$$\mathbf{H} = [\mathbf{1} \quad \mathbf{0}]$$

$$\mathbf{v}_k \sim N(0, \mathbf{R})$$
$$\mathbf{R} = [\sigma_z^2]$$

Simple Example (4/4)

- We know the initial starting state of the truck with perfect precision, so we initialize

$$\hat{\mathbf{x}}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

- And to tell the filter that we know the exact position, we give it a zero covariance matrix:

$$\mathbf{P}_0 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad \left(\mathbf{P}_0 = \begin{bmatrix} L & 0 \\ 0 & L \end{bmatrix} \right)$$

- If the initial position and velocity are not known perfectly the covariance matrix should be initialized with a suitably large number, say L, on its diagonal.

Outline

- Introduction to Kalman Filter
- Conceptual Overview
- The Theory of Kalman Filter
- Simple Example
- Extended Kalman Filter

Extended Kalman Filter

- The basic Kalman filter is limited to a linear assumption.
- The extended Kalman filter (EKF) is the nonlinear version of the Kalman filter which linearizes about an estimate of the current mean and covariance.

$$\mathrm{x_k = Ax_{k-1} + Bu_k + w_k} \quad \Longrightarrow \quad \mathrm{x_k} = f(\mathrm{x_{k-1} + u_k}) + \mathrm{w_k}$$

$$\mathrm{z_k = Hx_k + v_k} \quad \Longrightarrow \quad \mathrm{z_k} = h(\mathrm{x_k}) + \mathrm{v_k}$$

Theoretical Basis

Initial estimates for $\hat{x}_{k-1}$ and P_{k-1}

Prediction (Time Update)

(1) Project the state ahead

$$\hat{x}_k^- = f(\hat{x}_{k-1} + u_k)$$

(2) Project the error covariance ahead

$$P_k^- = A_k P_{k-1} A_k^T + Q$$

Correction (Measurement Update)

(1) Compute the Kalman Gain

$$K = P_k^- H_k^T (H_k P_k^- H_k^T + R)^{-1}$$

(2) Update estimate with measurement z_k

$$\hat{x}_k = \hat{x}_k^- + K(z_k - h(\hat{x}_k^-))$$

(3) Update Error Covariance

$$P_k = (1 - K H_k) P_k^-$$

$$A_k = \left.\frac{\partial f}{\partial x}\right|_{\hat{x}_{k-1}, u_k} \qquad H_k = \left.\frac{\partial h}{\partial x}\right|_{\hat{x}_k^-}$$

Reference

- Bishop, Gary, and Greg Welch. "An introduction to the Kalman filter." *Proc of SIGGRAPH, Course* 8.27599-23175 (2001): 41.
- Michael Williams. "Introduction to Kalman Filters." Powerpoint Slides, 5 June 2003