# Lecture 9, Oct 5, 2023

## Multidimensional Kalman Filtering

- We can generalize our model to multiple degrees of freedom with a separate measurement relation:
  - $\boldsymbol{x}_{k+1} = \boldsymbol{A}_k\boldsymbol{x}_k + \boldsymbol{B}_k\boldsymbol{u}_k + \boldsymbol{v}_k$ (the *state*, or *process equation*)
  - $\boldsymbol{z}_k + \boldsymbol{D}_k\boldsymbol{x}_k + \boldsymbol{w}_k$ (the *measurement model*)
  - Where $\boldsymbol{A}_k$ is the state update matrix, $\boldsymbol{B}_k$ is the control matrix, $\boldsymbol{z}_k$ is the measurement, $\boldsymbol{D}_k$ is the measurement matrix, $\boldsymbol{v}_k$ is the process (or model) noise and $\boldsymbol{w}_k$ is the measurement noise
  - Again assume $\boldsymbol{v}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}_k), \boldsymbol{w}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_k)$, i.e. zero-mean noise with covariances $\boldsymbol{Q}_k, \boldsymbol{R}_k$
  - In practice we can find $\boldsymbol{Q}_k$ and $\boldsymbol{R}_k$ through testing and characterization of the system; depending on the model, we may be able to estimate it mathematically
- Let $\hat{\boldsymbol{x}}_{k|j}$ be the estimate of $\boldsymbol{x}_k$ given measurements $\{\, \boldsymbol{z}_0, \dots, \boldsymbol{z}_j \,\}$, with $\boldsymbol{P}_{k|j}$ as its covariance
- Kalman filtering is a two-branch process divided into state and covariance estimations:
  - Given $\hat{\boldsymbol{P}}_{k|k}$ (the previous best estimate), $\boldsymbol{u}_k$ (the control input), $\boldsymbol{P}_{k|k}$ (the previous covariance)
  - State estimation:
    1. Predict the next state: $\hat{\boldsymbol{x}}_{k+1|k} = \boldsymbol{A}_k\hat{\boldsymbol{x}}_{k|k} + \boldsymbol{B}_k\boldsymbol{u}_k$
    2. Predict the next measurement: $\hat{\boldsymbol{z}}_{k+1|k} = \boldsymbol{D}_{k+1}\hat{\boldsymbol{x}}_{k+1|k}$
    3. Calculate the measurement residual: $\boldsymbol{s}_{k+1} = \boldsymbol{z}_{k+1} - \hat{\boldsymbol{z}}_{k+1|k}$
    4. Update the state estimate: $\hat{\boldsymbol{x}}_{k+1|k+1} = \hat{\boldsymbol{x}}_{k+1|k} + \boldsymbol{W}_{k+1}\boldsymbol{s}_{k+1}$
  - Covariance estimation:
    1. Predict the state covariance: $\boldsymbol{P}_{k+1|k} = \boldsymbol{A}_k\boldsymbol{P}_{k|k}\boldsymbol{A}_k^T + \boldsymbol{Q}_k$
    2. Predict the measurement covariance: $\boldsymbol{S}_{k+1} = \boldsymbol{D}_{k+1}\boldsymbol{P}_{k+1|k}\boldsymbol{D}_{k+1}^T + \boldsymbol{R}_{k+1}$
    3. Calculate the Kalman gain: $\boldsymbol{W}_{k+1} = \boldsymbol{P}_{k+1|k}\boldsymbol{D}_{k+1}^T\boldsymbol{S}_{k+1}^{-1}$
    4. Update the state covariance: $\boldsymbol{P}_{k+1|k+1} = \boldsymbol{P}_{k+1|k} - \boldsymbol{W}_{k+1}\boldsymbol{S}_{k+1}\boldsymbol{W}_{k+1}^T$
- The state covariance estimation part can be combined into a single equation, which is known as the *Ricatti equation*
- This only works if we have noise – if $\boldsymbol{R}_{k+1} = \boldsymbol{0}$, often $\boldsymbol{S}_{k+1}$ is not invertible; however if $\boldsymbol{R}_{k+1}$ is invertible, then due to the positive-definiteness of $\boldsymbol{P}_{k+1|k}$, we are guaranteed that $\boldsymbol{S}$ is invertible
- Example: body in free fall
  - Model: $\begin{bmatrix} x_{k+1} \\ v_{k+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_k \\ v_k \end{bmatrix} + \begin{bmatrix} -\frac{1}{2}g \\ -g \end{bmatrix}$
  - We will measure the height: $z_k = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_k \\ v_k \end{bmatrix} + w_k$
  - Take noise to be $\boldsymbol{Q}_k = \boldsymbol{0}, \boldsymbol{R}_k = 1$

## Optimality of Kalman Filtering

- If we expand the update relations we get:
  - $\hat{\boldsymbol{x}}_{k+1} = \hat{\boldsymbol{x}}_{k+1|k} + \boldsymbol{W}_{k+1}(\boldsymbol{D}_{k+1}\boldsymbol{x}_{k+1} + \boldsymbol{w}_{k+1} - \boldsymbol{D}_{k+1}\hat{\boldsymbol{x}}_{k+1|k})$
  - $P_{k+1|k+1} = (\boldsymbol{1} - \boldsymbol{W}_{k+1}\boldsymbol{D}_{k+1})(\boldsymbol{A}_k\boldsymbol{P}_{k|k}\boldsymbol{A}_k + \boldsymbol{Q}_k)(\boldsymbol{1} - \boldsymbol{W}_{k+1}\boldsymbol{D}_{k+1})^T + \boldsymbol{W}_{k+1}\boldsymbol{R}_{k+1}\boldsymbol{W}_{k+1}^T$
- We would want to minimize $\varepsilon_{k+1}^2 = \mathbb{E}\left[\|\hat{\boldsymbol{x}}_{k+1|k+1} - \boldsymbol{x}_{k+1}\|^2\right]$, i.e. the expected error
  - $$\begin{aligned} \varepsilon_{k+1}^2 &= \mathbb{E}\left[\|\hat{\boldsymbol{x}}_{k+1|k+1} - \boldsymbol{x}_{k+1}\|^2\right] \\ &= \mathbb{E}\left[(\hat{\boldsymbol{x}}_{k+1|k+1} - \boldsymbol{x}_{k+1})^T(\hat{\boldsymbol{x}}_{k+1|k+1} - \boldsymbol{x}_{k+1})\right] \\ &= \operatorname{tr}\mathbb{E}\left[(\hat{\boldsymbol{x}}_{k+1|k+1} - \boldsymbol{x}_{k+1})(\hat{\boldsymbol{x}}_{k+1|k+1} - \boldsymbol{x}_{k+1})^T\right] \\ &= \operatorname{tr}\boldsymbol{P}_{k+1|k+1} \end{aligned}$$
  - This means that to minimize the expected error, we should minimize the covariance
- To minimize the error, we solve for $\dfrac{\partial \varepsilon_{k+1}^2}{\partial \boldsymbol{W}_{k+1}} = \boldsymbol{0}$ to get the optimal $\boldsymbol{W}$
  - Note: $\dfrac{\partial \operatorname{tr}\boldsymbol{A}\boldsymbol{B}}{\partial \boldsymbol{B}} = \boldsymbol{A}^T$ for matrices $\boldsymbol{A}, \boldsymbol{B}$

  - If we do this, we get $\dfrac{\partial \operatorname{tr} \boldsymbol{P}_{k+1|k+1}}{\partial \boldsymbol{W}_{k+1}} = -2\boldsymbol{P}_{k+1|k+1}\boldsymbol{D}_{k+1}^T + 2\boldsymbol{W}_{k+1}\boldsymbol{S}_{k+1} = \boldsymbol{0}$ where $\boldsymbol{S}_{k+1}$ is defined above
  - Assuming $\boldsymbol{S}_{k+1}$ is invertible, solve to get $\boldsymbol{W}_{k+1} = \boldsymbol{P}_{k+1|k}\boldsymbol{D}_{k+1}^T\boldsymbol{S}_{k+1}^{-1}$
- Hence Kalman filtering is an *optimal* estimator

## Extended Kalman Filtering (EKF)

- What if we didn't have a linear process/measurement model?
- In general, we can have $\boldsymbol{x}_{k+1} = \boldsymbol{f}(\boldsymbol{x}_k, \boldsymbol{u}_k) + \boldsymbol{v}_k, \boldsymbol{z}_{k+1} = \boldsymbol{h}(\boldsymbol{x}_{k+1}) + \boldsymbol{w}_{k+1}$
  - Note we are assuming that noise is additive right now
- We simply linearize the system with the Jacobian
- For the predictions, we can directly do $\hat{\boldsymbol{x}}_{k+1|k} = \boldsymbol{f}(\hat{\boldsymbol{x}}_{k|k}, \boldsymbol{u}_k)$ and $\hat{\boldsymbol{z}}_{k+1|k} = \boldsymbol{h}(\hat{\boldsymbol{x}}_{k|k})$
- For the state covariance estimate, we will linearize about $\hat{\boldsymbol{x}}_{k+1|k}, \boldsymbol{u}_k$:
  - $\boldsymbol{A}_k = \left.\dfrac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}_k^T}\right|_{\hat{\boldsymbol{x}}_{k+1|k}, \boldsymbol{u}_k}, \boldsymbol{B}_k = \left.\dfrac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}_k^T}\right|_{\hat{\boldsymbol{x}}_{k+1|k}, \boldsymbol{u}_k}, \boldsymbol{D}_{k+1} = \left.\dfrac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}_k^T}\right|_{\hat{\boldsymbol{x}}_{k+1|k}}$
    - Note that now the matrices such as $\boldsymbol{A}_k$ are dependent on our state!
- The procedure is identical to that of normal Kalman filtering, except the nonlinear model is used for prediction and measurement, while the linearized Jacobians are used for covariance estimation