ECE 7251: Signal Detection and Estimation

Spring 2002

Prof. Aaron Lanterman
Georgia Institute of Technology

Lecture 16, 2/13/02:
The Kalman Filter

The Setup

$$\Theta_{k+1}^{[m\times1]} = A^{[m\times m]}\Theta_k^{[m\times1]} + U_k^{[m\times1]}$$ State equation

$$Y_k^{[n\times1]} = C^{[n\times m]}\Theta_k^{[m\times1]} + W_k^{[n\times1]}$$ Measurement eqn.

$U_k \sim \mathcal{N}(0, \underbrace{K_U}_{})$ — Process "noise" covariance

$W_k \sim \mathcal{N}(0, \underbrace{K_W}_{})$ — Measurement noise covariance

$\Theta_1 \sim \mathcal{N}(\underbrace{\hat{q}_{1|0}}_{}, \underbrace{P_{1|0}}_{})$

- $\hat{q}_{1|0}$: Initial guess, before taking any data
- $P_{1|0}$: Covariance indicating confidence of initial guess

$U_k,\ W_k,\ \Theta_1$ are uncorrelated with each other and for different k

Goal

- Goal: Find MMSE (conditional mean) estimates

$$\hat{q}_{k|k}(y) \equiv E[\Theta_k \mid y_{1,} \ldots y_k]$$

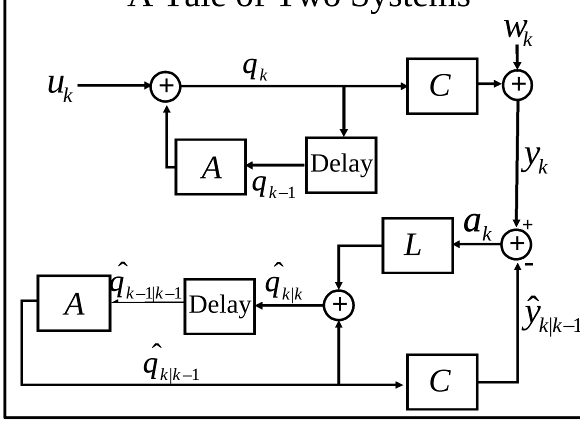
$$\hat{q}_{k+1|k}(y) \equiv E[\Theta_{k+1} \mid y_{1,} \ldots y_k] \quad \text{(Prediction)}$$

- Define

$$P_{k|k} \equiv \underbrace{E[(\Theta_k - \hat{q}_{k|k}(Y))(\Theta_k - \hat{q}_{k|k}(Y))^T]}_{\text{Filter error covariance}}$$

$$P_{k+1|k} \equiv \underbrace{E[(\Theta_{k+1} - \hat{q}_{k+1|k}(Y))(\Theta_{k+1} - \hat{q}_{k+1|k}(Y))^T]}_{\text{Prediction covariance}}$$

A Tale of Two Systems



Step 1A: Initialize State

- We've taken our first data point:

$$\hat{q}_{1|1} = E[\Theta_1 \mid y_1]$$
$$= E[\Theta_1] - \mathrm{Cov}[\Theta_1, Y_1]\mathrm{Cov}^{-1}[Y_1](y_1 - E[Y_1])$$
$$= E[\Theta_1] + P_{1|0}C^T(CP_{1|0}C^T + K_W)^{-1}(y_1 - C\hat{q}_{1|0})$$

- Recall that: $Y_1 = C\Theta_1 + W_1$

$$\Theta_1 \sim \mathcal{N}(\hat{q}_{1|0}, P_{1|0}),\ W \sim \mathcal{N}(0, K_W)$$

- Hence:

$$Y_1 \sim \mathcal{N}(C\hat{q}_{1|0}, CP_{1|0}C + K_W)$$
$$Cov[\Theta_1, Y_1] = P_{1|0}C^T$$

Step 1B: Initialize Covariance

- What is our confidence now that we've collected some data?

$$P_{1|1} \equiv E[(\Theta_1 - \hat{\Theta}_{1|1})(\Theta_1 - \hat{\Theta}_{1|1})^T]$$
$$= \mathrm{Cov}[\Theta_1] - \mathrm{Cov}[\Theta_1, Y_1]\mathrm{Cov}^{-1}[Y_1]\mathrm{Cov}[Y_1, \Theta_1]$$
$$= P_{1|0} - P_{1|0}C^T(CP_{1|0}C^T + K_W)^{-1}CP_{1|0}$$

Step 2: Prediction

$$\hat{q}_{k+1|k} \equiv E[\Theta_{k+1} \mid y_{1,} \ldots y_k] = A\hat{q}_{k|k}$$

$$P_{k+1|k} = AP_{k|k}A^T + K_U$$

- Recall that: $\Theta_{k+1} = A\Theta_k + U_k$
- Hence:

$$\Theta_{k+1|k} \sim \mathcal{N}(A\hat{q}_{k|k}, AP_{k|k}A + K_U)$$

Step 3A: State Update

- Recall $Y_k = C\Theta_k + W_k$
- Consider predicted data

$$\underbrace{Y_{k+1|k}}_{\text{Conditioned on all data up to } k} \sim N(C\hat{q}_{k+1|k}, CP_{k+1|k}C + K_W)$$

$$\hat{q}_{k+1|k+1} = E[\Theta_{k+1|k}] - \mathrm{Cov}[\Theta_{k+1}, Y_{k+1|k}] \times \mathrm{Cov}^{-1}[Y_{k+1|k}](y_{k+1} - E[Y_{k+1|k}])$$

$$= \underbrace{\hat{q}_{k+1|k}}_{=A\hat{q}_k} + \underbrace{P_{k+1|k}C^T(CP_{k+1|k}C^T + K_W)^{-1}}_{\text{Kalman Gain}}(y_{k+1} - C\underbrace{\hat{q}_{k+1|k}}_{=A\hat{q}_k})$$

Step 3B: Covariance Update

- Recall $Y_k = Cq_k + W_k$
- Consider predicted data

$$\underbrace{Y_{k+1|k}}_{\text{Conditioned on all data up to } k} \sim N(C\hat{q}_{k+1|k}, CP_{k+1|k}C + K_W)$$

$$P_{k+1|k+1} = \mathrm{Cov}[\Theta_{k+1|k}] - \mathrm{Cov}[\Theta_{k+1|k}, Y_{k+1|k}] \times \mathrm{Cov}^{-1}[Y_{k+1|k}]\mathrm{Cov}[\Theta_{k+1|k}, Y_{k+1|k}]$$

$$= P_{k+1|k} - \underbrace{P_{k+1|k}}_{=AP_{k|k}A^T + K_U}C^T(CP_{k+1|k}C^T + K_W)^{-1}CP_{k+1|k}$$

Putting it All Together

- The Kalman filter state update:

$$\underbrace{\hat{q}_{k+1}}_{\text{(dropping } k|k \text{ notation)}} = A\hat{q}_k + L_{k+1}(\underbrace{y_{k+1} - CA\hat{q}_k}_{= a_k \text{ (the innovations)}})$$

- Note that the Kalman gains don't involve the data, and can hence be computed offline ahead of time:

$$L_{k+1} = P_{k+1|k}C^T(CP_{k+1|k}C^T + K_W)^{-1}$$

$$P_{k+1|k} = AP_{k|k}A^T + K_U \quad (\ )$$

$$P_{k+1|k+1} = P_{k+1|k} - P_{k+1|k}C^T(CP_{k+1|k}C^T + K_W)^{-1}CP_{k+1|k} \quad (\ )$$

The Innovation Sequence

- Let the prediction of the data given the previous data be given by

$$\hat{y}_{k+1|k} \equiv E[Y_{k+1} \mid y_1, \ldots, y_k] = CA\hat{q}_k$$

- The <u>innovations</u> are defined as:

$$a_{k+1} = y_{k+1} - CA\hat{q}_k = y_{k+1} - \hat{y}_{k+1|k}$$

- By O.P., innovations orthogonal to the data:

$$E[a_k Y_k^T] = 0$$

- Innovations process is white:

$$E[a_k a_l] = 0 \text{ for } k \neq l$$

- When trying on *real* data, testing innovations for "whiteness" tells you how accurate your models are

Assorted Tidbits

- For non-Gaussian statistics (process and measurement noise), the Kalman filter is the best *linear* MMSE estimator
- Combining and yields the discrete Riccati equation (DRE):

$$P_{k+1|k+1} = \text{messy function of } P_{k|k}$$

- Under some conditions, DRE has a fixed point P_∞, and $P_{k|k} \to P_\infty$; in this case, Kalman filter acts like a <u>Wiener filter</u> for large k
- $P_{k|k}$ tells us our confidence in out state estimates. If it is small, then L_k is small, then the filter is <u>saturated</u>; we pay little attention to measurement.

Easy Variations

- Trivially extended to time-varying matrices A, B, C, K_V, and K_W,
- If u_k and v_k are correlated, form a new system:

$$\Theta_{k+1} = A\Theta_k + SK_W^{-1}(Y_k - C\Theta_k) + \tilde{u}_k$$
$$S = E[UW^T]$$

Measurement noise is now uncorrelated with new process noise; can use all the previous ideas

$SK_W^{-1}Y_k$ is called an "input injection" term