EKF for parameter estimation: 1a–1b

- You have learned how to use SPKF for parameter estimation
- Now, you will learn how to use EKF for parameter estimation

EKF step 1a: Parameter prediction time update

- Due to the linearity of the parameter dynamics equation, we have $\hat{\theta}_k^- = \hat{\theta}_{k-1}^+$ (same as for SPKF)

EKF step 1b: Error covariance time update

- Again, due to the linearity of the parameter dynamics equation, we have

$$\Sigma_{\tilde{\theta},k}^- = \Sigma_{\tilde{\theta},k-1}^+ + \Sigma_{\tilde{r}}$$

(also same as for SPKF)

EKF for parameter estimation: 1c

EKF step 1c: Output prediction

- The system output is predicted to be

$$\begin{aligned}\hat{d}_k &= \mathbb{E}[h(x_k, u_k, \theta, e_k) \mid \mathbb{D}_{k-1}] \\ &\approx h(x_k, u_k, \hat{\theta}_k^-, \bar{e}_k)\end{aligned}$$

- That is, it is assumed that propagating $\hat{\theta}_k^-$ and the mean estimation error is the best approximation to predicting the output

EKF for parameter estimation: 2a

EKF step 2a: Estimator gain matrix

- The output prediction error may then be approximated

$$\tilde{d}_k = d_k - \hat{d}_k = h(x_k, u_k, \theta, e_k) - h(x_k, u_k, \hat{\theta}_k^-, \bar{e}_k)$$

using a Taylor-series expansion on the first term

$$\begin{aligned}d_k &\approx h(x_k, u_k, \hat{\theta}_k^-, \bar{e}_k) \\ &\quad + \underbrace{\left.\frac{\mathrm{d}h(x_k, u_k, \theta, e_k)}{\mathrm{d}\theta}\right|_{\theta=\hat{\theta}_k^-}}_{\text{Defined as } \hat{C}_k^\theta}(\theta - \hat{\theta}_k^-) + \underbrace{\left.\frac{\mathrm{d}h(x_k, u_k, \theta, e_k)}{\mathrm{d}e_k}\right|_{e_k=\bar{e}_k}}_{\text{Defined as } \hat{D}_k^\theta}(e_k - \bar{e}_k)\end{aligned}$$

EKF for parameter estimation: 2a (cont.)

- From this, we can compute such necessary quantities as

$$\begin{aligned}\Sigma_{\tilde{d},k} &\approx \hat{C}_k^\theta \Sigma_{\tilde{\theta},k}^- (\hat{C}_k^\theta)^T + \hat{D}_k^\theta \Sigma_{\tilde{e}} (\hat{D}_k^\theta)^T \\ \Sigma_{\tilde{\theta}\tilde{d},k}^- &\approx \mathbb{E}[(\tilde{\theta}_k^-)(\hat{C}_k^\theta \tilde{\theta}_k^- + \hat{D}_k^\theta \tilde{e}_k)^T] \\ &= \Sigma_{\tilde{\theta},k}^- (\hat{C}_k^\theta)^T\end{aligned}$$

- These terms may be combined to get the Kalman gain

$$L_k^\theta = \Sigma_{\tilde{\theta},k}^- (\hat{C}_k^\theta)^T \big[\hat{C}_k^\theta \Sigma_{\tilde{\theta},k}^- (\hat{C}_k^\theta)^T + \hat{D}_k^\theta \Sigma_{\tilde{e}} (\hat{D}_k^\theta)^T\big]^{-1}$$

EKF for parameter estimation: 2a (cont.)

- Recall distinction between partial and total differentials
- By the chain rule of total differentials,

$$\begin{aligned}\frac{\mathrm{d}h(x_k,u_k,\theta,e_k)}{\mathrm{d}\theta} &= \frac{\partial h(x_k,u_k,\theta,e_k)}{\partial x_k}\frac{\mathrm{d}x_k}{\mathrm{d}\theta} + \frac{\partial h(x_k,u_k,\theta,e_k)}{\partial u_k}\underbrace{\frac{\mathrm{d}u_k}{\mathrm{d}\theta}}_{0} \\ &\quad + \frac{\partial h(x_k,u_k,\theta,e_k)}{\partial \theta}\frac{\mathrm{d}\theta}{\mathrm{d}\theta} + \frac{\partial h(x_k,u_k,\theta,e_k)}{\partial e_k}\underbrace{\frac{\mathrm{d}e_k}{\mathrm{d}\theta}}_{0} \\ &= \frac{\partial h(x_k,u_k,\theta,e_k)}{\partial \theta} + \frac{\partial h(x_k,u_k,\theta,e_k)}{\partial x_k}\frac{\mathrm{d}x_k}{\mathrm{d}\theta}\end{aligned}$$

- But, what is $\mathrm{d}x_k/\mathrm{d}\theta$?

EKF for parameter estimation: 2a (cont.)

- Can evaluate the required total derivative recursively as it evolves over time as the state evolves

$$\frac{\mathrm{d}x_k}{\mathrm{d}\theta} = \frac{\partial f(x_{k-1},u_{k-1},\theta,w_{k-1})}{\partial\theta} + \frac{\partial f(x_{k-1},u_{k-1},\theta,w_{k-1})}{\partial x_{k-1}}\frac{\mathrm{d}x_{k-1}}{\mathrm{d}\theta}$$

- The term $\mathrm{d}x_0/\mathrm{d}\theta$ is initialized to zero unless side information gives a better estimate of its value
- To calculate $\hat{C}_k^\theta$ for any specific model structure, we require methods to calculate all of the above partial derivatives for that model

EKF for parameter estimation: 2b–2c

EKF step 2b: Parameter estimate measurement update

- Compute *a posteriori* parameter estimate by updating *a priori* prediction using estimator gain and output innovation $d_k - \hat{d}_k$

$$\hat{\theta}_k^+ = \hat{\theta}_k^- + L_k^\theta (d_k - \hat{d}_k)$$

EKF step 2c: Error covariance measurement update

- Finally, the updated covariance is computed as

$$\Sigma_{\tilde{\theta},k}^+ = \Sigma_{\tilde{\theta},k}^- - L_k^\theta \Sigma_{\tilde{d},k} (L_k^\theta)^T$$

Summary

- You have now learned how to use EKF for parameter estimation when the state of the system is known
- Unlike EKF for state estimation, the distinction between total and partial derivatives when computing EKF update matrices now matters
- A recursive formulation allows computing $\mathrm{d}x_k/\mathrm{d}\theta$, which we initialize to zero unless side information is available
- We initialize parameter estimate with best information re. the parameter value: $\hat{\theta}_0^+ = \mathbb{E}[\theta_0]$, and parameter estimation error covariance matrix

$$\Sigma_{\tilde{\theta},0}^+ = \mathbb{E}\big[(\theta - \hat{\theta}_0^+)(\theta - \hat{\theta}_0^+)^T\big]$$

- A summary of the method is listed in the appendix

Nonlinear EKF for parameter estimation (1)

State-space model:

$$\begin{aligned}\theta_{k+1} &= \theta_k + r_k,\\ d_k &= h(x_k, u_k, \theta_k, e_k)\end{aligned}$$

where r_k and e_k are independent Gaussian noise processes with means zero and $\bar{e}$, respectively, and having covariance matrices $\Sigma_{\tilde{r}}$ and $\Sigma_{\tilde{e}}$, respectively

Definitions:

$$\hat{C}_k^\theta = \left.\frac{\mathrm{d}h(x_k,u_k,\theta,e_k)}{\mathrm{d}\theta}\right|_{\theta=\hat{\theta}_k^-} \qquad \hat{D}_k^\theta = \left.\frac{\mathrm{d}h(x_k,u_k,\theta,e_k)}{\mathrm{d}e_k}\right|_{e_k=\bar{e}_k}$$

Caution: Be careful to compute $\hat{C}_k^\theta$ using recursive chain rule described in lesson!

Nonlinear EKF for parameter estimation (2)

Initialization: For $k = 0$, set

$$\hat{\theta}_0^+ = \mathbb{E}[\theta_0]$$
$$\Sigma_{\tilde{\theta},0}^+ = \mathbb{E}\big[(\theta_0 - \hat{\theta}_0^+)(\theta_0 - \hat{\theta}_0^+)^T\big]$$
$$\frac{\mathrm{d}x_0}{\mathrm{d}\theta} = 0, \text{ unless side information is available}$$

Computation: For $k = 1, 2, \ldots$ compute:

State time update:	$\hat{\theta}_k^- = \hat{\theta}_{k-1}^+$
Covariance time update:	$\Sigma_{\tilde{\theta},k}^- = \Sigma_{\tilde{\theta},k-1}^+ + \Sigma_{\tilde{r}}$
Output prediction	$\hat{d}_k = h(x_k, u_k, \hat{\theta}_k^-, \bar{e}_k)$
Kalman gain matrix:	$L_k^\theta = \Sigma_{\tilde{\theta},k}^-(\hat{C}_k^\theta)^T[\hat{C}_k^\theta \Sigma_{\tilde{\theta},k}^-(\hat{C}_k^\theta)^T + \hat{D}_k^\theta \Sigma_{\tilde{e}}(\hat{D}_k^\theta)^T]^{-1}$
State measurement update:	$\hat{\theta}_k^+ = \hat{\theta}_k^- + L_k^\theta(d_k - \hat{d}_k)$
Covariance meas. update:	$\Sigma_{\tilde{\theta},k}^+ = \Sigma_{\tilde{\theta},k}^- - L_k^\theta \Sigma_{\tilde{d},k}(L_k^\theta)^T$