Kalman filter framework for estimating parameters

- You know that xKF can estimate state of a dynamic system given known parameters and noisy measurements
- It is also possible to use (nonlinear) KF to estimate parameters given a known state and noisy measurements
- This week, we first consider how to estimate system parameters if state is known
- Then, we will consider how to estimate *both* state and parameters of a system simultaneously using two different approaches

Parameter estimation: State equation

- We denote true parameters of a particular model by θ
- We'll use xKF to estimate $\hat{\theta}_k$ much like we estimated $\hat{x}_k$
  - Therefore, we require a model of the dynamics of the parameters
- By assumption, parameters change slowly, so we model them as constant with some small perturbation:

$$\theta_k = \theta_{k-1} + r_{k-1}$$

- The small white noise input r_k is fictitious, but models the slow drift in system parameter values plus infidelity of model structure

Parameter estimation: Output equation

- Output equation required by xKF must be a measurable function of parameter values: we use

$$d_k = h(x_k, u_k, \theta, e_k),$$

  where $h(\cdot)$ is output equation of the system model being identified, and e_k models the sensor noise and modeling error
- Note that d_k is usually the same measurement as y_k, but we maintain a distinction in case separate outputs are used
  - Then, $\mathbb{D}_k = \{d_0, d_1, \ldots, d_k\}$
  - Also, e_k and v_k often play the same role, but are also considered distinct here

Overall model equations

- Slightly revise the mathematical model of system dynamics

$$\begin{aligned} x_k &= f(x_{k-1}, u_{k-1}, \theta, w_{k-1}) \\ y_k &= h(x_k, u_k, \theta, v_k), \end{aligned}$$

  to include the parameters θ in the model explicitly
- Non-time-varying numeric values required by model may be embedded within $f(\cdot)$ and $h(\cdot)$, and are not included in θ
- Using this nonlinear state-space model, can implement parameter estimation using any nonlinear KF, such as EKF or SPKF

SPKF for parameter estimation: 1a

- Parameter estimation with SPKF is relatively straightforward, so we discuss it before we discuss EKF
- Define augmented θ^a that combines randomness of parameters, sensor noise
- This augmented vector is used in the estimation process as described below
- As always, we proceed by deriving the six steps of sequential probabilistic inference

SPKF step 1a: Parameter prediction time update

- The parameter prediction step is approximated as

$$\hat{\theta}_k^{a,-} = \mathbb{E}[\theta_{k-1}^a + r_{k-1} \mid \mathbb{D}_{k-1}] = \hat{\theta}_{k-1}^{a,+}$$

- This makes sense, since the parameters are assumed constant

SPKF for parameter estimation: 1b

SPKF step 1b: Error covariance time update

- Covariance prediction step is accomplished by first computing

$$\begin{aligned} \tilde{\theta}_k^{a,-} &= \theta_k^a - \hat{\theta}_k^{a,-} = \theta_{k-1}^a + r_k - \hat{\theta}_{k-1}^{a,+} \\ &= \tilde{\theta}_{k-1}^{a,+} + r_k \end{aligned}$$

- Then directly compute desired covariance

$$\begin{aligned} \Sigma_{\tilde{\theta},k}^{a,-} &= \mathbb{E}[\tilde{\theta}_k^{a,-}(\tilde{\theta}_k^{a,-})^T] = \mathbb{E}[(\tilde{\theta}_{k-1}^{a,+} + r_k)(\tilde{\theta}_{k-1}^{a,+} + r_k)^T] \\ &= \Sigma_{\tilde{\theta},k-1}^{a,+} + \Sigma_{\tilde{r}} \end{aligned}$$

- Time-updated covariance has additional uncertainty due to the fictitious noise "driving" the parameter values

SPKF for parameter estimation: 1c

SPKF step 1c: Predict system output d_k

- To predict system output, need sigma points describing output
- This, in turn, requires a set of $p+1$ sigma points describing $\theta_k^{a,-}$, which we will denote as $\mathcal{W}_k^{a,-}$

$$\mathcal{W}_k^{a,-} = \left\{\hat{\theta}_k^{a,-}, \hat{\theta}_k^{a,-} + \gamma\sqrt{\Sigma_{\tilde{\theta},k}^{a,-}}, \hat{\theta}_k^{a,-} - \gamma\sqrt{\Sigma_{\tilde{\theta},k}^{a,-}}\right\}$$

- From the augmented sigma points, the $p+1$ vectors comprising the parameters $\mathcal{W}_k^{\theta,-}$ and the $p+1$ vectors comprising the modeling error $\mathcal{W}_k^{e,-}$ are extracted

SPKF for parameter estimation: 1c (cont.)

- Output equation evaluated using all pairs of $\mathcal{W}_{k,i}^{\theta,-}$ and $\mathcal{W}_{k,i}^{e,-}$ (subscript i denotes that ith vector is being extracted from original set), yielding sigma points $\mathcal{D}_{k,i}$ for time step k

$$\mathcal{D}_{k,i} = h(x_k, u_k, \mathcal{W}_{k,i}^{\theta,-}, \mathcal{W}_{k,i}^{e,-})$$

- Finally, output prediction is computed as

$$\begin{aligned}\hat{d}_k^- &= \mathbb{E}\big[h(x_k, u_k, \theta, e_k) \mid \mathbb{D}_{k-1}\big] \\ &\approx \sum_{i=0}^{p} \alpha_i^{(\mathrm{m})} h(x_k, u_k, \mathcal{W}_{k,i}^{\theta,-}, \mathcal{W}_{k,i}^{e,-}) = \sum_{i=0}^{p} \alpha_i^{(\mathrm{m})} \mathcal{D}_{k,i}\end{aligned}$$

SPKF for parameter estimation: 2a

SPKF step 2a: Estimator gain matrix L_k^{θ}

- To compute estimator gain matrix, must first compute required covariance matrices

$$\Sigma_{\tilde{d},k} = \sum_{i=0}^{p} \alpha_i^{(\mathrm{c})} \big(\mathcal{D}_{k,i} - \hat{d}_k\big)\big(\mathcal{D}_{k,i} - \hat{d}_k\big)^T$$

$$\Sigma_{\tilde{\theta}\tilde{d},k}^- = \sum_{i=0}^{p} \alpha_i^{(\mathrm{c})} \big(\mathcal{W}_{k,i}^{\theta,-} - \hat{\theta}_k^{a,-}\big)\big(\mathcal{D}_{k,i} - \hat{d}_k\big)^T$$

- Then, we simply compute $L_k^{\theta} = \Sigma_{\tilde{\theta}\tilde{d},k}^- \Sigma_{\tilde{d},k}^{-1}$

SPKF for parameter estimation: 2b–2c

SPKF step 2b: Parameter estimate measurement update

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

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

SPKF step 2c: Error covariance measurement update

- The final step is calculated directly from the optimal formulation:

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

Summary

- To use xKF for parameter identification, must first define a relevant nonlinear state-space model
- You have now learned one method that can be used to do so
- You have also learned how to derive the SPKF for parameter estimation
  - Prediction steps similar to, but simpler than SPKF for state estimation
  - Update steps nearly identical to SPKF for state estimation
  - Appendix to this lesson lists all the steps
- Next step is to derive EKF for parameter estimation

Nonlinear SPKF 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:

$$\theta_k^a = \left[\theta_k^T,\ e_k^T\right]^T, \quad \mathcal{W}_k^a = \left[(\mathcal{W}_k^{\theta})^T,\ (\mathcal{W}_k^{e})^T\right]^T, \quad p = 2 \times \dim(\theta_k^a)$$

Nonlinear SPKF for parameter estimation (2)

Initialization: For $k = 0$, set

$$\hat{\theta}_0^+ = \mathbb{E}\big[\theta_0\big]$$
$$\hat{\theta}_0^{a,+} = \mathbb{E}\big[\theta_0^a\big] = \big[(\hat{\theta}_0^+)^T,\ \bar{e}\big]^T \qquad \Sigma_{\tilde{\theta},0}^{a,+} = \mathbb{E}\big[(\theta_0^a - \hat{\theta}_0^{a,+})(\theta_0^a - \hat{\theta}_0^{a,+})^T\big]$$
$$\Sigma_{\tilde{\theta},0}^+ = \mathbb{E}\big[(\theta_0 - \hat{\theta}_0^+)(\theta_0 - \hat{\theta}_0^+)^T\big] \qquad = \operatorname{diag}\big(\Sigma_{\tilde{\theta},0}^+,\ \Sigma_{\tilde{e}}\big)$$

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

State estimate time update: $\hat{\theta}_k^- = \hat{\theta}_{k-1}^+$

Error covariance time update: $\Sigma_{\tilde{\theta},k}^- = \Sigma_{\tilde{\theta},k-1}^+ + \Sigma_{\tilde{r}}$

Output estimate:
$$\mathcal{W}_k^{a,-} = \Big\{\hat{\theta}_k^{a,-},\ \hat{\theta}_k^{a,-} + \gamma\sqrt{\Sigma_{\tilde{\theta},k}^{a,-}},\ \hat{\theta}_k^{a,-} - \gamma\sqrt{\Sigma_{\tilde{\theta},k}^{a,-}}\Big\}$$
$$\mathcal{D}_{k,i} = h(x_k, u_k, \mathcal{W}_{k,i}^{\theta,-}, \mathcal{W}_{k,i}^{e,-})$$
$$\hat{d}_k = \textstyle\sum_{i=0}^{p} \alpha_i^{(\mathrm{m})} \mathcal{D}_{k,i}$$

Nonlinear SPKF for parameter estimation (3)

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

Estimator gain matrix:
$$\Sigma_{\tilde{d},k} = \textstyle\sum_{i=0}^{p} \alpha_i^{(\mathrm{c})}\big(\mathcal{D}_{k,i} - \hat{d}_k\big)\big(\mathcal{D}_{k,i} - \hat{d}_k\big)^T$$
$$\Sigma_{\tilde{\theta}\tilde{d},k}^- = \textstyle\sum_{i=0}^{p} \alpha_i^{(\mathrm{c})}\big(\mathcal{W}_{k,i}^{\theta,-} - \hat{\theta}_k^-\big)\big(\mathcal{D}_{k,i} - \hat{d}_k\big)^T$$
$$L_k^\theta = \Sigma_{\tilde{\theta}\tilde{d},k}^- \Sigma_{\tilde{d},k}^{-1}$$

State estimate meas. update: $\hat{\theta}_k^+ = \hat{\theta}_k^- + L_k^\theta\big(d_k - \hat{d}_k\big)$

Error covariance meas. update: $\Sigma_{\tilde{\theta},k}^+ = \Sigma_{\tilde{\theta},k}^- - L_k^\theta \Sigma_{\tilde{d},k} (L_k^\theta)^T$