Lecture 17: Parameter Augmented State Estimation

Thus far in the course we've encountered two forms of system identification (linear regression, and stepwise linear regression) as well as several estimation techniques (Luenberg observer and various Kalman filters). In this lecture we show that system identification can be compined with state estimation to estimate parameters of the system in real-time. This is particularly useful when the system parameters might change at runtime.

The basic idea of parameter augemnted state estimation is quite simple — we simply augment the state with a constant parameter which is treated as a new state but has trivial (zero) dynamics (since it is a constant). Then we can apply any state estimation technique. Of course, the system should be observable for the problem to be well posed.

Parameter Augemented State Estimation [Stengel, 1994, Sec. 4.7]

Consider a system model of the form

$$\dot{\boldsymbol{x}} = \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{x} + \boldsymbol{B}(\boldsymbol{\theta})\boldsymbol{u} \tag{1}$$

$$\boldsymbol{y} = \boldsymbol{C}\boldsymbol{x} \tag{2}$$

Suppose the true system parameter is $\boldsymbol{\theta}^*$ and our goal is to estimate $\boldsymbol{\theta}$. Form the stacked state vector

$$\boldsymbol{z} = \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{\theta} \end{bmatrix} \tag{3}$$

the dynamics are

$$\dot{\boldsymbol{z}} = \begin{bmatrix} \dot{\boldsymbol{x}} \\ \dot{\boldsymbol{\theta}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{x} + \boldsymbol{B}(\boldsymbol{\theta})\boldsymbol{u} \\ \boldsymbol{0} \end{bmatrix} \tag{4}$$

Note that the system is now nonlinear and in the form.

$$\dot{\boldsymbol{z}} = \boldsymbol{f}(\boldsymbol{z}, \boldsymbol{u}) \tag{5}$$

We can design a nonlinear estimator for this system the usual way. For example, using an Extended Kalman filter.

Example (Stengel, 4.7-1). Consider a second-order system model of a weathervane shown below

$$\dot{\boldsymbol{x}} = \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{x} + \boldsymbol{L}w \tag{6}$$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\omega_n^2 & -2\zeta\omega_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \omega_n^2 \end{bmatrix} w \tag{7}$$

where x_1 represents the weathervane angle, x_2 represents angular rate, and w is a scalar random white-noise wind input with (continuous-time) variance $q_c = 1000$. The damping ratio ζ is 0.1 and the natural frequency ω_n is 2 rad/s. We assume that the unknown parameter is ω_n^2. The weathervane instrumentation system measures both the angle and rate:

$$\boldsymbol{y} = \boldsymbol{C}\boldsymbol{x} + \boldsymbol{v} \tag{8}$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \tag{9}$$

where $\boldsymbol{v}$ is the measurement noise with continuous-time variance

$$\boldsymbol{R}_c = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix} \tag{10}$$

This system presents a difficult estimation problem because the input w is pure white noise known only by its mean and covariance (as opposed to a control input that is known to the estimator). We begin by augmenting our state vector as

$$\boldsymbol{z} = \begin{bmatrix} x_1 \\ x_2 \\ a \end{bmatrix} \tag{11}$$

where we have defined $a = -\omega_n^2$ and $b = -2\zeta\sqrt{-a}$ for convenience. Since we assume this parameter is a constant it has dynamics $\dot{a} = 0$. The dynamics of $\boldsymbol{z}$ are thus

$$\dot{\boldsymbol{z}} = \boldsymbol{f}(\boldsymbol{z}, w) \tag{12}$$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{a} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ a & b & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ a \end{bmatrix} + \begin{bmatrix} 0 \\ -a \\ 0 \end{bmatrix} w \tag{13}$$

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \end{bmatrix} = \begin{bmatrix} z_2 \\ z_3 z_1 + b z_2 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -z_3 w \\ 0 \end{bmatrix} \tag{14}$$

Note that the system above is nonlinear (even though we wrote it as a matrix) since a is a state variable that multiplies x_1 in the expression for $\dot{x}_2$. The outputs for the new system remain the same

$$\boldsymbol{y} = \boldsymbol{C}_A \boldsymbol{z} + \boldsymbol{v} \tag{15}$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ a \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \tag{16}$$

Now we will apply an extended Kalman filter. Begin by computing the necessary Jacobians

$$\boldsymbol{J_x f} = \begin{bmatrix} \frac{\partial f_1}{\partial z_1} & \frac{\partial f_1}{\partial z_2} & \frac{\partial f_1}{\partial z_3} \\ \frac{\partial f_2}{\partial z_1} & \frac{\partial f_2}{\partial z_2} & \frac{\partial f_2}{\partial z_3} \\ \frac{\partial f_3}{\partial z_1} & \frac{\partial f_3}{\partial z_2} & \frac{\partial f_3}{\partial z_3} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ z_3 & b & z_1 \\ 0 & 0 & 0 \end{bmatrix} \tag{17}$$

and

$$\boldsymbol{J}_w \boldsymbol{f} = \begin{bmatrix} \frac{\partial f_1}{\partial w} \\ \frac{\partial f_2}{\partial w} \\ \frac{\partial f_3}{\partial w} \end{bmatrix} = \begin{bmatrix} 0 \\ -z_2 \\ 0 \end{bmatrix} \tag{18}$$

We linearize each time around the current estimation point thus

$$\dot{\boldsymbol{z}} = \begin{bmatrix} 0 & 1 & 0 \\ \hat{z}_3 & b & \hat{z}_1 \\ 0 & 0 & 0 \end{bmatrix} \boldsymbol{z} + \begin{bmatrix} 0 \\ -\hat{z}_2 \\ 0 \end{bmatrix} w \tag{19}$$

To apply the discrete-time EKF we must discretize the system apply the simplified Euler

$$
\begin{aligned}
\boldsymbol{F}_k &= (\boldsymbol{I} + \boldsymbol{A}\Delta t) && (20)\\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 \\ (\hat{z}_3)_k & (\hat{b})_k & (\hat{z}_1)_k \\ 0 & 0 & 0 \end{bmatrix} \Delta t && (21)\\
&= \begin{bmatrix} 1 & \Delta t & 0 \\ (\hat{z}_3)_k \Delta t & 1 + (\hat{b})_k \Delta t & (\hat{z}_1)_k \Delta t \\ 0 & 0 & 1 \end{bmatrix} && (22)
\end{aligned}
$$

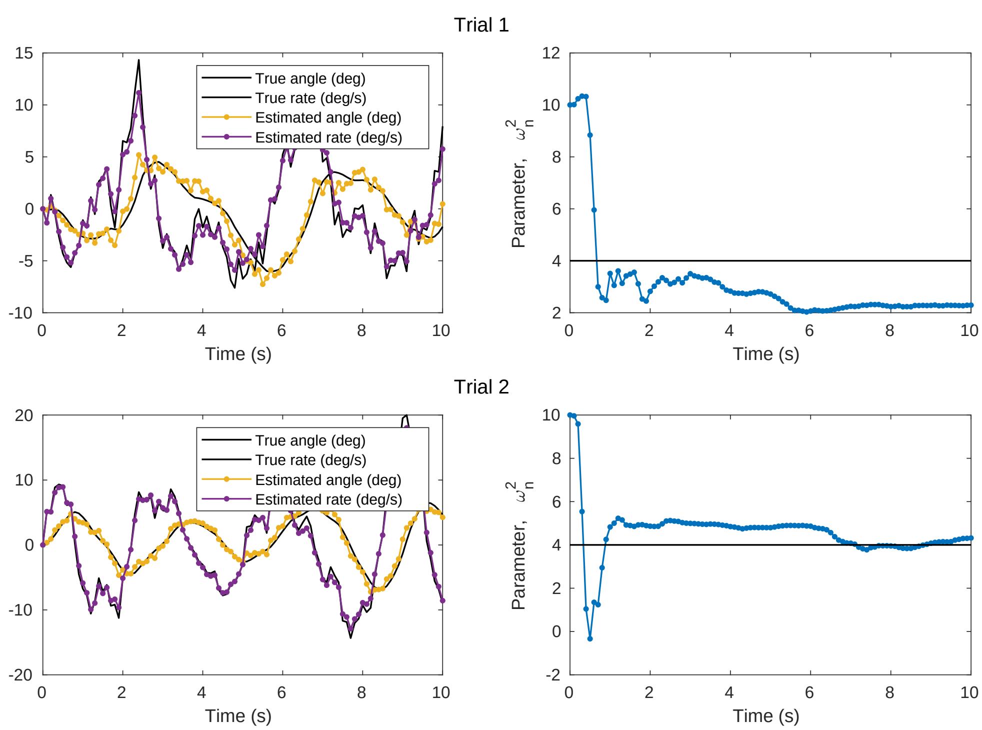
and

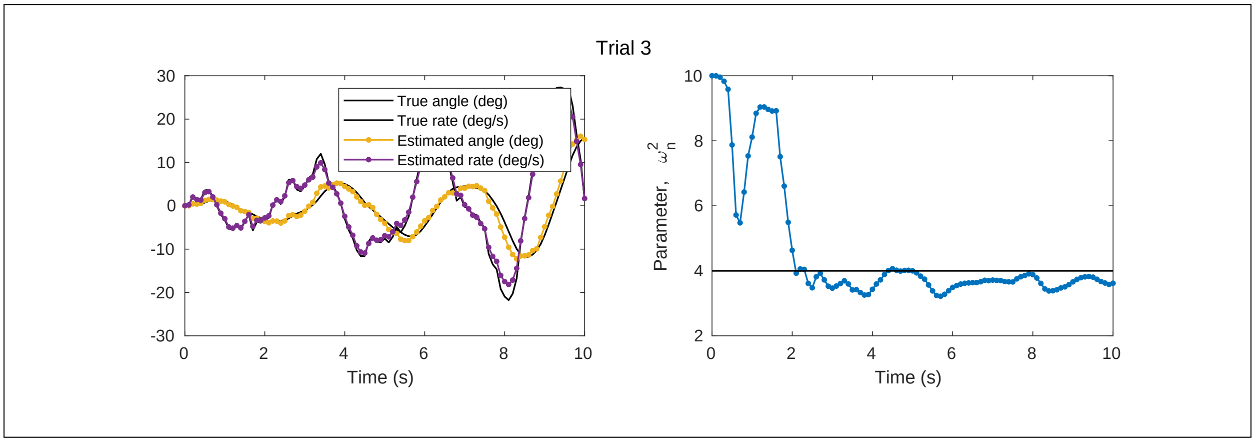
$$
\begin{aligned}
\boldsymbol{L}_k &= \boldsymbol{L}\Delta t && (23)\\
&= \begin{bmatrix} 0 \\ -(\hat{z}_2)_k \\ 0 \end{bmatrix} \Delta t && (24)
\end{aligned}
$$

So we have a system

$$
\begin{aligned}
\boldsymbol{z}_k &= \boldsymbol{F}_{k-1}\boldsymbol{z}_{k-1} + \boldsymbol{L}_{k-1} w_{k-1} && (25)\\
\boldsymbol{y}_k &= \boldsymbol{H}_k \boldsymbol{z}_k + \boldsymbol{r}_k && (26)
\end{aligned}
$$

which is in a form ready for the discrete-time extended Kalman filter.





References

[Stengel, 1994] Stengel, R. F. (1994). *Optimal Control and Estimation*. Courier Corporation.