ECE531 Lecture 10b: Dynamic Parameter Estimation: System Model

D. Richard Brown III

Worcester Polytechnic Institute

02-Apr-2009
Introduction

- So far, we have only considered estimation problems with fixed parameters, e.g.
  - $Y_k \overset{i.i.d.}{\sim} \mathcal{N}(\theta, \sigma^2)$.
  - $Y_k = a \cos(\omega k + \phi) + W_k$ for $\theta = [a, \phi, \omega]^\top$.
  - $Y_k \overset{i.i.d.}{\sim} \theta e^{-\theta y_k}$.

- Many problems require us to estimate dynamic or time-varying parameters, e.g.
  - Radar (position and velocity of target changing over time)
  - Communications (amplitude and phase of signal changing over time)
  - Stock market (price of shares next week changing over time)

- Step 1: We need to develop a general model for
  - How time-varying parameters evolve over time and
  - How observations are generated according from the parameter state.

- Step 2: We need to develop good techniques for estimating dynamic parameters. We can expect these techniques to be based on our good static parameter estimators, i.e. MVU and/or ML, and also use some knowledge of the time-varying model.
Discrete Time Model for Dynamic Parameters

The time index is denoted as $n = 0, 1, \ldots$. All vectors are considered to be random unless otherwise specified.

$U[n] \in \mathbb{R}^s$

$X[n] \in \mathbb{R}^m$

$Z[n] \in \mathbb{R}^k$

$V[n] \in \mathbb{R}^k$

$Y[n] \in \mathbb{R}^k$

$f[n] : \mathbb{R}^s \times \mathbb{R}^m \rightarrow \mathbb{R}^m$

$h[n] : \mathbb{R}^m \rightarrow \mathbb{R}^k$
Notation and Terminology

- $U[n]$ is the dynamical system “input”, $n = 0, 1, \ldots$.
- $X[n]$ is the “state” of the dynamical system, $n = 0, 1, \ldots$.
- $Z[n]$ is the dynamical system “output”, $n = 0, 1, \ldots$.
- $f[n]$ is a time varying function that updates the state based on the current state and the current input, i.e.
  \[ X[n + 1] = f[n](X[n], U[n]) \]
- $h[n]$ is a time varying function that generates the current output based on the current state, i.e.
  \[ Z[n] = h[n](X[n]) \]
- $V[n]$ is the “measurement noise” $n = 0, 1, \ldots$.
- $Y[n]$ is the “observation” $n = 0, 1, \ldots$.
Example: One-Dimensional Motion

Suppose we have a particle moving on a line with position $x$ and velocity $v$ updated according to

$$
x[n + 1] = x[n] + Tv[n]$$
$$v[n + 1] = v[n] + Ta[n]$$

where $a[n]$ represents the acceleration and $T$ is the time between samples. The implicit assumption here is that $T$ is small enough such that the discrete time model holds.

- The state $X = [x, v]^\top$.
- The input $U = a$.
- The state update equation can be written as

$$X[n + 1] = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} X[n] + \begin{bmatrix} 0 \\ T \end{bmatrix} U[n]$$

- Suppose we observe the position of the particle in noise. Then

$$Y[n] = \begin{bmatrix} 1 & 0 \end{bmatrix} X[n] + V[n]$$
Example: One-Dimensional Motion (White Gaussian Input)
Remarks

- In these types of dynamical systems, $X[k]$ is completely determined by the earlier state $X[\ell]$, $\ell < k$, and the inputs \{${U[\ell], \ldots, U[k-1]}$\}.

- The state at time $\ell$ completely summarizes the system in the sense that you don’t need to know the details of what happened prior to time $\ell$ if you know $X[\ell]$.

- We are going to study problems in which we wish to estimate the dynamic state $X[\ell]$ given a sequence of observations $Y[0], \ldots, Y[k]$. These problems can be categorized into three types:
  
  1. **Filtering**: $\ell = k$ (estimate the current state)
  2. **Prediction**: $\ell > k$ (estimate a future state)
  3. **Smoothing**: $\ell < k$ (estimate a previous state)

- Note that the notation in dynamic parameter estimation problems is different than static parameter estimation problems: To be consistent with your textbook, $X[\ell]$ represents the quantity we wish to estimate.
Restriction 1: Squared Error Cost Assignment

We will only consider the squared error cost assignment, i.e.

\[
\text{MSE} = \mathbb{E} \left\{ \| \hat{X}[\ell] - X[\ell] \|_2^2 \right\}
\]

where \( \hat{X}[\ell] \) is the estimate of the state \( X[\ell] \). We assume that we know the joint distribution of the input \( U[n] \) and the distribution of the initial state \( X[0] \).

Given the observation \( Y[0], \ldots, Y[k] \), what estimator \( \hat{X}[\ell] \) minimizes the MSE?

Hint: Is this Bayesian estimation or non-random parameter estimation?
Memory and Computational Requirements

- We are often interested in producing state estimates in real time, i.e., upon receiving the observation $Y[k]$, we estimate $X[\ell]$; upon receiving observation $Y[k+1]$, we estimate $X[\ell+1]$; and so on.

- In general, computing

$$\hat{X}_{\text{mmse}}[\ell] = \mathbb{E} \{X[\ell] \mid Y[0], \ldots, Y[k]\}$$

would require us to
  - Keep all the past observations in memory
  - Compute the estimate as a function of all the past observations

- In other words, the general dynamic state estimation problem has linearly growing memory and computational burdens.

- Additional restrictions are going to be necessary to make the dynamic state estimation problem computationally feasible.
Restriction 2: Linear Dynamical Model

We are going to restrict our attention to systems with state update equations and output equations of the form

\[
X[n + 1] = F[n]X[n] + G[n]U[n] \quad n = 0, 1, \ldots
\]
\[
Y[n] = H[n]X[n] + V[n] \quad n = 0, 1, \ldots
\]

where, for each \( n \), \( F[n] \in \mathbb{R}^{m \times m} \), \( G[n] \in \mathbb{R}^{m \times s} \), and \( H[n] \in \mathbb{R}^{k \times m} \).

- We’ve already seen that one-dimensional motion fits within this linear model.
- The same is true for two- and three-dimensional motion.
- Many nonlinear systems can approximately fit in this model by linearizing \( f \) and \( h \) around a nominal state (Taylor series expansion).
Linear Dynamical Model


\[ = \text{etc.} \]

Repeating this process leads to an expression for the state at time \( n + 1 \) in terms of the initial state \( X[0] \) and the inputs \( U[0], U[1], \ldots \):

\[ X[n + 1] = \left\{ \prod_{k=0}^{n} F[n - k] \right\} X[0] + \sum_{j=0}^{n} \left\{ \prod_{k=0}^{n-j-1} F[n - k] \right\} G[j]U[j] \]

where

\[ \prod_{k=0}^{t} F[n - k] := F[n]F[n - 1]F[n - 2] \cdots F[n - t] \]

is called the state transition matrix from time \( n - t \) to time \( n + 1 \).
Linear Dynamical Model: Time Invariant Case

When

\[ F[n] \equiv F \]
\[ G[n] \equiv G \]

then the discrete time dynamical system is time invariant and the state at time \( n + 1 \) is simply

\[ X[n + 1] = F^{n+1}X[0] + \sum_{j=0}^{n} F^{n-j}G[j]U[j] \]
Restriction 3: Gaussian Input and Measurement Noise

We will revisit this restriction later, but for now we will only consider systems in which the input sequence $U[0], U[1], \ldots$ and the measurement noise sequence $V[0], V[1], \ldots$ are independent sequences of independent zero mean Gaussian random vectors, i.e.

$$
\begin{align*}
\mathbb{E}\{U[k]\} &= 0 \\
\mathbb{E}\{U[k]U^\top[j]\} &= \begin{cases} Q[k] & k = j \\ 0 & \text{otherwise} \end{cases} \\
\mathbb{E}\{V[k]\} &= 0 \\
\mathbb{E}\{V[k]V^\top[j]\} &= \begin{cases} R[k] & k = j \\ 0 & \text{otherwise} \end{cases} \\
\mathbb{E}\{U[k]V^\top[j]\} &= 0 \quad \text{for all } j \text{ and } k
\end{align*}
$$

We also assume that the initial state $X[0] \sim \mathcal{N}(m[0], \Sigma[0])$ is a Gaussian random vector independent of $U[0], U[1], \ldots$ and $V[0], V[1], \ldots$. 
As we will see in the development of the Discrete-Time Kalman-Bucy filter, restrictions 2 and 3 will allow us to write the MMSE estimator

\[ \hat{X}_{\text{mmse}}[\ell] = E\{X[\ell] | Y[0], \ldots, Y[k]\} \]

as

\[ \hat{X}_{\text{mmse}}[\ell] = E\{X[\ell] | \hat{X}[\ell - 1], Y[k]\} \]

without any loss of optimality.

In other words, in our restricted linear/Gaussian model, the optimal state estimator only depends on the previous state estimate and the current observation.

Unlike the general dynamic state estimation problem, the memory requirements and computational burden do not grow in time.

We will also see that the discrete-time Kalman-Bucy filter possesses certain types of optimality even when restrictions 2 and 3 do not hold.