ECE531 Screencast 7.1: Introduction to Dynamic Parameter Estimation

D. Richard Brown III

Worcester Polytechnic Institute

Introduction

So far, we have only considered estimation problems with static/fixed parameters, e.g.

- $Y_k \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\theta, \sigma^2)$.
- $Y_k = a\cos(\omega k + \phi) + W_k$ for $\theta = [a, \phi, \omega]^\top$.
- $Y_k \overset{\text{i.i.d.}}{\sim} \theta e^{-\theta y_k}$.

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)

Key ideas:

- We don't just want to perform "one-shot" estimates.
- We don't want the complexity to "blow up" as we receive more observations.

Preliminary Notation and Terminology

For **fixed** parameters, we have typically used the notation Θ or θ.

For **dynamic** parameters, the convention is to use the notation $X[\ell] \in \mathbb{R}^m$ to refer to the parameter at time ℓ. It is also common to call $X[\ell]$ the "state" at time ℓ. X is capitalized because it is assumed to be random.

The observation at time ℓ is denoted as $Y[n] \in \mathbb{R}^k$. Note that, in general, observations can be vectors. We will sometimes use the notation

$$\mathcal{Y}_a^b := \begin{bmatrix} Y[a] \\ \vdots \\ Y[b] \end{bmatrix} \in \mathbb{R}^{k(b-a+1)}$$

to denote the "super vector" of all observations from $Y[a]$ to $Y[b]$.

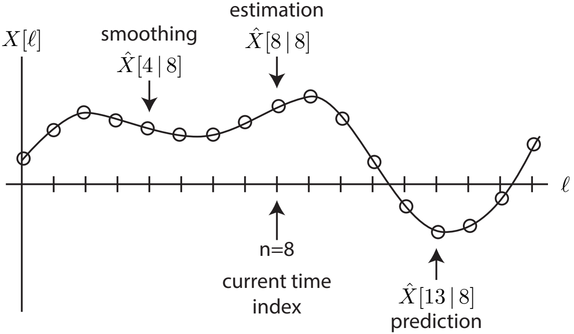
Since we are in the context of Bayesian estimation, the MMSE estimate of the state at time ℓ given observations $Y[0], \dots, Y[n]$ is denoted as

$$\hat{X}[\ell \,|\, n] = \mathrm{E}\left\{X[\ell] \,|\, \mathcal{Y}_0^n\right\}$$

Prediction, Estimation, and Smoothing

We are going to study problems in which we wish to estimate the dynamic state $X[\ell]$ given a sequence of observations $Y[0], \dots, Y[n]$. These problems can be categorized into three types (assume $m > 0$):

1. **Prediction**:$\hat{X}[\ell \,|\, \ell - m]$ (estimate a future state)
2. **Filtering/Estimation**: $\hat{X}[\ell \,|\, \ell]$ (estimate the current state)
3. **Smoothing**: $\hat{X}[\ell \,|\, \ell + m]$ (estimate a previous state)



The Brute Force Approach: Batch MMSE Estimation

Assume a Bayesian linear Gaussian model with $X[\ell]$ and $\mathcal{Y}_0^\ell := \left[Y^\top[0], \dots, Y^\top[\ell]\right]^\top$ jointly Gaussian distributed. We wish to estimate the state $X[\ell]$ from these observations. If we just did this as a batch operation, what is the MMSE estimate of $X[\ell]$?

We know the answer to this is to just compute the conditional mean:

$$\begin{aligned}\hat{X}[\ell \,|\, \ell] &= \mathrm{E}\left\{X[\ell] \,|\, \mathcal{Y}_0^\ell\right\} \\ &= \mathrm{E}\{X[\ell]\} + \mathrm{cov}\{X[\ell], \mathcal{Y}_0^\ell\} \left[\mathrm{cov}\{\mathcal{Y}_0^\ell, \mathcal{Y}_0^\ell\}\right]^{-1} \left(\mathcal{Y}_0^\ell - \mathrm{E}\{\mathcal{Y}_0^\ell\}\right)\end{aligned}$$

Recall that each $Y[n] \in \mathbb{R}^k$.

- What are the dimensions of the matrix inverse?
- What happens as we get more observations?

Kalman Filter: The Main Idea

The Kalman Filter is a computationally efficient way to calculate MMSE estimates of dynamic parameters governed by certain stochastic dynamics.

We've seen something similar already: sequential LMMSE estimation. This was for fixed parameters, however.

Our approach:

- Step 1: We need to develop a general **dynamic model** for
  - How time-varying parameters (the states) evolve over time and
  - How observations are generated from the states.
- Step 2: We need to develop good techniques for estimating dynamic parameters. These techniques should leverage some knowledge of the dynamic model.