Technion, Faculty of Electrical Engineering Fall 2008/9

Estimation and Identification in Dynamic Systems (048825)

Prof. Nahum Shimkin

Lecture Notes

1 Introduction: Basic Problems of Interest

Our main interest in this course will be in the following problem:

- State estimation in dynamic systems, for which the state cannot be fully observed.

Some related problems that we shall also consider:

- Parameter estimation in dynamic systems (system identification).
- Joint state and parameter estimation.

Our emphasis will be on algorithms which are optimal in a statistical (stochastic) sense.

The basic system models that we will deal with are:

- Continuous-state systems: We will develop the celebrated *Kalman Filter* for state estimation in linear state-space models, and its extensions.
- Discrete-state models: the so-called Hidden Markov Models.

We will also consider various extensions of these basic models and problems.

We next give a brief outline of the basic problems and illustrative applications.

1.1 State Estimation in Linear Systems

Consider a discrete-time linear state-space model of the form:

$$\begin{aligned} x_{k+1} &= F x_k + G u_k + v_k \,, \\ z_k &= H x_k + w_k \,. \end{aligned}$$

Here: $x \in \mathbb{R}^n$ is the state vector, which is unknown to us

$z \in \mathbb{R}^m$ is the measurement vector

u is a known input signal

v and w are unobserved noise sequences

F, G, H are the system matrices

The basic state-estimation problem: The systems matrices are given, and so are some properties of the noise sequences. Our goal is to find an estimate $\hat{x}_{k+1}$ for the state vector x_{k+1}, given the measurements $\{z_k,\, z_{k-1}\, \ldots\}$.

The proposed solution is of the following (state-observer) form:

$$\hat{x}_{k+1} = F\hat{x}_k + Gu_k + K_k(z_k - H\hat{x}_k)\,.$$

This is a recursive filter, which can be operated in an on-line mode (i.e., the estimate is updated each time a new measurement is obtained).

K_k is a gain matrix, to be "properly" chosen.

The Kalman Filter is obtained by an optimal choice of these gains, under appropriate statistical model assumptions and error criteria.

Examples: We next sketch a few simplified examples for problems that can be cast in this form.

Example 1: Position Estimation

Consider an object moving in 1-dimensional space, with position $p(t)$.
We are given *noisy* (inaccurate) measurements of this position at some discrete times:

$$z(t) = p(t) + n_z(t), \quad t = t_0, t_1, \ldots \,.$$

Required: to estimate the position $\hat{p}(t)$.

To formulate this problem in state space form, several options are available:

1. Random acceleration model (2nd order model):

$$\begin{aligned} \frac{d}{dt} p(t) &= v(t) \\ \frac{d}{dt} v(t) &= a(t) \equiv n_v(t) \end{aligned}$$

   where $n_v(t)$ is "white", 0-mean noise signal with known statistics. This noise reflects the expected object "maneuverability".

   We have arrived at the following state model:

$$\frac{d}{dt} \begin{bmatrix} p(t) \\ v(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} p(t) \\ v(t) \end{bmatrix} + \begin{bmatrix} 0 \\ n_v(t) \end{bmatrix}$$

   where $x(t) = (p(t), v(t))'$ is the state vector.
   This state equation may be discretized to obtain a discrete-time state model over the measurement times, of the form:

$$x(t_{k+1}) = A(k)x(t_k) + n(k) \,.$$

The measurement equation is

$$z(t) = [1, 0] \begin{bmatrix} p(t) \\ v(t) \end{bmatrix} \equiv [1, 0]\, x(t) \quad t = t_k$$

2. Random acceleration-change model (3nd order model):

   When velocity cannot change abruptly, the following model is more suitable:

$$\begin{aligned} \dot{p}(t) &= v(t) \\ \dot{v}(t) &= a(t) \\ \dot{a} &= n_a(t) \end{aligned}$$

   Furthermore, when acceleration cannot change abruptly we can add a simple low path filter:

$$\dot{a} = -\beta a + n_a(t)$$

   with β a properly chosen costant. With state $x(t) = [p(t), v(t), a(t)]'$, we obtain the following state model:

$$\dot{x}(t) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -\beta \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 0 \\ n_a(t) \end{bmatrix}$$

   and measurement equation:

$$z = [1, 0, 0]x + n_z \,.$$

3. Additional measurements:

   We may, for example, have also direct velocity measurements of the moving object. Then the measurement equations are:

$$z_1(t_k) = p(t_k) + n_1(t_k) z_2(t_k) = v(t_k) + n_2(t_k)$$

In matrix form (for $x = [p, v, a]'$):

$$z(t_k) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} x(t) + n_z(t)$$

with $n_z = (n_1, n_2)'$.

Position estimation (with Kalman filtering) has many variants and applications, including:

- *Navigation:* Using odometry, inertial sensors, GPS, vision,...
- *Tracking:* Using radar, vision, cellular phones....

Example 2: Signal Detection

Consider a discrete-time signal $s(k)$ which is transmitted through a noisy channel with ISI; the received signal is

$$z(k) = \sum_{i=0}^{N-1} h_i s(k-i) + n_z(k)$$

with n_z a white noise sequence, say $n_z(k) \sim N(0, \sigma_z)$.

It is required to recover the transmitted signal $s(k)$ from the measurements $z(k'),\ k' \leq k$. This is a classical signal filtering problem.

To use statistical methods, we use a statistical model for the transmitted signal: e.g., s is a white noise sequence with $s(k) \sim N(0, \sigma_s)$. For $N = 2$, the state variables and equations are:

$$\begin{aligned}
x_1(k) &= s(k) & &\Rightarrow & x_1(k+1) &= n_s(k+1) \\
x_2(k) &= s(k-1) & &\Rightarrow & x_2(k+1) &= x_1(k) \\
x_3(k) &= s(k-2) & &\Rightarrow & x_3(k+1) &= x_2(k)
\end{aligned}$$

and in matrix form:

$$x(k+1) = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} x(k) + \begin{bmatrix} n_s(k+1) \\ 0 \\ 0 \end{bmatrix}$$

$$z(k) = [h_0, h_1, h_2] x(k) + n_z(k)\,.$$

Example 3: Higher harmonics detection

We are given a sinusoidal signal, with basic frequency $f_1 = 50$. We need to detect the higher-order (say, 3nd order) harmonic content of the signal. (A 2nd order harmonic is usually absent in power systems.)

With $\omega_1 = 2\pi f_1$, we write this signal as

$$s(t) = A\sin(\omega_1 t + \phi_1) + B\sin(3\omega_1 t + \phi_3) \quad (+\text{other terms})\,.$$

It is assumed that the amplitude and phases may (slowly) vary with time. The signal is measured every $T = 0.1$ sec. We wish to track $B(t)$ and $\phi_1(t)$ over time.

The problem may obviously be approached in the frequency domain, using standard filtering methods. We give here the alternative Kalman-filter formulation.

We start by writing the harmonic signal model as:

$$s(t) = A_1(t)\cos(\omega_1 t) + A_2(t)\sin(\omega_1 t) + B_1(t)\cos(3\omega_1 t) + B_2(t)\sin(3\omega_1 t)$$

The state is taken as the four amplitudes: $x = [A_1, A_2, B_1, B_2]'$. Focusing on the measurements instances $t_k = kT$, we have the following model:

$$\begin{aligned} A_1(t_{k+1}) &= A_1(t_k) + n_{A1}(k) \\ &\vdots \\ B_2(t_{k+1}) &= B_2(t_k) + n_{B2}(k) \end{aligned}$$

The noise variances are taken as small quantities related to the allowed rate of change of the amplitudes. The state equation is then

$$x(t_{k+1}) = Ix(t_k) + n(k)$$

where I is the unit matrix.

The measurement model is:

$$z(t_k) = s(t_k) + n_z(k)$$

with n_z the measurement error. It may be taken as a white Gaussian sequence, with $n_z(k) \sim N(0, \sigma_z)$. This gives

$$\begin{aligned} z(t_k) &= [\cos(\omega_1 t_k), \sin(\omega_1 t_k), \cos(3\omega_1 t_k), \sin(3\omega_1 t_k)]x(t_k) + n_z(k) \\ &\doteq H(k)x(t_k) + n_z(k)\,. \end{aligned}$$

1.2 Hidden Markov Models (HMMs)

HMMs are state models with discrete state, which cannot be directly observed, and with discrete or continuous measurements.

Let $x_k \in \{1, 2, \ldots, N\}$ be a Markov chain specified by the transition law

$$p(x_{k+1} = j \,|x_k = i\} = p_{ij}$$

and initial distribution $p(x_0)$. Let the z_k be the measurement, say discrete, related to x_k through,

$$p(z_k = z \,|\, x_k = i) = q(z|i)\,.$$

The basic problems here are:

1. Given the model parameters, and measurements $(z_n, z_{n-1}, \ldots, z_1\}$, estimate the state sequence $\{x_n, x_{n-1}, \ldots, x_1\}$.

We shall develop the Maximum Likelihood (ML) estimator for that problem, which is efficiently implemented using the so-called *Viterbi algorithm*.

2. Given the measurements $(z_n, z_{n-1}, \ldots, z_1\}$, estimate the model parameters, namely: find the model that best describes the data.

The standard solution for this problem is joint state and parameter estimation, using the *EM* (Expectation Maximization) algorithm.

Applications: An important application for HMMs is speech processing, in particular speech recognition. The HMM may be used at different levels of speech modeling, such as:

(1) Word level: The state x is a complete word; the measurement is the recorded

sound of the word; and the dynamics $p(j|i)$ represents the likelihood of word j appearing after word i.

(2) Phonetic level: Using a phonetic alphabet to model the inner structure of each word.

In that context, "estimating the state sequence" means identifying the spoken word sequence; and "estimating the model parameters" relates to the training phase when the model is tuned to a specific speaker.