# Mean-Squared-Error Estimation

Dec 6, 2022

Let $X$, $Y$ be two random variables. Suppose we want to estimate $Y$ given that $X=x$. What is the "best" estimate of $Y$ given $X=x$?

One criterion commonly used is the minimum-mean-squared-error (MMSE) criterion: Find a function $c(x)$ such that $\varepsilon = E[(Y-c(X))^2]$

is minimized.

The result is $c(x) = E[Y|X=x]$

Proof. $$\varepsilon = \iint_{\mathbb{R}^2} (y-c(x))^2 f_{XY}(x,y)\,dx\,dy$$

$$= \int_{\mathbb{R}} f_X(x) \underbrace{\left[\int_{\mathbb{R}} (y-c(x))^2 f_{Y|X}(y|x)\,dy\right]}_{h(x)} dx$$

Since $f_X(x)$, $h(x)$ are $\geq 0 \ \forall x$, can minimize $\varepsilon$ by minimizing $h(x)$ $\forall x \in \mathbb{R}$. To minimize $h(x)$:

$$\frac{\partial}{\partial c(x)} h(x) = -2\int_{-\infty}^{\infty}(y-c(x))f_{Y|X}(y|x)\,dy$$

$$= 0$$

So $\int_{-\infty}^{\infty}(y-c(x))f_{Y|X}(y|x)\,dy = 0$, or

$$E[Y|X=x] - c(x) = 0$$

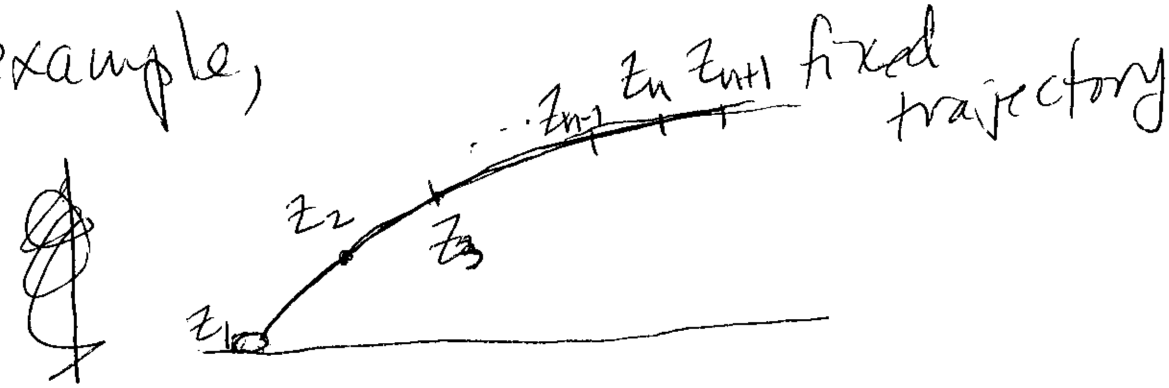
So $c(x) = E[Y|X=x]$

Sometimes $c(x)$ is constrained to be a linear function. This is the case for the Kalman filter.

# The Kalman Filter

Estimates the state of a dynamic system from a sequence of noisy measurements of the state.

For example,




Want to estimate the distance $z_n$ along the trajectory that object has traveled by time $n$ (actually time $t_n$). Two-stage model:

The state model is

$$z_n = a_{n-1} z_{n-1} + w_{n-1}, \quad n = 1, 2, \ldots$$

for some initial state $z_0$

$Z_n \sim$ state variable

(In practice $Z_n$ is a vector with, e.g., 3 position rvs, 3 velocity rvs, etc.)

$a_n \sim$ constant sequence (not random) that quantifies the state transitions)

$W_{n-1} \sim$ state model error, modeled as a zero-mean uncorrelated sequence of rvs, so

$$E[W_n] = 0 \quad \forall n$$

$$E[W_j W_k] = 0 \quad \forall j \neq k$$

So $E[W_n^2] \sim$ model error variance

Want to estimate cov predict $Z_n$ from $Z_1, \dots, Z_{n-1}$.

The observation model is

$$X_n = Z_n + N_n, \quad \text{where}$$

$X_n$ is our measurement of $Z_n$

$N_n$ is measurement error (noise), modeled as zero-mean, uncorrelated

$$E[N_n] = 0 \quad \forall n$$

$$E[N_j N_k] = 0 \quad \forall j \neq k$$

$E[N_n^2]$ noise variance or power.

Also assume

$$E[N_j W_k] = 0 \quad \forall j, k$$

We actually estimate $Z_n$ from $X_1, \dots, X_{n-1}$

The Kalman filter finds the linear MMSE estimate of $Z_n$ from measured values of $X_1, \ldots, X_{n-1}$. Let

$$Y_n = \sum_{j=1}^{n} h_j^{(n-1)} X_{n-j} = h_1^{(n-1)} X_{n-1} + \cdots + h_n^{(n-1)} X_0$$

where $X_0 = Z_0$

Then we want to find $h_j^{(n-1)}$ for $j = 1, \ldots, n$ to minimize

$$E\left[(\underbrace{Y_n}_{\text{estimate of } Z_n} - \underbrace{Z_n}_{\text{state at time } t_n})^2\right]$$

Can show that the optimal filter $h_j^{(n-1)}$ satisfies

$$R_{ZX}(n,l) = \sum_{j=1}^{n} h_j^{(n-1)} R_{XX}(n-j,l)$$