EECS 126: Probability & Random Processes
Fall 2021

Tracking

Shyam Parekh

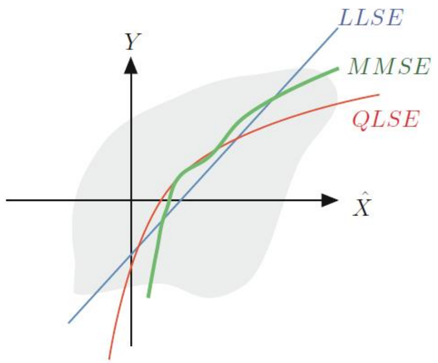
Topics Covered in Lectures on Tracking

- Tracking Problem (Section 9.1 & 9.2)
- LLSE (Section 9.3)
- Linear Regression (Section 9.4)
- MMSE (Section 9.6)
- Updating LLSE (Section 10.1)
- Kalman Filter (Kalman Filter Notes*)

* Lecture will be based on the Kalman Filter Notes posted on the course website.
Sections 9.8 and 10.2 can be used as references.

Problem Statement

- Let (X, Y) be a pair of continuous RVs related to a system, and we want to estimate X based on the observed value Y by
  - $\hat{X} = g(Y)$ such that expected cost $C(g) := E(c(X, g(Y)))$ is minimized for a given cost function c.
  - Squared Error cost: $c(X, \hat{X}) = |X - \hat{X}|^2$ for $X \in \mathcal{R}$ or $\| X - \hat{X} \|^2$ for $X \in \mathcal{R}^d$.
  - With squared error cost and minimizing $C(g)$ over arbitrary function g, $\hat{X}$ is the Minimum Mean Squared Error (MMSE) Estimate of X given Y.
  - With squared error cost and minimizing $C(g)$ where g limited to linear functions of Y (i.e., $\hat{X} = a + bY$ for some a and b), $\hat{X}$ is the Linear Least Squares Error (LLSE) Estimate of X given Y.



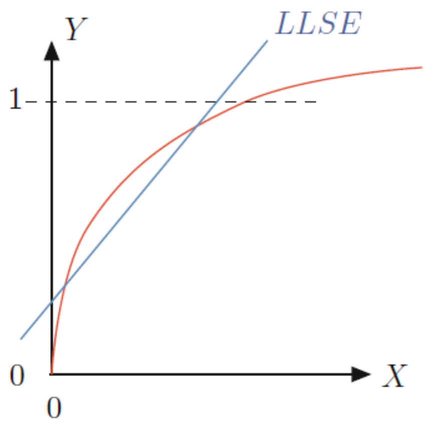
- Different formulations:
  - Known joint distribution of (X, Y).
  - Offline: Observe set of samples $(X_i, Y_i), i = 1, 2, \dots, K$.
  - Online: Observe successive samples (X_n, Y_n).
- Examples:
  - Estimate location based on GPS signals.
  - Estimate speed based on radar signals.
  - Estimate state of vehicle based on sensor signals.

Linear Least Squares Error (LLSE) Estimate

- We are looking for $\hat{X} = a + bY$ that minimizes $E(|X - a - bY|^2)$ over $\forall\ a, b \in \mathcal{R}$ assuming we know joint distribution of (X, Y).
  - $L[X|Y]$ denotes this estimate.
- Theorem: Assuming $var(Y) \neq 0$,

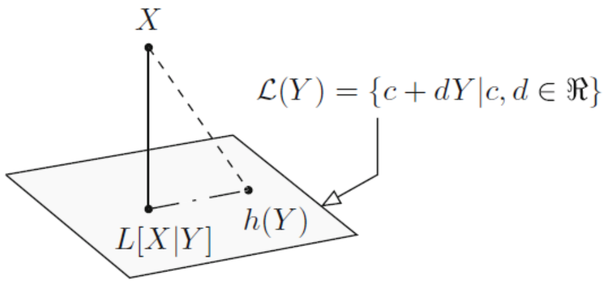
$$L[X|Y] = E(X) + \frac{cov(X,Y)}{var(Y)}(Y - E(Y)).$$

- Example: $Y = \alpha X + Z$, where X and Z are zero-mean & independent RVs.
- Example: $X = \alpha Y + \beta Y^2$, where $Y \equiv_D U[0,1]$.



Projection

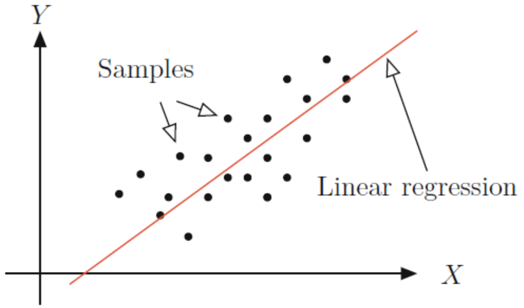
- $L[X|Y]$ is the projection of X onto the subspace $\mathcal{L}(Y)$ of linear functions of Y.
  - Define $E(VW)$ as the inner product of two RVs V and W.
  - V and W are orthogonal if $E(VW) = 0$.
  - Projection Property: $X - L[X|Y]$ is orthogonal to every linear function of Y.



  - $L[X|Y]$ is the closest point to X in $\mathcal{L}(Y)$.

Regression

- In stead of knowing the joint distribution, suppose we observe IID samples $(X_i, Y_i), i = 1, \dots, K$.
- Our goal is to choose a and b that minimizes $\frac{1}{K}\sum_{i=1}^{K} |X_i - a - bY_i|^2$.



- Theorem: $a + bY = E_K(X) + \frac{cov_K(X,Y)}{var_K(Y)}(Y - E_K(Y))$, where

$E_K(X) = \frac{1}{K}\sum_{i=1}^{K} X_i$

$E_K(Y) = \frac{1}{K}\sum_{i=1}^{K} Y_i$

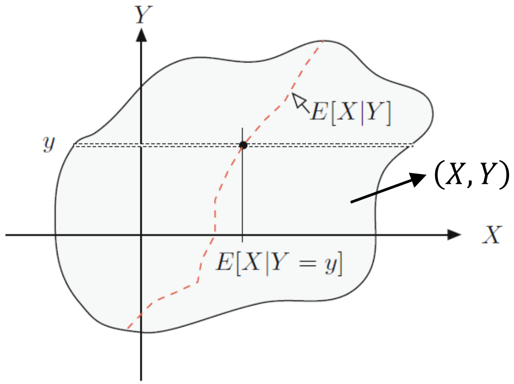
$cov_K(X,Y) = \frac{1}{K}\sum_{i=1}^{K} X_i Y_i - E_K(X)E_K(Y)$

$var_K(Y) = \frac{1}{K}\sum_{i=1}^{K} Y_i^2 - (E_K(Y))^2$

- Theorem: As the number of samples increases the linear regression approaches LLSE estimate.

Minimum Mean Squared Error (MMSE) Estimate

- For now, assume we know the joint distribution of (X, Y).
- Problem: Find function g such that $g(Y)$ minimizes $E(|X - g(Y)|^2)$.
  - The best solution $g(Y)$ is called MMSE of X given Y.
- Theorem: MMSE of X given Y is given by $g(Y) = E[X|Y]$.
- Recall $E[X|Y]$ is an RV, and $E[X|Y = y] = \int_{-\infty}^{\infty} x f_{X|Y}\,[\mathrm{x|y}]\mathrm{dx}$, where $f_{X|Y}[\mathrm{x|y}] = \frac{f_{X,Y}(x,y)}{f_Y(y)}$.



- Lemma (Orthogonality Property of MMSE):
  i. For any function $\phi(.)$, $E\big((X - E[X|Y])\phi(Y)\big) = 0$.
  ii. Also, if $g(Y)$ is such that $E\Big(\big(X - g(Y)\big)\phi(Y)\Big) = 0\ \forall\ \phi(.)$, $g(Y) = E[X|Y]$.
- Fact: $\phi(Y)$ is MMSE of X given Y if and only if $E[X - \phi(Y)|Y] = 0$.

Properties of Conditional Expectation

- Properties:
  a. Linearity: $E[a_1X_1 + a_2X_2|Y] = a_1E[X_1|Y] + a_2E[X_2|Y]$.
  b. Factoring: $E[h(Y)X|Y] = h(Y)E[X|Y]$.
  c. Independence: If X and Y are independent, $E[X|Y] = E(X)$.
  d. Smoothing: $E(E[X|Y]) = E(X)$.
  e. Tower: $E[E[X|Y,Z]|Y] = E[X|Y]$.
- Example: Let X, Y be IID $U[0,1]$. Find $E[(X + 2Y)^2|\mathrm{Y}]$.
- Example: Let X, Y, Z be IID. Find $E[X|X + Y + Z]$.

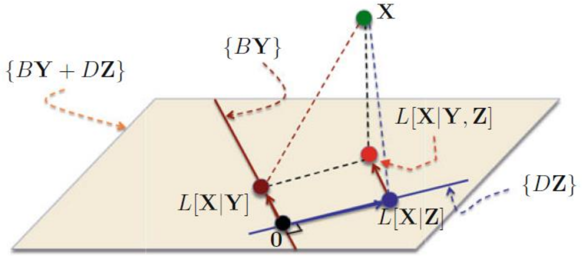
MMSE for Jointly Gaussian RVs

- Theorem: Let X, Y be JG RVs. Then,

$$\mathrm{E[X|Y]} = L[X|Y] = E(X) + \frac{cov(X,Y)}{var(Y)}(Y - E(Y)).$$

Important Results for LLSE

- Theorem (LLSE Orthogonal Update): Assume that X, Y, Z are zero-mean RVs, and that Y, Z are orthogonal. Then, $L[X|Y,Z] = L[X|Y] + L[X|Z]$.
- Theorem (LLSE General Update): Assume that X, Y, Z are zero-mean RVs. Then, $L[X|Y,Z] = L[X|Y] + L[X|Z - L[Z|Y]]$.
- LLSE Properties:
  - *a.* $L[a_1X_1 + a_2X_2|Y] = a_1L[X_1|Y] + a_2L[X_2|Y]$.
  - *b.* $L[L[X|Y,Z]|Y] = L[X|Y]$.
  - c. If X, Y are uncorrelated, $L[X|Y] = E(X)$.



Scalar Kalman Filter (KF)

- Consider a system with state x_n and output y_n:

  $x_n = ax_{n-1} + v_n$, $|a| < 1$, and

  $y_n = cx_n + w_n, \mathrm{n} \geq 1$, where

  $x_0, \{v_n\}_{n=1}^{\infty}, \{w_n\}_{n=1}^{\infty}$ are orthogonal zero-mean RVs,

  with $E(v_n^2) = \sigma_v^2, E(w_n^2) = \sigma_w^2$.

  Find $\hat{x}_{\mathrm{n|n}} := L[x_n | y_1, \ldots, y_n], \mathrm{n} \geq 1$, recursively.
- Examples: Particle Positon, Chemical Reaction, Econometry, …
- Solution (assuming $c = 1$):

  Compute $\hat{x}_{1|1}$ and iterate for $n \geq 2$ using

  $\hat{x}_{\mathrm{n|n}} = a\hat{x}_{\mathrm{n\text{-}1|n\text{-}1}} + k_n(y_n - a\hat{x}_{\mathrm{n\text{-}1|n\text{-}1}})$, where k_n's can be computed offline.

  Compute $\sigma_{1|1}^2$ and iterate for $n \geq 2$ using

  $E((x_n - \hat{x}_{\mathrm{n|n\text{-}1}})^2) =: \sigma_{n|n-1}^2 = a^2\sigma_{n-1|n-1}^2 + \sigma_v^2$

  $k_n = \frac{\sigma_{n|n-1}^2}{\sigma_{n|n-1}^2 + \sigma_w^2}$

  $E((x_n - \hat{x}_{\mathrm{n|n}})^2) =: \sigma_{n|n}^2 = \sigma_{n|n-1}^2(1 - k_n)$
- Example: Obtain $\hat{x}_{1|1}$ and $\sigma_{1|1}^2$ for initializing the KF recursions when $x_0 = 0$.

Vector Kalman Filter (KF)

- Consider a system with state X_n and output Y_n:

  $X_n = AX_{n-1} + V_n$, and

  $Y_n = CX_n + W_n, \mathrm{n} \geq 1$, where

  $X_0, \{V_n\}_{n=1}^{\infty}, \{W_n\}_{n=1}^{\infty}$ are orthogonal zero-mean Random Vectors,

  with Σ_v and Σ_W being the covariance matrices for each V_n and W_n, resp.

  Find $\hat{X}_{\mathrm{n|n}} \coloneqq L[X_n | Y_1, \ldots, Y_n], \mathrm{n} \geq 1$, recursively.

- Solution:

  Compute $\hat{X}_{1|1}$ and iterate for $n \geq 2$ using

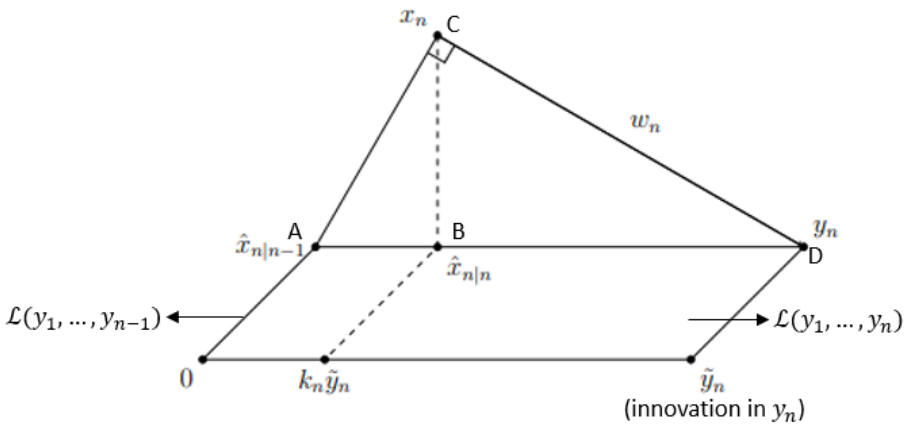
  $\hat{X}_{\mathrm{n|n}} = A\hat{X}_{\mathrm{n\text{-}1|n\text{-}1}} + K_n(Y_n - CA\hat{X}_{\mathrm{n\text{-}1|n\text{-}1}})$, where K_n's can be computed offline.

  Compute $\Sigma_{1|1}$ and iterate for $n \geq 2$ using

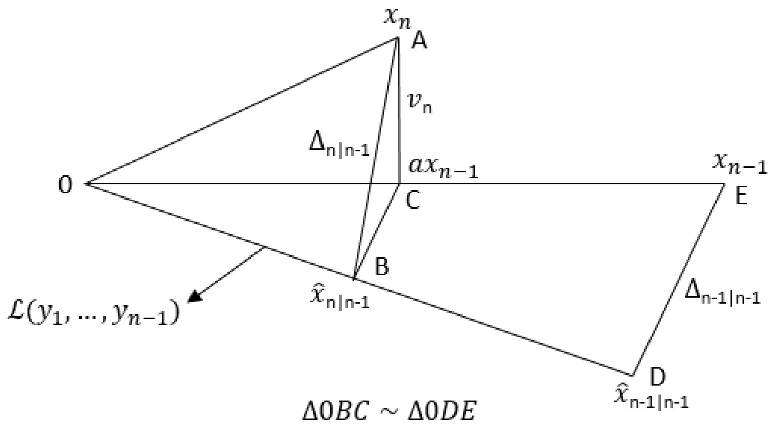
  $cov(X_n - \hat{X}_{\mathrm{n|n\text{-}1}}) \coloneqq \Sigma_{n|n-1} = A\Sigma_{n-1|n-1}A' + \Sigma_V$

  $K_n = \Sigma_{n|n-1}C'(C\Sigma_{n|n-1}C' + \Sigma_W)^{-1}$

  $cov(X_n - \hat{X}_{\mathrm{n|n}}) \coloneqq \Sigma_{n|n} = (I - K_nC)\Sigma_{n|n-1}$



$\Delta ABC \sim \Delta ACD$



x_n
A
v_n
$\Delta_{n|n-1}$
0
ax_{n-1}
C
x_{n-1}
E
B
$\hat{x}_{n|n-1}$
$\mathcal{L}(y_1, \ldots, y_{n-1})$
$\Delta_{n-1|n-1}$
D
$\hat{x}_{n-1|n-1}$
$\Delta 0BC \sim \Delta 0DE$