Endterm Review

EECS 126

Vipul Gupta

UC Berkeley

Warm-up

Consider two random variables X and Y. Is the following statement true or false. If $L[X|Y] = E[X|Y]$, then X and Y are jointly Gaussian. Either argue that it is correct, or provide a counterexample.

Warm-up

Consider two random variables X and Y. Is the following statement true or false. If $L[X|Y] = E[X|Y]$, then X and Y are jointly Gaussian. Either argue that it is correct, or provide a counterexample.

The statement is wrong. For example, take $X = Y = U[0, 1]$. Or any X and Y that have a linear dependence on each other. Or they can even be independent.

Still Warming up

Consider a Poisson process $\{N_t, t > 0\}$. Let T_n be the random variable denoting the time of n-th arrival. Find $MMSE[T_2|T_{10}]$.

Still Warming up

Consider a Poisson process $\{N_t, t > 0\}$. Let T_n be the random variable denoting the time of n-th arrival. Find $MMSE[T_2|T_{10}]$.

Given T_{10}, the previous arrivals are uniformly distributed between 0 and T_{10}. Thus, the second arrival has expected value of $2T_{10}/10$.

Some title related to MLE and MMSE

WiFi is not working for Kurtland, so he shows up at an Internet cafe at time 0 and spends his time exclusively typing emails (what a nerd!). The times that his emails are sent are modeled by a Poisson process with rate λ emails per hour.

(a) Let Y_1 and Y_2 be the times at which Kurtland's first and second emails are sent. Find the joint pdf of Y_1 and Y_2.

(b) Find $MMSE[Y_2|Y_1]$ and $LLSE[Y_2|Y_1]$. Hint: Don't use part (a).

(c) You watch Kurtland for an hour and observe that he has sent exactly 5 emails. Find the MLE of λ. (Any intuitions on what the answer should be?)

“Some title” solution

(a) Let Y_1 and Y_2 be the times at which Kurtland's first and second emails are sent. Find the joint pdf of Y_1 and Y_2.
The joint pdf is

$$f(y_2, y_1) = f(y_1)f(y_2|y_1) = \lambda e^{-\lambda y_1} \lambda e^{-\lambda(y_2 - y_1)} 1\{0 \leq y_1 \leq y_2\}$$
$$= \lambda^2 e^{-\lambda y_2} 1\{0 \leq y_1 \leq y_2\}.$$

(b) Find $MMSE[Y_2|Y_1]$ and $LLSE[Y_2|Y_1]$.
By memoryless property, MMSE estimate is $\mathbb{E}[Y_2|Y_1] = Y_1 + 1/\lambda$, which is linear and hence also equal to MMSE.

(c) You watch Kurtland for an hour and observe that he has sent exactly 5 emails. Find the MLE of λ.
$\arg\max_\lambda Pr(5 \text{ emails}|\lambda) = \arg\max_\lambda \frac{\lambda^5 e^{-\lambda}}{5!}$. Thus, $\lambda = 5$, and hence, average emails per hour is 5 which is intuitive.

Quadratic Estimator

Smart Alvin thinks he has uncovered a good model for the relative change in daily stock price of XYZ Inc., a publicly traded company in the New York Stock Exchange. His model is that the relative change in price, X, depends on the relative change in price of oil, Y, and some unpredictable factors, modeled collectively as a random variable Z. That is,

$$X = Y + 2Z + Y^2$$

In his model, Y is continuous RV uniformly distributed between -1 and 1 and Z is independent of Y with mean $\mathbb{E}[Z] = 0$ and $Var(Z) = 1$.

(a) Smart Alvin first decides to use a Linear Least Square Estimator of X given Y. Find $L[X|Y]$. What is the MSE of Smart Alvin's LLSE?

(b) Smart Alvin now decides to use a more sophisticated quadratic least squares estimator for X given Y, i.e. an estimator of the form $Q[X|Y] = aY^2 + bY + c$. Find $Q[X|Y]$ (intuition?).

(c) Which estimator has a lower mean squared error (MSE)?

Quadratic Estimator solution

(a) Smart Alvin first decides to use a Linear Least Square Estimator of X given Y. Find $L[X|Y]$. What is the MSE of Alvin's LLSE?

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

We calculate each term: $\mathbb{E}(X) = \mathbb{E}(Y^2) = 1/3, \mathbb{E}(Y) = 0, var(Y) = 1/3, cov(X,Y) = E(XY) - E(X)E(Y) = E(Y^2 + Y^3 + 2ZY) = 1/3.$

So $L[X|Y] = 1/3 + Y$.

MSE $=$
$\mathbb{E}[(X - L(X|Y))^2] = \mathbb{E}[(Y^2 - 1/3)^2] + 4Var(Z) = Var(Y^2) + 4Var(Z)$

Quadratic Estimator solution

(b) Smart Alvin now decides to use a more sophisticated quadratic least squares estimator for X given Y , i.e. an estimator of the form $Q[X|Y] = aY^2 + bY + c$. Find $Q[X|Y]$.

First, note that the pdf of Y and Z is symmetric around 0. Now, by orthogonality principle we have

$$\mathbb{E}[X - (aY^2 + bY + c)] = 0 \Rightarrow 1/3 - a/3 - c = 0$$
$$\mathbb{E}[(X - (aY^2 + bY + c))Y] = 0 \Rightarrow 1/3 - b/3 = 0$$
$$\mathbb{E}[(X - aY^2 - bY - c)Y^2] = 0 \Rightarrow (1 - a) \times 1/5 - c/3 = 0.$$

For the last equation we used $E[Y^4] = 2\int_0^1 \frac{1}{2} y^4 dy = 1/5$ and $\mathbb{E}[XY^2] = E[Y^3 + 2ZY^2 + Y^4] = E[Y^4] = 2/5$. This gives $Q(X|Y) = Y^2 + Y$.

(c) Which MSE is better?
QSE is a better estimate as its MSE $= \mathbb{E}[(X - Q(X|Y))^2] = 4Var(Z)$.

Hypothesis testing

Consider a Poisson point process. The null hypothesis is that it is a Poisson process of rate λ_0, and the alternate hypothesis is that it is a Poisson process of rate λ_1. Here $\lambda_1 > \lambda_0 > 0$.

Suppose we observe the total number of points n in the process over the time interval $[0, T]$. Describe the optimal
(a) Bayesian and
(b) Neyman Pearson (NP)
hypothesis test for this problem. For NP test, assume the maximum probability of false alarm to be ϵ, where $0 < \epsilon < 1$.

Hypothesis testing solution

The likelihood ratio between the hypotheses is the function on this set given by the ratio of the respective pmfs:

$$l(n) = \frac{(\lambda_1 T)^n e^{-\lambda_1 T}/n!}{(\lambda_0 T)^n e^{-\lambda_0 T}/n!} = (\frac{\lambda_1}{\lambda_2})^n e^{-(\lambda_1 - \lambda_0)T}.$$

This is a monotone increasing function of n.

(a) Bayesian test is generally simpler. Choose process 1 if $l(n) >= 1$, i.e.

$$n >= \frac{(\lambda_1 - \lambda_0)T}{log(\lambda_1) - log(\lambda_0)}.$$

Hypothesis testing solution

$$l(n) = \frac{(\lambda_1 T)^n e^{-\lambda_1 T}/n!}{(\lambda_0 T)^n e^{-\lambda_0 T}/n!} = (\frac{\lambda_1}{\lambda_2})^n e^{-(\lambda_1 - \lambda_0)T}.$$

(b) The optimal Neyman Pearson test is a (randomized) threshold rule based on this likelihood ratio. Since the likelihood ratio is a monotone increasing function of n, the optimal rule will decide hypothesis 1 is true if the observed number of points in $[0, T]$ is large enough. More precisely, depending on ϵ, we find $n_0 \geq 0$ and $0 < \delta < 1$ such that

$$\sum_{n=n_0+1}^{\infty} \frac{\lambda_0^n}{n!} e^{-\lambda_0} + \delta \frac{\lambda_0^{n_0}}{n_0!} e^{-\lambda_0} = \epsilon$$

The optimal rule for allowed probability of false alarm ϵ decides that hypothesis 1 is true whenever the observed number of points exceeds n_0, while if the observed number of points equals n_0 it decides that hypothesis 1 is true with probability δ. Use PYTHON/MATLAB to solve. LAB idea!

Tricky MMSE!

Let X, Y be i.i.d. $N(0,1)$. Find $\mathbb{E}[X|(X+Y)^3]$.

Tricky MMSE!

Let X, Y be i.i.d. $N(0,1)$. Find $\mathbb{E}[X|(X+Y)^3]$.

Hint: What is $\mathbb{E}[X|X+Y]$?

Tricky MMSE!

Let X, Y be i.i.d. $N(0,1)$. Find $\mathbb{E}[X|(X+Y)^3]$.

Let $Z = (X+Y)^3$. Given Z, one finds $X+Y = Z^{1/3}$. By symmetry, $\mathbb{E}[X|X+Y] = (X+Y)/2$. Hence,

$$\mathbb{E}[X|Z] = \frac{1}{2}Z^{1/3}.$$

Jointly Gaussian

Let X_1, X_2, X_3 be jointly Gaussian with mean $[1, 4, 6]^T$ and covariance matrix $\begin{bmatrix} 3 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$. Find $MMSE(X_1|X_2, X_3)$.

Jointly Gaussian

Let X_1, X_2, X_3 be jointly Gaussian with mean $[1, 4, 6]^T$ and covariance matrix $\begin{bmatrix} 3 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$. Find $MMSE(X_1|X_2, X_3)$.

$MMSE(X_1|X_2, X_3) = \mathbb{E}(X_1|X_2, X_3)$ for jointly Gaussian RVs can be expressed as

$$E[X_1|X_2, X_3] = a_0 + a_1(X_2 - 4) + a_2(X_3 - 6).$$

(We subtract 4 and 6 from X_2 and X_3, respectively, to make them zero-centered to help with calculations). The equation $\mathbb{E}[\mathbb{E}[X_1|X_2, X_3]] = \mathbb{E}[X_1]$ gives $a_0 = 1$. The requirements that $X_1 - (a_0 + a_1(X_2 - 4) + a_2(X_3 - 6))$ be uncorrelated with $(X_2 - 4)$ and $(X_3 - 6)$ gives us two more equations. Solving them using the covariance matrix information yields

$$E\left[X_1|X_2, X_3\right] = 1 + (X_2 - 4) - (X_3 - 6) = X_2 - X_3 + 3.$$