UC Berkeley
Department of Electrical Engineering and Computer Science

EE 126: Probablity and Random Processes

Problem Set 9
Fall 2007

Issued: Thursday, November 1, 2007 **Due:** Friday, November 9, 2007

Reading: Bertsekas & Tsitsiklis, Chapter 4, §5.1

Chapter 4 problems

Problem 9.1

For this problem, we consider the form of the linear least squares estimator of X when two measurements are available. More specifically, find a_1, a_2, and b such that the linear predictor $g(Y_1, Y_2) = a_1Y_1 + a_2Y_2 + b$ minimizes

$$\mathbb{E}[(X - g(Y_1, Y_2))^2] \;=\; \mathbb{E}[(X - a_1Y_1 - a_2Y_2 - b)^2].$$

To make the problem a little more tractable, assume that Y_1 and Y_2 are uncorrelated (i.e., $\mathbb{E}[Y_1Y_2] = \mathbb{E}[Y_1]\mathbb{E}[Y_2]$).

Solution:

In this problem, we want to find the linear estimator $g(Y_1, Y_2)$ that minimizes $\mathbb{E}[(X - g(Y_1, Y_2))^2]$. This is an extension of the case of single measurement. Our estimator is of the form

$$g(Y_1, Y_2) = a_1Y_1 + a_2Y_2 + b$$

and therefore our goal is to find a_1, a_2 and b that solve

$$\text{minimize } (\mathbb{E}[(X - a_1Y_1 - a_2Y_2 - b)^2]) \qquad (*)$$

To achieve $(*)$, we must satisfy

$$\begin{array}{rclcl}
\frac{\partial \mathbb{E}[(X-a_1Y_1-a_2Y_2-b)^2]}{\partial b} & = & \mathbb{E}[2(X - a_1Y_1 - a_2Y_2 - b)(-1)] & = & 0 \\
\frac{\partial \mathbb{E}[(X-a_1Y_1-a_2Y_2-b)^2]}{\partial a_1} & = & \mathbb{E}[2(X - a_1Y_1 - a_2Y_2 - b)(-Y_1)] & = & 0 \\
\frac{\partial \mathbb{E}[(X-a_1Y_1-a_2Y_2-b)^2]}{\partial a_2} & = & \mathbb{E}[2(X - a_1Y_1 - a_2Y_2 - b)(-Y_2)] & = & 0
\end{array}$$

and $\partial^2\mathbb{E}[\cdot]/\partial b^2$, $\partial^2\mathbb{E}[\cdot]/\partial {a_1}^2$, and $\partial^2\mathbb{E}[\cdot]/\partial {a_2}^2$ must be > 0.

By the linearity of expectation,

$$\begin{array}{rcll}
b & = & \mathbb{E}[X] - a_1\mathbb{E}[Y_1] - a_2\mathbb{E}[Y_2] & (1) \\
a_1\mathbb{E}[Y_1^2] & = & \mathbb{E}[XY_1] - a_2\mathbb{E}[Y_1Y_2] - b\mathbb{E}[Y_1] & (2) \\
a_2\mathbb{E}[Y_2^2] & = & \mathbb{E}[XY_2] - a_1\mathbb{E}[Y_1Y_2] - b\mathbb{E}[Y_2]. & (3)
\end{array}$$

We now have 3 equations to solve for 3 unknowns.

Consider $(2) - \mathbb{E}[Y_1] \cdot (1)$, we obtain

$$a_1\mathbb{E}[Y_1^2] - b\mathbb{E}[Y_1] = \mathbb{E}[XY_1] - a_2\mathbb{E}[Y_1Y_2] - b\mathbb{E}[Y_1] - \mathbb{E}[X]\mathbb{E}[Y_1] + a_1(\mathbb{E}[Y_1])^2 + a_2\mathbb{E}[Y_1]\mathbb{E}[Y_2].$$

Arranging algebra and use the fact that $\mathbb{E}[Y_1Y_2] = \mathbb{E}[Y_1]\mathbb{E}[Y_2]$

$$a_1(\mathbb{E}[Y_1^2] - (\mathbb{E}[Y_1])^2) = \mathbb{E}[XY_1] - \mathbb{E}[X]\mathbb{E}[Y_1].$$

Similarly,

$$a_2(\mathbb{E}[Y_2^2] - (\mathbb{E}[Y_2])^2) = \mathbb{E}[XY_2] - \mathbb{E}[X]\mathbb{E}[Y_2].$$

Therefore,

$$\begin{aligned} a_1 &= (\mathbb{E}[XY_1] - \mathbb{E}[X]\mathbb{E}[Y_1])/\sigma_{Y_1}^2 \\ a_2 &= (\mathbb{E}[XY_2] - \mathbb{E}[X]\mathbb{E}[Y_2])/\sigma_{Y_2}^2 \\ b &= \mathbb{E}[X] - a_1\mathbb{E}[Y_1] - a_2\mathbb{E}[Y_2]. \end{aligned}$$

Writing this expression in the similar term as the case of single measurement:

$$g(Y_1, Y_2) = \mathbb{E}[X] + \frac{\mathrm{Cov}(X,Y_1)}{\sigma_{Y_1}^2}(Y_1 - \mathbb{E}[Y_1]) + \frac{\mathrm{Cov}(X,Y_2)}{\sigma_{Y_2}^2}(Y_2 - \mathbb{E}[Y_2])$$

where we use the fact that $\mathrm{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$.

Note: Convince yourself that the second order condition is satisfied.

Problem 9.2

Bob has gone hiking, and is lost in the forest. In order to try and find a road, he decides on the following *distance/coin-flip* strategy. At time instants $t = 1, 2, 3 \ldots$, he chooses a distance uniformly at random between t and $t + 1$. Independently of the chosen random distance, he then flips a fair coin; if it comes up heads, he moves the chosen random distance to the right (positive on the real line), and otherwise for a tails toss, he moves the chosen random distance to the left (negative on the real line). Both the random distance and the coin flip are independent random variables for different time instants. Assume that he starts at the origin at time instant $t = 0$.

(a) Let Y_s be Bob's position after repeating his distance/coin-flip strategy for a fixed number of s time instants. Compute its expected value and variance as a function of s.

Now suppose that Bob repeats his distance/coin-flip strategy for a *random number* S of time rounds, after which he stops. Assume that $S \sim \mathrm{Geo}(p)$ has a geometric distribution with parameter p, and let $X \in \mathbb{R}$ be his final position. For any question below, you may feel free to express your answer (if appropriate) in terms of the moments $\mu_i = \mathbb{E}[S^i], i = 1, 2, 3 \ldots$.

(b) Suppose that you observe that $S = s$. What is the minimum mean squared error (MMSE) estimator of X given this information?

(c) What is the expected value and variance of his position X?

(d) Now suppose that you observe that Bob finishes at position $X = x$. Given this information, what is the linear least squares estimator (LLSE) of the number of time rounds S that he repeated his distance/coin-flip strategy?

(e) What is the linear least-squares estimate of S based on X^2?

Solution:

(a) Let X_t denote the distance that Bob travels at time instant t. X_t is uniformly distributed on two line segments $[-(t+1), -t]$ and $[t, t+1]$. Obviously, $\mathbb{E}(X_t) = 0$, and

$$Var(X_t) = \mathbb{E}(X_t^2) = \int_t^{t+1} x^2 dx = \frac{(t+1)^3}{3} - \frac{t^3}{3}$$

Since, $Y_s = \sum_{t=1}^{s} X_t$ and $\{X_t\}_{t=1}^{s}$ are independent random variables,

$$\begin{aligned} \mathbb{E}(Y_s) &= \sum_{t=1}^{s} \mathbb{E}(X_t) = 0 \\ Var(Y_s) &= \sum_{t=1}^{s} Var(X_t) = \frac{(s+1)^3 - 1}{3} \end{aligned}$$

(b) The minimum mean squared error estimator of X given $S = s$ is the conditional expection of X given $S = s$. This value is the same as $E(Y_s)$, which is 0.

(c)

$$\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X|S)) = \mathbb{E}(0) = 0$$

$$\begin{aligned} Var(X) &= \mathbb{E}(Var(X|S)) + Var(\mathbb{E}(X|S)) \\ &= \mathbb{E}(Var(Y_s)) + Var(0) \\ &= \mathbb{E}(\frac{(S+1)^3 - 1}{3}) \\ &= \mu_1 + \mu_2 + \mu_3/3 \end{aligned}$$

(d) The covariance of X and S is

$$\begin{aligned} Cov(X, S) &= \mathbb{E}(XS) - \mathbb{E}(X)\mathbb{E}(S) \\ &= \mathbb{E}(XS) \\ &= \mathbb{E}(E(XS|S = s)) \\ &= 0 \end{aligned}$$

Thus, the linear least square estimator of S given X is $\mathbb{E}(S) = \mu_1$.

(e) The linear least square estimator of S given X^2 can be written as:

$$\mathbb{E}(S) + \frac{Cov(X^2, S)}{Var(X^2)}(X^2 - \mathbb{E}(X^2))$$

where $\mathbb{E}(S) = \mu_1$ and $\mathbb{E}(X^2) = Var(X) = \mu_1 + \mu_2 + \mu_3/3$. By definition, the covariance between X^2 and S is

$$\begin{aligned} Cov(X^2, S) &= \mathbb{E}(X^2 S) - \mathbb{E}(X^2)\mathbb{E}(S) \\ &= \mathbb{E}(\mathbb{E}(X^2 S|S)) - \mathbb{E}(X^2)\mathbb{E}(S) \\ &= \mathbb{E}(S\frac{(S+1)^3 - 1}{3}) - \mathbb{E}(X^2)\mathbb{E}(S) \\ &= (\mu_2 + \mu_3 + \mu_4/3) - (\mu_1 + \mu_2 + \mu_3/3)\mu_1 \end{aligned}$$

To compute the variance of X^2, we use the law of conditional variances.

$$Var(X^2) = \mathbb{E}(Var(X^2|S)) + Var(\mathbb{E}(X^2|S))$$

$$\begin{aligned} Var(\mathbb{E}(X^2|S)) &= Var(S^3/3 + S^2 + S) \\ &= \mathbb{E}((S^3/3 + S^2 + S)^2) - (\mathbb{E}(S^3/3 + S^2 + S))^2 \\ &= (\mu_6/9 + 2/3\mu_5 + 5/3\mu_4 + 2\mu_3 + \mu_2) - (\mu_3/3 + \mu_2 + \mu_1)^2 \end{aligned}$$

To compute $Var(X^2|S = s)$, we rewrite X^2 as a sum over X_iX_j.

$$\begin{aligned} & Var(X^2|S = s) \\ = \; & Var((\sum_{t=1}^{s} X_t)^2) \\ = \; & \sum_{t=1}^{s} Var(X_t^2) + \sum_{i\neq j} Var(X_iX_j) + \sum_{i\neq j} Cov(X_i^2, X_j^2) + \\ & \sum_{k=1}^{s}\sum_{i\neq j} Cov(X_iX_j, X_k^2) + \sum_{(i,j)\neq(k,l), i\neq j, k\neq l} Cov(X_iX_j, X_kX_l) \end{aligned}$$

We will see that the most of covariances are zero. Since X_i and X_j are independent, $Cov(X_i^2, X_j^2) = 0$. If $k \neq i$ and $k \neq j$,

$$Cov(X_iX_j, X_k^2) = \mathbb{E}(X_iX_jX_k^2) - \mathbb{E}(X_i)\mathbb{E}(X_j)\mathbb{E}(X_k^2) = \mathbb{E}(X_i)\mathbb{E}(X_j)\mathbb{E}(X_k^2) - 0 = 0$$

If k equals to one of i and j, let's assume $k = i$,

$$Cov(X_iX_j, X_k^2) = \mathbb{E}(X_i^3X_j) - \mathbb{E}(X_i)\mathbb{E}(X_j)\mathbb{E}(X_i^2) = \mathbb{E}(X_i^3)\mathbb{E}(X_j) - 0 = 0$$

Given the fact that $(i, j) \neq (k, l)$, $Cov(X_iX_j, X_kX_l)$ is non-zero, ONLY when $k = j, l = i$.

$$\begin{aligned} & Var(X^2|S = s) \\ = \; & \sum_{t=1}^{s} Var(X_t^2) + 2\sum_{i\neq j} Var(X_iX_j) \\ = \; & \boxed{\sum_{t=1}^{s}(\mathbb{E}(X_t^4) - (\mathbb{E}(X_t^2))^2) + 2\sum_{i\neq j}\mathbb{E}(X_i^2)\mathbb{E}(X_j^2)} \\ = \; & \sum_{t=1}^{s}\mathbb{E}(X_t^4) - 3(\sum_{t=1}^{s}\mathbb{E}(X_t^2)^2) + 2(\sum_{t=1}^{s}\mathbb{E}(X_t^2))^2 \\ = \; & \frac{(S+1)^5 - 1}{5} - \frac{1}{15}(9S^5 + 45S^4 + 85S^3 + 75S^2 + 31S) + \frac{2(3S^2 + 3S + 1)^2}{9} \\ = \; & (10 + 12S + 15S^2 + 15S^3 - 18S^5)/45 \end{aligned}$$

Note that we actually use *mathematica*, a symbolic tool for math, to get the summation: $\sum_{t=1}^{s}\mathbb{E}(X_t^2)^2 = \sum_{t=1}^{s}(t^3/3 + t^2 + t)^2$.

Therefore, the variance of X^2 is

$$(5\mu_6 + 21\mu_5 + 75\mu_4 + 100\mu_3 + 60\mu_2 + 13\mu_1 + 5)/45 - (\mu_3/3 + \mu_2 + \mu_1)^2$$

Putting all results into one expression, the linear least square estimator of S given X^2 is

$$\mu_1 + \frac{(\mu_2+\mu_3+\mu_4/3)-(\mu_1+\mu_2+\mu_3/3)\mu_1}{(5\mu_6+21\mu_5+75\mu_4+100\mu_3+60\mu_2+13\mu_1+5)/45-(\mu_3/3+\mu_2+\mu_1)^2}\Big(X^2-(\mu_1+\mu_2+\mu_3/3)\Big)$$

Students received full credit for reasoning up to the boxed line above.

Chapter 5 problems

Problem 9.3

Each of n packages is loaded independently onto either a red truck (with probability p) or onto a green truck (with probability $1-p$). Let R be the total number of items selected for the red truck and let G be the total number of items selected for the green truck.

(a) Determine the PMF, expected value, and variance of the random variable R.

(b) Evaluate the probability that the first item to be loaded ends up being the only one on its truck.

(c) Evaluate the probability that at least one truck ends up with a total of exactly one package.

(d) Evaluate the expected value and the variance of the difference $R-G$.

(e) Assume that $n \geq 2$. Given that both of the first two packages are loaded onto the red truck, find the conditional expectation, variance, and PMF of the random variable R.

Solution:

(a) R is a binomial random variable with parameters p and n, hence

$$p_R(r) = \binom{n}{r}(1-p)^{n-r}p^r, \qquad \text{for } r = 0,\ 1,\ 2,\ \ldots,\ n.$$

Note that R can be written as a sum of i.i.d. Bernoulli random variables as follows:

$$R = X_1 + X_2 + \cdots X_n,$$

where for each i

$$p_{X_i}(x) = \begin{cases} p, & \text{if } x = 1; \\ (1-p), & \text{if } x = 0. \end{cases}$$

It is easy to verify that $E(X_i) = p$ and $\text{var}(X_i) = p(1-p)$. Then

$$E(R) = E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n) = np$$

and, because of the independence of X_is,

$$\text{var}(R) = \text{var}(X_1) + \text{var}(X_2) + \cdots + \text{var}(X_n) = n\ \text{var}(X_i) = np(1-p).$$

(b) Denote the event of interest by A. A is the union of the following two mutually exclusive events:

- The first item is placed in the red truck and the remaining $n-1$ are placed in the green truck.
- The first item is placed in the green truck and the remaining $n-1$ are placed in the red truck.

Thus the probability of interest is the sum of the probabilities of the two events above:

$$P(A) = p(1-p)^{n-1} + (1-p)p^{n-1}.$$

(c) Denote the event of interest by B. It occurs if exactly one OR both of the trucks end up with exactly 1 package. Thus,

$$P(B) = \begin{cases} 1, & \text{if } n = 1; \\ p(1-p) + (1-p)p, & \text{if } n = 2; \\ \begin{pmatrix} n \\ 1 \end{pmatrix} (1-p)^{n-1}p + \begin{pmatrix} n \\ 1 \end{pmatrix} (p)^{n-1}(1-p) & \text{if } n = 3, 4, 5, \ldots. \end{cases}$$

(d) $E(D) = E(R-G) = E(R-(n-R)) = E(2R-n) = 2E(R) - n = 2np - n.$

Since $D = 2R - n$ where n is a constant, $\text{var}(D) = 4\,\text{var}(R) = 4np(1-p)$.

(e) Let C be the event that both of the first two packages to be loaded go onto the red truck. The random variable R given event C becomes

$$R_{|C} = 2 + X_3 + X_4 + \cdots + X_n.$$

Hence,

$$E(R \mid C) = E(2 + X_3 + X_4 + \cdots + X_n) = 2 + (n-2)E(X) = 2 + (n-2)p.$$

Similarly the conditional variance of R is

$$\text{var}(R|C) = \text{var}(2 + X_3 + X_4 + \cdots + X_n) = (n-2)\ \text{var}(X) = (n-2)p(1-p).$$

Finally the conditional PMF for R given C is simply the unconditional PMF for R shifted to the right by 2. Hence,

$$p_{R|C}(r|c) = \begin{pmatrix} n-2 \\ r-2 \end{pmatrix} (1-p)^{(n-r)}p^{(r-2)} \qquad \text{for } r = 2,\ 3,\ \ldots,\ n.$$

Problem 9.4

Fred is giving out samples of dog food. He makes calls door to door, but he leaves a sample (one can) only on those calls for which the door is answered *and* a dog is in residence. On any call the probability of the door being answered is 3/4, and the probability that any household has a dog is 2/3. Assume that the events "Door answered" and "A dog lives here" are independent and also that the outcomes of all calls are independent.

(a) Determine the probability that Fred gives away his first sample on his third call.

(b) Given that he has given away exactly four samples on his first eight calls, determine the conditional probability that Fred will give away his fifth sample on his eleventh call.

(c) Determine the probability that he gives away his second sample on his fifth call.

(d) Given that he did not give away his second sample on his second call, determine the conditional probability that he will leave his second sample on his fifth call.

(e) We will say that Fred "needs a new supply" immediately *after* the call on which he gives away his last can. If he starts out with two cans, determine the probability that he completes at least five calls before he needs a new supply.

(f) If he starts out with exactly m cans, determine the expected value and variance of D_m, the number of homes with dogs which he passes up (because of no answer) before he needs a new supply.

Solution:
A successful call occurs with probability $p = \frac{3}{4} \cdot \frac{2}{3} = \frac{1}{2}$.

(a) Fred will give away his first sample on the third call if the first two calls are failures and the third is a success. Since the trials are independent, the probability of this sequence of events is simply

$$(1-p)(1-p)p = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$$

(b) The event of interest requires failures on the ninth and tenth trials and a success on the eleventh trial. For a Bernoulli process, the outcomes of these three trials are independent of the results of any other trials and again our answer is

$$(1-p)(1-p)p = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$$

(c) We desire the probability that L_2, the second-order interarrival time is equal to five trials. We know that $p_{L_2}(l)$ is a Pascal PMF, and we have

$$p_{L_2}(5) = \binom{5-1}{2-1} p^2 (1-p)^{5-2} = 4 \cdot \left(\frac{1}{2}\right)^5 = \frac{1}{8}$$

(d) Here we require the conditional probability that the experimental value of L_2 is equal to 5, given that it is greater than 2.

$$\begin{aligned} p_{L_2|L_2>2}(5|L_2>2) &= \frac{p_{L_2}(5)}{P(L_2>2)} = \frac{p_{L_2}(5)}{1-p_{L_2}(2)} \\ &= \frac{\binom{5-1}{2-1} p^2(1-p)^{5-2}}{1-\binom{2-1}{2-1} p^2 (1-p)^0} = \frac{4 \cdot \left(\frac{1}{2}\right)^5}{1-\left(\frac{1}{2}\right)^2} = \frac{1}{6} \end{aligned}$$

(e) The probability that Fred will complete at least five calls before he needs a new supply is equal to the probability that the experimental value of L_2 is greater than or equal to 5.

$$\begin{aligned} P(L_2 \geq 5) = 1 - P(L_2 \leq 4) &= 1 - \sum_{l=2}^{4} \binom{l-1}{2-1} p^2 (1-p)^{l-2} \\ &= 1 - (\frac{1}{2})^2 - \binom{2}{1}(\frac{1}{2})^3 - \binom{3}{1}(\frac{1}{2})^4 = \frac{5}{16} \end{aligned}$$

(f) Let discrete random variable F represent the number of failures before Fred runs out of samples on his mth successful call. Since L_m is the number of trials up to and including the mth success, we have $F = L_m - m$. Given that Fred makes L_m calls before he needs a new supply, we can regard each of the F unsuccessful calls as trials in another Bernoulli process with parameter r, where r is the probability of a success (a disappointed dog) obtained by

$$\begin{aligned} r &= \text{P(dog lives there | Fred did not leave a sample)} \\ &= \frac{\text{P(dog lives there AND door not answered)}}{1 - \text{P(giving away a sample)}} = \frac{\frac{1}{4} \cdot \frac{2}{3}}{1 - \frac{1}{2}} = \frac{1}{3} \end{aligned}$$

We define X to be a Bernoulli random variable with parameter r. Then, the number of dogs passed up before Fred runs out, D_m, is equal to the sum of F Bernoulli random variables each with parameter $r = \frac{1}{3}$, where F is a random variable. In other words,

$$D_m = X_i + X_2 + X_3 + \cdots + X_F.$$

Note that D_m is a sum of a random number of independent random variables. Further, F is independent of the X_i's since the X_i's are defined in the conditional universe where the door is not answered, in which case, whether there is a dog or not does not affect the probability of that trial being a failed trial or not. From our results in class, we can calculate its expectation and variance by

$$\begin{aligned} E[D_m] &= E[F]E[X] \\ \text{var}(D_m) &= E[F]\text{var}(X) + (E[X])^2\text{var}(F), \end{aligned}$$

where we make the following substitutions.

$$\begin{aligned} E[F] &= E[L_m - m] = \frac{m}{p} - m = m. \\ \text{var}(F) &= \text{var}(L_m - m) = \text{var}(L_m) = \frac{m(1-p)}{p^2} = 2m. \\ E[X] &= r = \frac{1}{3}. \\ \text{var}(X) &= r(1-r) = \frac{2}{9}. \end{aligned}$$

Finally, substituting these values, we have

$$\begin{aligned} E[D_m] &= m \cdot \frac{1}{3} = \frac{m}{3} \\ \mathrm{var}(D_m) &= m \cdot \frac{2}{9} + \left(\frac{1}{3}\right)^2 \cdot 2m = \frac{4m}{9} \end{aligned}$$