ACM 116 Problem Set 3 Solutions

Hannes Helgason

Problem 1

(a) Since A and B are i.i.d. zero-mean random variables we get that the mean of $X(t)$ is

$$E(X(t)) = \cos(\omega t)EA + \sin(\omega t)EB = 0,$$

independent of t, and the autocorrelation,

$$\begin{aligned} E(X(t)X(t+\tau)) &= E\Big(A^2\cos(\omega t)\cos(\omega(t+\tau)) + AB\cos(\omega t)\sin(\omega(t+\tau)) \\ &\quad + B^2\sin(\omega t)\sin(\omega(t+\tau)) + BA\sin(\omega t)\cos(\omega(t+\tau))\Big) \\ &= E(A^2)\Big(\cos(\omega t)\cos(\omega(t+\tau)) + \sin(\omega t)\sin(\omega(t+\tau))\Big) \\ &= E(A^2)\cos(\omega\tau) \end{aligned}$$

depends only on the time difference τ (we assume that the variance of A and B is finite, and therefore so is the variance of the process). Thus, the process $X(t)$ is weakly stationary.

(b) If a process $X(t)$ is strongly stationary, all of its moments have to be independent of t. The third moment of $X(t)$,

$$E(X^3(t)) = (EA^3)(\cos^3(\omega t) + \sin^3(\omega t)),$$

is not a constant with respect to t unless $EA^3 = 0$. Clearly we can find many examples of distributions for A so that the third moment is not zero while $E(A) = 0$. For example take A to be a discrete random variable with $P(A = -1) = 1/3$, $P(A = 1/2) = 2/3$. Then $E(A) = -1 \times 1/3 + 1/2 \times 2/3 = 0$ and $E(A^3) = -1 \times 1/3 + 1/8 \times 2/3 = -1/4$.

We could just as well consider the fourth order moment:

$$E(X^4(t)) = (EA^4)(\cos^4(\omega t) + \sin^4(\omega t)) + 6(EA^2)^2\cos^2(\omega t)\sin^2(\omega t).$$

and taking $A \sim Unif[-1, 1]$ we get $E(X^4(t)) = 1/5(\cos^4(\omega t) + \sin^4(\omega t)) + 2/3\cos^2(\omega t)\sin^2(\omega t)$ so the fourth moment is in general not independent of t.

Above is one way to get a contradiction to show that in general the process is in general not strongly stationary. Another possible way would be to look at process for the times $t = 0$ and $t = \pi/\omega$, that is $X(0) = A$ and $X(\pi/\omega) = -A$. This implies that if the distribution of A is not symmetric around zero, the process $X(t)$ is not strictly stationary.

Problem 2

(a) Since $X(t)$ and $Y(t)$ are independent, with zero mean and have identical covariance functions we have that

$$E(Z^2(t)) = E\big(a^2X^2(t) + 2abX(t)Y(t) + b^2Y^2(t)\big) = (a^2 + b^2)E(X^2(t)) = (a^2 + b^2)R_X(0).$$

Thus, the second order moment of the process is a finite constant, independent of time t. The cross-correlation between $X(t)$ and $Y(t)$ is

$$E(X(t+\tau)Y(t)) = E(X(t+\tau))E(Y(t)) = 0,$$

so we get

$$\begin{aligned} E(Z(t)Z(t+\tau)) &= a^2E(X(t)X(t+\tau)) + abE(X(t)Y(t+\tau)) + abE(Y(t)X(t+\tau)) + b^2E(Y(t)Y(t+\tau)) \\ &= a^2R_X(\tau) + b^2R_Y(\tau), \end{aligned}$$

which gives us that the autocorrelation of $Z(t)$ depends only on τ, and since the mean of $Z(t)$, $E(Z(t)) = 0$, is constant, we conclude that $Y(t)$ is weakly stationary.

(b) If $X(t)$ and $Y(t)$ are also jointly Gaussian, then $Z(t)$, a linear combination of these two processes, is Gaussian with mean

$$E(Z(t)) = 0,$$

variance

$$E(Z^2(t)) = (a^2 + b^2)R_X(0) = (a^2 + b^2)\sigma_X^2,$$

where σ_X^2 is the variance of $X(t)$ and autocorrelation function equal to

$$R_Z(\tau) = E(Z(t)Z(t+\tau)) = (a^2 + b^2)R_X(\tau).$$

(Note that to get a full description of a Gaussian process, in addition to the mean, we need to know its autocorrelation, not just the variance.) For given times $t_1, t_2, \ldots, t_n$, the density of $\mathbf{Z} := [Z(t_1), Z(t_2), \ldots Z(t_n)]^T$ is $N(\mathbf{0}, \Sigma)$, where the entry (i, j) of the covariance matrix is $\Sigma_{i,j} = E[Z(t_i)Z(t_j)] = (a^2 + b^2)R_X(t_j - t_i)$.

Problem 3

(a) Lets first assume that the processes X_t and Z_t both have mean zero. We require the estimate of X_t, $\hat{X}_t$, to be linear:

$$\hat{X}_t = \sum_{k=t-p}^{t+p} h_{t-k}Y_k = \sum_{k=-p}^{p} h_kY_{t-k}.$$

To determine the optimum filter, we need to find h_k such that the mean square error is minimum. Or equivalently, the optimal filter must satisfy the orthogonality condition:

$$0 = E((X_t - \hat{X}_t)Y_j) \quad \forall j \in I.$$

That is, $E(X_tY_j) = E(\hat{X}_tY_j)$, $\forall j \in I$. Plugging in for $\hat{X}_t$ then gives:

$$\begin{aligned} E(X_tY_j) &= E\left(\sum_{k=-p}^{p} h_kY_{t-k}Y_j\right) \\ &= \sum_{k=-p}^{p} h_kE(Y_{t-k}Y_j) \quad \forall j \in I. \end{aligned}$$

If we assume that X_t and Y_t are jointly wide-sense stationary (and both wide-sense stationary) then we can write the equation above in the form

$$\gamma_{X,Y}(t-j) = \sum_{k=-p}^{p} h_k\gamma_Y(t-k-j) \quad \forall j \in I,$$

or by the substitution $m = t - j$,

$$\gamma_{X,Y}(m) = \sum_{k=-p}^{p} h_k\gamma_Y(m-k) \quad \forall m \in I.$$

(b) The matrix equation for the $2p+1$ coefficients is

$$\begin{pmatrix} \gamma_Y(0) & \gamma_Y(1) & \cdots & \gamma_Y(2p) \\ \gamma_Y(1) & \gamma_Y(0) & \cdots & \gamma_Y(2p-1) \\ \vdots & & \ddots & \vdots \\ \gamma_Y(2p) & & \cdots & \gamma_Y(0) \end{pmatrix} \begin{pmatrix} h_{-p} \\ h_{-p+1} \\ \vdots \\ h_p \end{pmatrix} = \begin{pmatrix} \gamma_{X,Y}(-p) \\ \gamma_{X,Y}(-p+1) \\ \vdots \\ \gamma_{X,Y}(p) \end{pmatrix}$$

(c) Since Z_t has mean zero and is independent of X_t we have $E(X_tY_j) = E(X_tX_j) + E(X_t)E(Z_j) = E(X_tX_j)$ and therefore

$$\gamma_{Y,X}(\tau) = E(Y_tX_{t+\tau}) = \gamma_X(\tau) = 4 \cdot 2^{-|\tau|}.$$

Using in addition the fact that Z_t is white noise with variance $\sigma^2 = 1$ we get

$$\begin{aligned} \gamma_Y(\tau) &= E(Y_tY_{t+\tau}) = E(X_tX_{t+\tau}) + E(Z_tZ_{t+\tau}) \\ &= \gamma_X(\tau) + \delta(\tau). \end{aligned}$$

Thus, the matrix equation for the optimal filter is

$$\begin{pmatrix} \gamma_Y(0) & \gamma_Y(1) & \gamma_Y(2) \\ \gamma_Y(1) & \gamma_Y(0) & \gamma_Y(1) \\ \gamma_Y(2) & \gamma_Y(1) & \gamma_Y(0) \end{pmatrix} \begin{pmatrix} h_{-1} \\ h_0 \\ h_1 \end{pmatrix} = \begin{pmatrix} \gamma_{X,Y}(-1) \\ \gamma_{X,Y}(0) \\ \gamma_{X,Y}(1) \end{pmatrix}$$

or

$$\begin{pmatrix} 5 & 2 & 1 \\ 2 & 5 & 2 \\ 1 & 2 & 5 \end{pmatrix} \begin{pmatrix} h_{-1} \\ h_0 \\ h_1 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 2 \end{pmatrix}$$

which gives

$$\begin{pmatrix} h_{-1} \\ h_0 \\ h_1 \end{pmatrix} = \begin{pmatrix} 1/11 \\ 8/11 \\ 1/11 \end{pmatrix}.$$

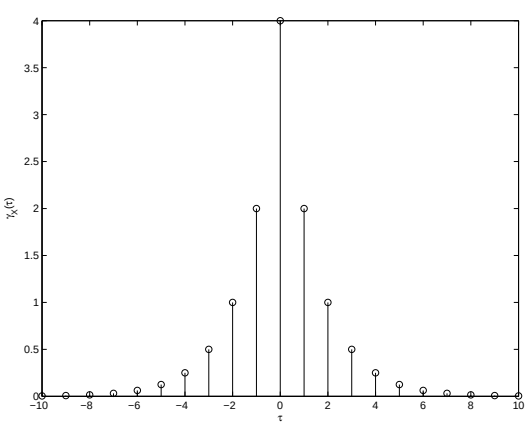
(d) Since the error $e_t = X_t - \hat{X}_t$ is orthogonal to every Y_j, it is indeed orthogonal to every linear combination of Y_j's and therefore $E(e_t\hat{X}_t) = 0$ and the mean squared error is

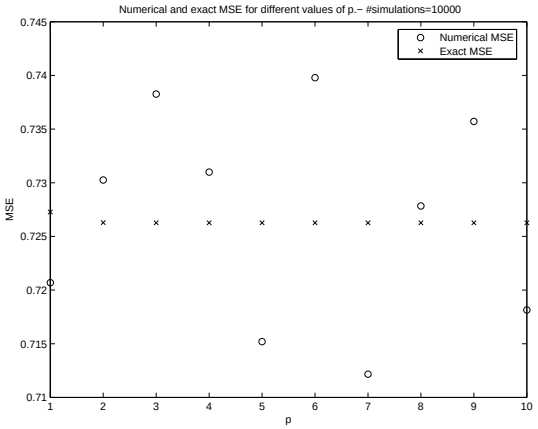
$$\begin{aligned} E(e_t^2) &= E(e_t(X_t - \hat{X}_t)) = E(e_tX_t) = E(X_t^2) - E(\hat{X}_tX_t) \\ &= \gamma_X(0) - \sum_{k=-p}^{p} h_k E(Y_{t-k}X_t) = \gamma_X(0) - \sum_{k=-p}^{p} h_k \gamma_{Y,X}(k). \end{aligned}$$

This gives us that the mean squared error is

$$\begin{aligned} E(e_t^2) &= 4 - \sum_{k=-p}^{p} h_k 4 \cdot 2^{-|k|} = 4 - 4(1/2h_{-1} + h_0 + 1/2h_1) \\ &= \frac{8}{11} \approx 0.727. \end{aligned}$$

(e) The `Matlab` code for this problem can be found at the end of this solution set. From the plot of the autocorrelation we see that it decreases very rapidly and for $\tau = 10$ it's close to 0.004. Thus, if we want to estimate the process X_t at time $t = t_0$, we don't expect to gain much by using measurements taken at times greater than 10 time-steps from t_0, or even less. On the figure that shows the exact MSE and numerical estimate of MSE for different values of p, and as we can see by increasing p, above say 2, doesn't improve MSE considerably.





Problem 4

(a) Let X_i be the time to be served by server i. As each X_i is exponentially distributed

$$f_{X_i}(x) = \begin{cases} \mu_i e^{-\mu_i x}, & \text{for } x \geq 0 \\ 0, & \text{otherwise.} \end{cases}$$

By conditioning on X_1 we get

$$\begin{aligned} P(X_3 > X_1) &= \int_0^\infty P(X_3 > X_1 \mid X_1 = x)\mu_1 e^{-\mu_1 x} dx \\ &= \int_0^\infty P(X_3 > x)\mu_1 e^{-\mu_1 x} dx \\ &= \int_0^\infty e^{-\mu_3 x}\mu_1 e^{-\mu_1 x} dx \\ &= \frac{\mu_1}{\mu_3 + \mu_1} \end{aligned}$$

(b) We wish to calculate $P(X_1 + X_2 < X_3)$. By conditioning on X_1 we get, for $x > 0$,

$$\begin{aligned} P(X_1 + X_2 < x) &= \int_{x_1=0}^{x} \left(\int_{x_2=0}^{x-x_1} \mu_2 e^{-\mu_2 x_2} dx_2 \right) \mu_1 e^{-\mu_1 x_1} dx_1 \\ &= \begin{cases} \frac{\mu_1\left(1-e^{-\mu_2 x}\right) - \mu_2\left(1-e^{-\mu_1 x}\right)}{\mu_1 - \mu_2} & \text{if } \mu_1 \neq \mu_2 \\ 1 - e^{-\mu x} - \mu x e^{-\mu x} & \text{if } \mu_1 = \mu_2 = \mu \end{cases} \end{aligned}$$

Now by conditioning on X_3 we get

$$\begin{aligned} P(X_1 + X_2 < X_3) &= \int_0^\infty P(X_1 + X_2 < X_3 \mid X_3 = x)\mu_3 e^{-\mu_3 x} dx \\ &= \int_0^\infty P(X_1 + X_2 < x)\mu_3 e^{-\mu_3 x} dx \\ &= \frac{\mu_1 \mu_2}{(\mu_1 + \mu_3)(\mu_2 + \mu_3)} \end{aligned}$$

in both cases, $\mu_1 \neq \mu_2$ and $\mu_1 = \mu_2$.

(c) Let $T = X_1 + X_2 + X_3 + W$ be the total time we spend in the system, where W is the waiting time while we wait for server 3, after being serviced by server 2. Then, since if $X_1 + X_2 \geq X_3$, $W = 0$, the

expected waiting time for server 3 is

$$\begin{aligned} E(W) &= E(W \mid X_1 + X_2 < X_3)P(X_1 + X_2 < X_3) + E(W \mid X_1 + X_2 \geq X_3)P(X_1 + X_2 \geq X_3) \\ &= E(W \mid X_1 + X_2 < X_3)P(X_1 + X_2 < X_3) \\ &= \frac{1}{\mu_3}\frac{\mu_1\mu_2}{(\mu_1+\mu_3)(\mu_2+\mu_3)} \end{aligned}$$

were we have used the memoryless property. Thus, the expected amount of time we spend in the system is

$$E(T) = E(X_1) + E(X_2) + E(X_3) + E(W) = \frac{1}{\mu_1} + \frac{1}{\mu_2} + \frac{1}{\mu_3} + \frac{1}{\mu_3}\frac{\mu_1\mu_2}{(\mu_1+\mu_3)(\mu_2+\mu_3)}$$

Another solution, that involves a bit more calculations: In what follows, we take $\mu_1 \neq \mu_2$ - similar calculations can be done in the case where $\mu_1 = \mu_2$. First we compute the density function by differentiation:

$$\begin{aligned} f_{X_1+X_2}(x) &= \frac{d}{dx}P(X_1 + X_2 < x) \\ &= \frac{\mu_1\mu_2}{\mu_1 - \mu_2}\left(e^{-\mu_2 x} - e^{-\mu_1 x}\right) \end{aligned}$$

After being served by server 1 and 2, the original customer in server 3 either has or has not been served. The time remaining is the sum of 1 or 2 exponential distributions, as per the memoryless property.

$$\begin{aligned} E(\text{time left} \mid \text{at server 3 at time } t) &= \begin{cases} \frac{1}{\mu_3} & \text{if server 3 empty, probability } 1 - e^{-\mu_3 t} \\ \frac{2}{\mu_3} & \text{otherwise, probability } e^{-\mu_3 t} \end{cases} \\ E(\text{time left} \mid \text{at server 3 at time } t) &= \frac{1}{\mu_3}\left((1 - e^{-\mu_3 t}) + 2e^{-\mu_3 t}\right) \\ &= \frac{1 + e^{-\mu_3 t}}{\mu_3} \end{aligned}$$

We then integrate this expected time left over all possible times of being at server 3.

$$\begin{aligned} E(\text{time left}) &= \int_0^\infty E(\text{time left} \mid \text{at server 3 at time } t) f_{X_1+X_2}(t)dt \\ &= \int_0^\infty \left(t + \frac{1 + e^{-\mu_3 t}}{\mu_3}\right)\frac{\mu_1\mu_2}{\mu_1 - \mu_2}\left(e^{-\mu_2 t} - e^{-\mu_1 t}\right)dt \\ &= \frac{1}{\mu_1} + \frac{1}{\mu_2} + \frac{2}{\mu_3} + \frac{\mu_2}{(\mu_1 - \mu_2)(\mu_1 + \mu_3)} - \frac{\mu_1}{(\mu_1 - \mu_2)(\mu_2 + \mu_3)} \end{aligned}$$

Problem 5

(a) The number of rainfalls in the first 5 days follows a $Poi(5 \cdot 0.2 = 1)$ distribution.

$$\begin{aligned} P(\text{empty after 5 days}) &= P(\text{no rain in the first 5 days}) \\ &= e^{-1} \\ &\approx 0.368 \end{aligned}$$

(b) The only possible scenarios for the reservoir to be empty are if it doesn't rain in the first 5 days or if exactly one rainfall of 5000 units occurs sometime in the first 5 days, but no additional rain falls in the 10 day period. Now

$$P(\text{no rain in the first 5 days}) = e^{-1}$$

$$P(\text{exactly 1 rain in the first 5 days}) = e^{-1}\frac{1^1}{1!}$$

$$P(\text{no rain in days } 6 - 10) = e^{-1}$$

giving,

$$\begin{aligned} P(\text{empty sometime in first 10 days}) &= e^{-1} + e^{-1}0.8e^{-1} \\ &\approx 0.476 \end{aligned}$$

Problem 6

Consider the time between 2005 and 2050, a 45 year span. We will make the following approximate assumptions and introduce notation:

$$\begin{aligned} N &= \#\text{ people who become lawyers by 2050} \sim Poi(500 \cdot 45) \\ I_i &= \text{indicator for whether lawyer } i \text{ is still in practice in 2050} \\ T_i &= \text{starting time of } i\text{-th lawyer} \end{aligned}$$

Assume that time each lawyer practices follows an exponential distribution with rate 1/30. Then we have

$$E(I_i \mid T_i = t_i) = e^{-\frac{45-t_i}{30}}.$$

Fix an $N = n$ and consider $E\Big(\sum_{i=1}^{n} I_i\Big)$. Even though the T_i's are distributed as an ordering of n uniform random variables, this expectation is the same as if the T_i's were n unordered uniform random variables on $[0, 45]$. For ease of calculation, we will proceed as if they were unordered. For a $U_i \sim Unif[0, 45]$,

$$P(\text{lawyers who starts at } U_i \text{ still practices in 2050}) = \int_0^{45} e^{-\frac{45-x}{30}} \frac{1}{45} dx$$

Hence,

$$\begin{aligned} E\left(\sum_{i=1}^{n} I_i\right) &= n \int_0^{45} e^{-\frac{45-x}{30}} \frac{1}{45} dx \\ &= n\frac{30}{45}\big(1 - e^{-\frac{45}{30}}\big) \end{aligned}$$

For a random N, we have

$$\begin{aligned} E\left(\sum_{i=1}^{N} I_i\right) &= E\left(\textstyle\sum_{n=1}^{\infty} n\frac{30}{45}\big(1 - e^{-\frac{45}{30}}\big) e^{-500\cdot 45} \frac{(500\cdot 45)^n}{n!}\right) \\ &= \tfrac{30}{45}\big(1 - e^{-\frac{45}{30}}\big) E(N) \\ &= \tfrac{30}{45}\big(1 - e^{-\frac{45}{30}}\big)(500 \cdot 45) \\ &\approx 11653 \end{aligned}$$

This result assumes that there are no lawyers in 2005. We may simply add $N_0 e^{-\frac{45}{30}}$ to the result if there are originally N_0 lawyers.

Appendix

Matlab-code for Problem 3 (e)

```
% ACM 116, Homework 3, fall 2005
% numerical simulations for problem 3 (e)

B = 10000;
ps = 1:10;
```

```
squaredError = zeros(B,length(ps));
% exact MSE
MSEexact = zeros(1,length(ps));

for j=1:length(ps),
  p = ps(j);
  nobs = 2*p + 1; % number of observations

  gamma = 4*2.^(-abs(0:(nobs-1)));
  Sigma = toeplitz(gamma,gamma);

  % get filter coefficients
  crosscorr = 4*2.^(-abs((-p):p)).';
  h = (Sigma+eye(nobs))\crosscorr;

  % calculate the exact MSE
  MSEexact(j) = gamma(1) - crosscorr.'*h;

  % repeat procedure B times to get an
  % estimate of the MSE
  for k=1:B,
    % using a built-in function in Matlab to
    % generate samples
    X = mvnrnd(zeros(1,nobs),Sigma);

    % add noise to get observations
    Y = X + randn(1,nobs);

    Xhat = fliplr(Y)*h;
    squaredError(k,j) = (Xhat - X(p+1)).^2;
  end
end
% numerical MSE
MSEnum = mean(squaredError);

figure(1);
tau = -10:10;
stem(tau,4*2.^(-abs(tau)));
xlabel('\tau');ylabel('\gamma_X(\tau)');

figure(2);
plot(ps,MSEnum,'o',ps,MSEexact,'x');
xlabel('p');ylabel('MSE');
legend('Numerical MSE','Exact MSE');
title(strcat(...
    'Numerical and exact MSE for different values of p.',...
    '- #simulations=',num2str(B)));
```