Statistics 519, Winter Quarter 2020

Problem Set 7

Problem 22 (2 points for each of the 3 parts). Let $\{Z_t\}$ be IID $\mathcal{N}(0, \sigma^2)$ noise, and define

$$X_t = \begin{cases} Z_t, & \text{if } t \text{ is even;} \\ \frac{Z_{t-1}^2}{\sigma\sqrt{2}} - \frac{\sigma}{\sqrt{2}}, & \text{if } t \text{ is odd.} \end{cases}$$

a. Show that $\{X_t\}$ is $\mathrm{WN}(0, \sigma^2)$.

b. What is the best linear predictor of X_{n+1} given X_n, and what is its associated MSE?

c. What is the best predictor of X_{n+1} given X_n, and what is its associated MSE? Comment briefly on how the best predictor compares with the best linear predictor.

Problem 23 (2 points). Verify the result stated at the bottom of overhead XI–19, namely, that the innovations $U_1, U_2, \ldots, U_n$ are uncorrelated, i.e., that $\mathrm{cov}\,\{U_l, U_m\} = 0$ for all l and m such that $1 \le l < m \le n$.

Problem 24 (3 points). Let $\{X_t\}$ be an invertible MA(1) process; i.e., we can write $X_t = Z_t + \theta Z_{t-1}$, where $\{Z_t\} \sim \mathrm{WN}(0, \sigma^2)$, and $|\theta| < 1$. Use the Levinson–Durbin recursions to show that, for this process, the PACF at lag h and one-step-ahead mean square error of prediction based upon h past values are given by

$$\phi_{h,h} = -\frac{(-\theta)^h(1-\theta^2)}{1-\theta^{2h+2}} \quad \text{and} \quad v_h = \sigma^2 \frac{1-\theta^{2h+4}}{1-\theta^{2h+2}}$$

(see overhead XII–16).

Problem 25 (3 points for each of the 3 parts). Let $X_t = Z_{1,t} + \frac{2}{3} Z_{1,t-1}$, where $\{Z_{1,t}\} \sim \mathrm{WN}(0, \frac{9}{13})$, and let $Y_t = \frac{6}{13} Y_{t-1} + Z_{2,t}$, where $\{Z_{2,t}\} \sim \mathrm{WN}(0, \frac{133}{169})$.

a. Using appropriate formulae presented in class, determine the ACVF and ACF for both processes at lags $h = 0$, 1, 2, 3 and 4.

b. For the MA(1) process $\{X_t\}$, carry out the Levinson–Durbin recursions and innovations algorithm for $n = 1, 2, 3$ and 4, and report your results for $\{v_0, v_1, \ldots, v_4\}$, $\{\phi_{n,j} : n =$

$1,\ldots,4,\ \ j=1,\ldots,n\}$ and $\{\theta_{n,j} : n=1,\ldots,4,\ \ j=1,\ldots,n\}$ in the form of a table resembling the following display:

$$
\begin{array}{l|llll|llll}
v_0 = ??? & & & & & & & & \\
v_1 = ??? & \phi_{1,1} = ??? & & & & \theta_{1,1} = ??? & & & \\
v_2 = ??? & \phi_{2,1} = ??? & \phi_{2,2} = ??? & & & \theta_{2,1} = ??? & \theta_{2,2} = ??? & & \\
v_3 = ??? & \phi_{3,1} = ??? & \phi_{3,2} = ??? & \phi_{3,3} = ??? & & \theta_{3,1} = ??? & \theta_{3,2} = ??? & \theta_{3,3} = ??? & \\
v_4 = ??? & \phi_{4,1} = ??? & \phi_{4,2} = ??? & \phi_{4,3} = ??? & \phi_{4,4} = ??? & \theta_{4,1} = ??? & \theta_{4,2} = ??? & \theta_{4,3} = ??? & \theta_{4,4} = ???
\end{array}
$$

Repeat all the above for the AR(1) process $\{Y_t\}$. Comment briefly about any interesting patterns you notice in the displays. (Feel free to use the `R` functions `LD.recursions` and `innovations.algorithm` available via the R Code page of the course Web site.)

c. Determine the variances of $\widehat{X}_1, \ldots, \widehat{X}_5$, i.e., the best linear predictors for the MA(1) process based upon 0, 1, 2, 3 and 4 past values. Compare these to the variances of $\widehat{X}_{1|0,\infty}, \ldots, \widehat{X}_{5|4,\infty}$, i.e., the best linear predictors based upon the infinite past (see overhead XI–43). Repeat all of the above for the AR(1) process.

Problem 26 (1 point for each of the 5 parts). Plot the sample PACF at lags $h = 1,\ldots,20$ for the five time series listed below, along with 95% confidence bounds for the unknown true PACF based on the null hypothesis of white noise (see course overhead XII–11). Comment on what the sample PACF tells you about the suitability of an AR model for each of the time series.

a. El Niño–Southern Oscillation (ENSO) index (overhead I–30)

b. residuals $\{r_t\}$ from Lake Huron level time series (overhead III–5)

c. residuals $\{r_t\}$ from accidental deaths time series (overhead III–94)

d. wind speed time series (overhead V–12)

e. annual North Pacific Index (NPI) time series. This series is an area-weighted sea level pressure over latitudes 30° N to 65° N and longitudes 160° E to 140° W and over November to March for each year from 1900 to 2019. This time series is of interest in the study of on-going changes in the arctic climate.

Note: all five time series can be accessed via the Data page of the course Web site (items 8, 10, 13, 14 and 16 in the list on that page) or in `R` using the command

```
scan("http://faculty.washington.edu/dbp/s519/Data/xxx.txt")
```

by replacing '`xxx`' with '`ENSO`' to get series a; with '`rs-LH`' to get series b; with '`rs-deaths`' to get series c; with '`wind`' to get series d; and with '`NPI`' to get series e.

Solutions are due Friday, February 28, at the beginning of the class.