Stat 551 Exam 1

October 25, 2012

There are 14 parts on this exam. I will score every part out of 10 points and take your best 10 of 14 scores. (Budget your time accordingly and find things you can do.)
1. Two independent mean 0 second order stationary time series \( \{u_t\} \) and \( \{v_t\} \) have respective autocovariance functions \( \gamma_1(s) = \sigma_1^2 \rho_1(s) \) and \( \gamma_2(s) = \sigma_2^2 \rho_2(s) \).

a) Is the series \( \{w_t\} \) defined by \( w_t = u_t v_t \) second order stationary? If so, what is its autocorrelation function?

\[
E W_t = E(u_t v_t) = E(u_t)E(v_t) = 0 \cdot 0 = 0 \quad \forall t
\]

\[
\text{Cov}(W_{t+s}, W_t) = E(W_{t+s}W_t) = E(u_{t+s}v_{t+s}u_t v_t) = E(u_{t+s}u_t)(v_{t+s}v_t)
\]

\[
= \gamma_1(s) \gamma_2(s)
\]

This is independent of \( t \), so the series \( \{w_t\} \) is second order stationary. The autocorrelation function is above, so the autocorrelation function for \( \{w_t\} \) is

\[
\rho(s) = \frac{\gamma_1(s) \gamma_2(s)}{\gamma_1(0) \gamma_2(0)} = \rho_1(s) \rho_2(s)
\]

b) For constants \( a \) and \( b \), is the series \( \{w_t\} \) defined by \( w_t = au_t + bv_t \) second order stationary? If so, what is its autocorrelation function?

\[
E W_t = E(a u_t + b v_t) = a E u_t + b E v_t = a \cdot 0 + b \cdot 0 = 0 \quad \forall t
\]

\[
\text{Cov}(W_{t+s}, W_t) = E(a u_{t+s} + b v_{t+s}, a u_t + b v_t)
\]

\[
= E(a^2 u_{t+s} u_t + ab u_{t+s} v_t + ab v_{t+s} u_t + b^2 v_{t+s} v_t)
\]

\[
= a^2 \gamma_1(s) + b^2 \gamma_2(s)
\]

Which is independent of \( t \), so the process is second order stationary with autocovariance function above. The autocorrelation function is then

\[
\rho(s) = \frac{a^2 \gamma_1(s) + b^2 \gamma_2(s)}{a^2 \gamma_1(0) + b^2 \gamma_2(0)} = \frac{a^2 \sigma_1^2 \rho_1(s) + b^2 \sigma_2^2 \rho_2(s)}{a^2 \sigma_1^2 + b^2 \sigma_2^2}
\]

\[
= \frac{a^2 \rho_1(s)}{a^2 \sigma_1^2 + b^2 \sigma_2^2} + \frac{b^2 \rho_2(s)}{a^2 \sigma_1^2 + b^2 \sigma_2^2}
\]
2. Suppose that the time series \( \{ y_t \} \) is second order stationary with mean 0, variance 1, and has the autocorrelation function \( \rho(s) = \exp(-|s|) \).

a) Write out the covariance matrix for \( Y_4 = (y_1, y_2, y_3, y_4)' \) (call it \( \Sigma_4 \)) and give a matrix expression for the covariance matrix of \( \nabla Y_4 = (y_2 - y_1, y_3 - y_2, y_4 - y_3) \). (You do not need to simplify this latter, but all relevant numerical values need to be plugged in.)

\[
\Sigma_4 = \begin{pmatrix}
1 & e^{-1} & e^{-2} & e^{-3} \\
1 & e^{-1} & e^{-2} & e^{-3} \\
e^{-2} & e^{-1} & e^{-2} & e^{-3} \\
e^{-3} & e^{-2} & e^{-3} & e^{-3}
\end{pmatrix}
\]

\[
\text{Cov} \nabla Y_4 = \begin{pmatrix}
-1 & 1 & 0 & 0 \\
0 & -1 & 1 & 0 \\
0 & 0 & -1 & 1 \\
0 & 0 & 0 & 1
\end{pmatrix}
\]

b) Give a matrix expression for a 95% prediction interval for \( y_5 \) based on \( Y_4 \) (defined in part a)). Except for \( Y_4 \) and \( \Sigma_4 \) (for which you may use these symbols) all other matrices and vectors need to have numbers plugged, but you need NOT simplify/do matrix computations.

\[
(\begin{pmatrix}
y_4 \\
y_6
\end{pmatrix}) \text{ has covariance matrix}
\]

\[
\Sigma_4 = \begin{pmatrix}
\Sigma_4 & e^{-5} & e^{-4} & e^{-3} \\
e^{-5} & e^{-4} & e^{-3} & e^{-2} \\
e^{-4} & e^{-3} & e^{-2} & e^{-3} \\
e^{-3} & e^{-2} & e^{-3} & e^{-3}
\end{pmatrix}
\]

So the Gaussian conditional mean of \( y_6 | Y_4 \) is \( \hat{y}_6 = 0 + (e^{-5}, e^{-4}, e^{-3}, e^{-2}) \Sigma_4^{-1} Y_4 \)

and the conditional variance is \( \nu = 1 - (e^{-5}, e^{-4}, e^{-3}, e^{-2}) \Sigma_4^{-1} \begin{pmatrix}
e^{-5} \\
e^{-4} \\
e^{-3} \\
e^{-2}
\end{pmatrix} \)

and the prediction interval is

\( \hat{y}_6 \pm 1.96 \sqrt{\nu} \)
3. The function \( \rho_c(s) = \exp(-cs) \) is a valid autocorrelation function for any \( c > 0 \). Suppose that the time series \( \{y_t\} \) is second order stationary with mean 0 and variance \( \sigma^2 \), but neither \( c \) nor \( \sigma^2 \) is known. Data available are \( \mathbf{Y}_4 = (y_1, y_2, y_3, y_4)' \). Write out a function of \( c \) and \( \sigma^2 \) (say \( L(c, \sigma^2) \)) that might plausibly be maximized in order to produce good estimates of these two model parameters. You should use matrix notation.

One might use \( L(c, \sigma^2) \) the Gaussian log likelihood. That is, with

\[
\sum_{i=1}^{4} \frac{c_i \sigma^2}{2} = \sigma^2 \begin{pmatrix} 1 & e^{-c} & e^{-2c} & e^{-3c} \\ e^{-c} & 1 & e^{-c} & e^{-2c} \\ e^{-2c} & e^{-c} & 1 & e^{-c} \\ e^{-3c} & e^{-2c} & e^{-c} & 1 \end{pmatrix}
\]

one might choose to optimize

\[
L(c, \sigma^2) = -\frac{4}{2} \log \left| \text{det} \left( \sum_{i=1}^{4} \frac{c_i \sigma^2}{2} \right) \right| - \frac{1}{2} \mathbf{Y}_4 \left( \sum_{i=1}^{4} \frac{c_i \sigma^2}{2} \right)^{-1} \mathbf{Y}_4
\]

4. Write in both operator notation and in difference equation notation (involving series values like \( y_t \)) the form of a model for \( \mathbf{Y} \) where a mean 0 SARIMA \( (1,1,1) \times (0,1,0)_9 \) noise process obscures a transfer function mean based on a predictor series \( \mathbf{x} \) involving a time delay of \( r = 4 \), a "numerator order" of \( m = 1 \) (a single lag), and "denominator order" of \( l = 0 \) (no lags).

\[
(\mathbf{I} - \phi \mathbf{B})(\mathbf{I} - \theta \mathbf{B}^4)(\mathbf{I} - \beta \mathbf{B}^4 \mathbf{x}) = (\mathbf{I} + \theta \mathbf{B}) \mathbf{Y}
\]

\[
\text{Since } (\mathbf{I} - \phi \mathbf{B})(\mathbf{I} - \theta \mathbf{B}^4)(\mathbf{I} - \beta \mathbf{B}^4 \mathbf{x}) = (\mathbf{I} - (1 + \phi) \mathbf{B} + \mathbf{B}^2)(\mathbf{I} - \beta \mathbf{B}^4 \mathbf{x})
\]

\[
= \mathbf{I} - (1 + \phi) \mathbf{B} + \mathbf{B}^2 - \mathbf{B}^4 + (1 + \phi) \mathbf{B}^5 - \mathbf{B}^6
\]

in other notation this is

\[
(\mathbf{y}_t - (1 + \phi) \mathbf{y}_{t-1} + \mathbf{y}_{t-2} - \mathbf{y}_{t-4} + (1 + \phi) \mathbf{y}_{t-5} - \mathbf{y}_{t-6}) - \beta (\mathbf{x}_{t-4} - (1 + \phi) \mathbf{x}_{t-5} + \mathbf{x}_{t-6} - \mathbf{x}_{t-8} + (1 + \phi) \mathbf{x}_{t-9} - \mathbf{x}_{t-10}) = \mathbf{e}_t + \theta \mathbf{e}_{t-1}
\]
5. Consider the simplest of all ARIMA models, one where $\mathbf{D}Y$ is mean 0 variance $\sigma^2$ white noise. (This, of course, gives a mean and covariance matrix for $\mathbf{D}Y$. ) Suppose that (in more or less standard time series fashion) for forecasting purposes one assumes that $Y_i$ is uncorrelated with $\mathbf{D}Y_n$. In order to finish making a complete joint model for $Y_n$ one needs to make some assumption about the marginal of $Y_i$. We'll suppose that $\text{E}Y_i = \mu$ and $\text{Var}Y_i = \eta^2$.

a) What are the mean vector and covariance matrix of $(y_1,(\mathbf{D}Y_n)',y_{n+1})$ under the above assumptions? (Note that $y_{n+1} = y_1 + \sum_{t=2}^{n+1}(y_t - y_{t-1})$. Writing the covariance matrix in partitioned form where the upper left block is $n \times n$ should make this quickly doable.)

$$
\begin{align*}
\text{Var } W_n &= \begin{pmatrix}
\mu^2 & \eta^2 & \eta^2 & \cdots & \eta^2 \\
\eta^2 & \sigma_1^2 & \sigma_2^2 & \cdots & \sigma_n^2 \\
\eta^2 & \sigma_2^2 & \sigma_3^2 & \cdots & \sigma_n^2 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\eta^2 & \sigma_{n-1}^2 & \sigma_n^2 & \cdots & \sigma_n^2 + \eta^2
\end{pmatrix} \\
\text{E } W_n &= (\mu,0,0,\ldots,0,\mu)
\end{align*}
$$

b) Under these assumptions, what is the best linear predictor of $y_{n+1}$ given $(y_1,(\mathbf{D}Y_n)')$ (or equivalently $Y_n$)? (Write this in matrix form and simplify.)

This is just the Gaussian conditional mean of the last entry of $(y_1,(\mathbf{D}Y_n)',y_{n+1})'$. This is

$$
\begin{align*}
\mathbf{m} &= (\eta^2, \sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2) \text{ diag } \left( \frac{1}{\eta^2}, \frac{1}{\sigma_1^2}, \frac{1}{\sigma_2^2}, \ldots, \frac{1}{\sigma_n^2} \right) \\
&= \mathbf{m} + \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} y_1 - \mu \\ y_2 - y_1 - \eta \\ y_3 - y_2 - \eta \\ \vdots \\ y_n - y_{n-1} - \eta \end{pmatrix} \\
&= \mathbf{y}_n
\end{align*}
$$
6. Consider the ARMA form \( \Phi(B)Y = \Theta \varepsilon \) for \( \varepsilon \) mean 0 white noise, where
\[
\Phi(B) = I - .25B^2 \quad \text{and} \quad \Theta(B) = I - .5B
\]
This apparent ARMA(2,1) form is not the simplest possible ARMA form for such a \( Y \). Indeed the standard argument employed to show an ARMA model is both causal and invertible cannot be applied to the present representation. However, there is another representation that can be used.

a) Argue very carefully that there is another representation \( \Phi^*(B)Y = \Theta^* \varepsilon \) that reveals that \( Y \) satisfying the original equation is actually a causal and invertible ARMA process of lower order than (2,1). (You will want to expand \( \Phi(B) \) as the "product" of two "backshift monomial" operators.)

Note that \( \Phi(B) = I - .25B^2 = (I - .5B)(I + .5B) \) so that the original equation is
\[
(I - .5B)(I + .5B)Y = (I - .5B)\varepsilon
\]
Now the operator \((I - .5B)\) is invertible, i.e. \( \exists \) a causal linear filter \( Z \) such that \( Z(I - .5B) = I \). So operating on both sides of the original equation with \( Z \), we have
\[
(I + .5B)Y = \varepsilon
\]
and \( Y \) is AR(1) with AR parameter \( \phi = -.5 \).

b) In light of a), what are
- the autocorrelation function,
- the partial autocorrelation function, and
- the inverse autocorrelation function
for the process \( Y \)?

\[
\rho(s) = (-.5)^{|s|}, \quad \alpha(s) = \begin{cases} 0 & \text{if } |s| > 1 \\ \frac{s}{1 + .5s} & \text{if } |s| = 1 \\ 1 & \text{if } s = 0 \end{cases}
\]

\( \gamma(1) = \frac{\gamma(0)}{\gamma(0)} = \rho(1) = -.5 \)

The inverse autocorrelation function is the autocorrelation function for a MA(1) process with MA parameter \( \theta \), namely
\[
\rho_{MA}(s) = \begin{cases} .5 & \text{if } s = 0 \\ \frac{.5}{1 + .5s} & \text{if } s = 1 \\ 0 & \text{otherwise} \end{cases}
\]
7. Consider the mean 0 ARMA(1, 2) series \( \{y_t\} \) satisfying the difference equation

\[ y_t - .5y_{t-1} = .1\varepsilon_{t-2} - .2\varepsilon_{t-1} + \varepsilon_t \]

a) Find the first 4 coefficients \( \psi_j \) (namely \( \psi_0, \psi_1, \psi_2, \psi_3 \)) in the ("MA(\( \infty \))") representation

\[ y_t = \sum_{s=0}^{\infty} \psi_s \varepsilon_{t-s}. \]

Write \( (1 - .5z) = \phi(z) \), \( (1 - 2z + 1z^2) = \Theta(z) \)

So \( \phi^{-1}(z) = \sum_{j=0}^{\infty} (\cdot5)^j z^j \) and \( \phi^{-1}(z) \Theta(z) = \)

\[ (1 + .5z + 25z^2 + 125z^3 + \cdots) (1 - 2z + 1z^2) \]

\[ = 1 + .5z - 2z + (.5)(-2)z^2 + 25z^2 + 1z^2 \]

\[ + 125z^3 + (.25)(-2)z^3 + (.5)(.1)z^3 + \cdots \]

\[ = 1 + .3z + 25z^2 + 125z^3 \]

\( \psi_0 = 1 \), \( \psi_1 = .3 \), \( \psi_2 = 25 \), \( \psi_3 = .125 \)

b) Use your answer to a) and find the values of the lag 1 and 2 autocorrelations for this model, \( \rho(1) \) and \( \rho(2) \).

Use the recursions (since \( p = 1 \) and \( q = 2 \) here):

\[ \gamma(0) = \phi_1 \gamma(-1) = \sigma^2 (\psi_0 + \theta_1 \psi_1 + \theta_2 \psi_2) \]

\[ \gamma(0) - .5\gamma(1) = \sigma^2 (1 + (-.2)(.3) + (.1)(.25)) \]

\[ \sigma^2 (1 - .5\rho(1)) = \sigma^2 (.965) \]

\[ -.5\rho(1) = -.035 \quad \text{i.e.} \quad \rho(1) = .07 \]

\[ \gamma(2) - \phi_1 \gamma(1) = \sigma^2 (\theta_2 \psi_0) \]

\[ \sigma^2 (\rho(2) - .5(.07)) = \sigma^2 (.1(1)) \]

\[ \rho(2) = .1 + (.5)(.07) = .135 \]
8. Below are pieces of more or less standard advice in applied time series analysis. Explain (in 100 words or less each) their origins. (Say why they are correct.)

a) When seasonality of period \( s \) is likely, one should try seasonal differencing at lag \( s \) before resorting to applications of "ordinary" differencing.

Seasonal differencing will remove a linear trend from data in addition to dealing with seasonality. It may not be necessary to do ordinary differencing after seasonal differencing. (On the other hand, ordinary differencing will not deal with seasonality.)

b) If upon fitting an ARMA \((p,q)\) model, the fitted \( p \)th order complex polynomial \( \phi(z) \) has a factor \((1 - \hat{\alpha}z)\) where \( \hat{\alpha} \approx 1 \) (say \( \phi(z) \approx (1 - z)\phi^*(z) \) for \( \phi^*(z) \) a \((p-1)\)st order complex polynomial with real coefficients), then the original series has not been adequately differenced.

\[ \phi(z) \text{ of the form } (1 - z)\phi^*(z) \text{ means that } \]

\[ \Phi(B)Y = \Phi'(B)(1-B)Y = \Phi'(B)\Delta Y \]

which says that \( Y \) is ARIMA \((p-1,1,0)\) and while \( \Phi'(B) \) is certainly not invertible it's at least possible that \( \Phi'(B) \) is invertible (so that the model for \( \Delta Y \) is causal). So while the model for \( Y \) is not simple/causal/tractable, the corresponding one for \( \Delta Y \) may be so.