Solutions to Exam 2

1(a). Given the definition of $P$, we need to consider the behaviour of $\hat{\sigma}_T^2(p)$ for $p = 0, 1$ when the true dgp is actually a stationary $AR(1)$ process.

If $p = 0$ then the assumed model is $AR(0)$ and hence white noise. Therefore, the residual from an $AR(0)$ model is just the series itself, $v_t$. Therefore, we have

$$\hat{\sigma}_T^2(0) = T^{-1} \sum_{t=1}^T v_t^2$$

Since $\{v_t\}$ is a covariance stationary process and ergodic for the second moments, it follows that

$$\hat{\sigma}_T^2(0) \xrightarrow{p} \gamma_0 = \frac{\sigma^2}{1-\theta^2} = \sigma^2 + \frac{\theta^2 \sigma^2}{(1-\theta^2)}$$

(1)

If $p = 1$ then

$$\hat{\sigma}_T^2(1) = T^{-1} \sum_{t=2}^T (v_t - \hat{\theta}_1(1)v_{t-1})^2 \xrightarrow{p} \sigma^2$$

(2)

where $\hat{\theta}_1(1) = \sum_{t=2}^T v_{t-1}v_t / \sum_{t=2}^T v_{t-1}^2$.

Define $SIC_T(p) = \ln[\hat{\sigma}_T^2(p)] + pln[T]/T$. Given (1)-(2) and $\lim_{T \to \infty} \ln[T]/T = 0$, it follows that

$$SIC_T(0) \xrightarrow{p} \sigma^2 + \frac{\theta^2 \sigma^2}{(1-\theta^2)} > \sigma^2$$

(3)

$$SIC_T(1) \xrightarrow{p} \sigma^2$$

(4)

where the inequality in (3) uses the fact that $|\theta| < 1$. It follows from (3)-(4) that $\lim_{T \to \infty} P(\hat{p}_{SIC} = 1) = 1$.

1(b). For a given value of $p$, Akaike’s information criterion is $AIC(p)$ where AIC is

$$AIC(p) = \ln[\hat{\sigma}_T^2(p)] + \frac{2p}{T}$$

and the estimated order is $\hat{p}_{AIC} = \text{argmin}_{p \in P} \ AIC(p)$.

SIC is consistent, that is $\lim_{T \to \infty} P(\hat{p}_{SIC} = 1) = 1$. AIC is not consistent but has a zero probability of underfitting in the limit, $\lim_{T \to \infty} P(\hat{p}_{AIC} = 0) = 0$ but $\lim_{T \to \infty} P(\hat{p}_{AIC} > 1) > 0$ - recall that in this question the true order is one.
1(c). Consider the AR(p) model

\[ v_t = \theta_1 v_{t-1} + \theta_2 v_{t-2} + \ldots + \theta_p v_{t-p} + w_t \]

If it is assumed that \( w_t \sim IN(0, \sigma^2) \) then the Likelihood ratio test for \( H_0: \theta_n = \theta_{n+1} = \ldots = \theta_p = 0 \) is

\[ LR_T = T \ln(\hat{\sigma}^2(n)/\hat{\sigma}^2(p)) \]

where \( \hat{\sigma}^2(m) = RSS(m)/T \) and RSS(m) is the residual sum of squares from the OLS regression of \( v_t \) on \( v_{t-1}, v_{t-2}, \ldots, v_{t-m} \). Therefore, we have

\[ T \{ SIC(n) - SIC(p) \} = LR_T + (n - p) \ln(T) \]

2(a). Consider \( A_m = (I_k - \Theta)(I_k + \Theta + \Theta^2 + \ldots + \Theta^m) \). Multiplying out, it can be seen that

\[ A_m = I_k + \Theta - \Theta + \Theta^2 - \Theta^2 \ldots - \Theta^{m+1} \]

Using the Jordan decomposition for \( \Theta^{m+1} \) in (5) we have

\[ A_m = I_k - C^{-1}\Lambda^{m+1}C \]

where \( C \) is the matrix of eigenvectors of \( \Theta \) and \( \Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_p) \). It follows from (6) that

\[ \lim_{m \to \infty} A_m = I_k - C^{-1}\Lambda^{m+1}C \]

Now \( \Lambda^{m+1} = diag(\lambda_1^{m+1}, \lambda_2^{m+1}, \ldots, \lambda_p) \) and since \( |\lambda_i| < 1 \) for \( i = 1, 2, \ldots, p \), it follows that \( \lim_{m \to \infty} \Lambda^{m+1} = 0_{k \times k} \), where \( 0_{k \times k} \) is the \( k \times k \) null matrix. Therefore, it follows from (7) that \( \lim_{m \to \infty} A_m = I_k \) and so

\[ I_k = \lim_{m \to \infty} (I_k - \Theta)(I_k + \Theta + \Theta^2 + \ldots + \Theta^m) \]

\[ I_k = (I_k - \Theta) \lim_{m \to \infty} (I_k + \Theta + \Theta^2 + \ldots + \Theta^m) \]

It follows from (8) and the nonsingularity of \( I_k - \Theta \) that

\[ (I_k - \Theta)^{-1} = I_2 + \lim_{m \to \infty} \sum_{i=1}^{m} \Theta^i \]

2(b). Using the VMA(\( \infty \)) representation given in the question and \( E[w_t] = 0 \), it follows that

\[ Var|v_t| = E[\left( \sum_{i=0}^{\infty} \Theta^i w_{t-i} \right) \left( \sum_{j=0}^{\infty} \Theta^j w_{t-j} \right)] \]

\[ = E[\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \Theta^i w_{t-i} \Theta^j w_{t-j}] \]

\[ = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \Theta^i E[w_{t-i} w_{t-j}] \Theta^j \]

\[ = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \Theta^i \sigma^2 \]

\[ = \sigma^2 \sum_{i=0}^{\infty} \Theta^i \]

\[ = \sigma^2 / (1 - \Theta) \]
Now \( w_t \sim i.i.d.(0, \Omega) \) implies

\[
E[w_{t-i}w_{t-j}] = \begin{cases} \Omega, & \text{for } i = j \\ 0_{k \times k}, & \text{for } i \neq j \end{cases}
\] (9)

\[
= \sum_{i=1}^{\infty} \Theta^i \Omega \Theta^i.
\] (10)

where \( 0_{k \times k} \) is the \( k \times k \) null matrix. Therefore, it follows that \( \text{Var}[v_t] = \Omega + \sum_{i=1}^{\infty} \Theta^i \Omega \Theta^i \).

2(c). If \( v_t = \sum_{i=0}^{\infty} \Psi_i w_{t-i} \) then the impulse response function for \( v_1 \) in response to \( w_2 \) is

\[
I_{1,2}(s) = \{\Psi_s\}_{1,2}
\] (11)

where \( \{.,.\}_{i,j} \) denotes the \( i-j \)th element of the matrix within the parentheses. For the stationary \( VAR(1) \) model, \( \Psi_s = \Theta^s \). Since \( v_t \sim VAR(1), \) \( v_2 \) does not Granger cause \( v_1 \) if and only if \( E[v_{1,t}|v_{t-1},v_{t-2},...] \) does not depend on \( \{v_{2,t-i}, i = 1, 2, \ldots\} \) and so \( \Theta \) takes the form

\[
\Theta = \begin{bmatrix} \theta_{1,1} & 0 \\ \theta_{2,1} & \theta_{2,2} \end{bmatrix}
\] (12)

It follows from (12) that

\[
\Theta^s = \begin{bmatrix} \theta_{1,1}^s & 0 \\ m_s & \theta_{2,2}^s \end{bmatrix}
\] (13)

for some \( m_s \). Combining (14) and (13), it follows that

\[
I_{1,2}(s) = 0, \text{ for all } s > 0
\] (14)

3