MTH442 Assignment 3 Solutions

Jiyanshu Dhaka

Q1:

Claim: If $\gamma(0) > 0$ and $\gamma(h) \to 0$ as $h \to \infty$, then Γ_T is positive definite.

Proof by contradiction: Assume there is a Γ_T that is singular. Because $\gamma(0) > 0$, $\Gamma_1 = \{\gamma(0)\}$ is non-singular. Thus, there is an $r \geq 1$ such that Γ_r is non-singular. Consider the ordered sequence $\Gamma_1, \Gamma_2, \ldots$ and suppose Γ_{r+1} is the first singular Γ_T in the sequence. Then φ_{r+1} is a linear combination of $\varphi = (\varphi_1, \ldots, \varphi_r)'$, say, $\varphi_{r+1} = b'\varphi$ where $b = (b_1, \ldots, b_r)'$. Because of stationarity, it must also be true that $\varphi_{r+h+1} = b'\varphi_h$, where $\varphi_h = (\varphi_h, \ldots, \varphi_{r+h-1})'$ for all $h \geq 1$. This means that for any $n \geq r+1$, φ_n is a linear combination of $\varphi_1, \ldots, \varphi_r$, i.e., $\varphi_n = b_n'\varphi$ where $b_n = (b_{n1}, \ldots, b_{nr})'$. Thus, $\gamma(0) = \mathrm{var}(\varphi_n) = b_n'\Gamma_r b_n = b_n' Q \Lambda Q' b_n$ where QQ' is the identity matrix and $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_r)$ is the diagonal matrix of the positive eigenvalues $(0 < \lambda_1 \leq \cdots \leq \lambda_r)$ of Γ_r. From this result, we conclude

$$\gamma(0) \geq \lambda_1 b_n' Q Q' b_n = \lambda_1 \sum_{j=1}^{r} b_{nj}^2;$$

this shows that for each j, b_{nj} is bounded in n. In addition, $\gamma(0) = \mathrm{cov}(\varphi_n, \varphi_n) = \mathrm{cov}(\varphi_n, b_n'\varphi)$, from which it follows that

$$0 < \gamma(0) \leq \sum_{j=1}^{r} |b_{nj}||\gamma(n-j)|.$$

From this inequality, it is seen that because the b_{nj} are bounded, it is not possible to have $\gamma(0) > 0$ and $\gamma(h) \to 0$ as $h \to \infty$, which contradicts the assumption that Γ_T is singular.

Thus, Γ_T must be positive definite.

Solution to Q4

We start by taking the prediction equations with $n = h$ and dividing both sides by $\gamma(0)$, to obtain:

$$R_h \phi_h = \rho_h,$$

where R_h is the $h \times h$ autocorrelation matrix, and $\rho_h = (\rho(1), \ldots, \rho(h))$ is the vector of lagged autocorrelations. Partition the equation as in the hint with $\phi_h = \begin{pmatrix} \phi_{h-1} \\ \phi_{hh} \end{pmatrix}$, and solve:

$$\begin{pmatrix} R_{h-1} & \tilde{\rho}_{h-1} \\ \tilde{\rho}_{h-1}^\top & 1 \end{pmatrix} \begin{pmatrix} \phi_{h-1} \\ \phi_{hh} \end{pmatrix} = \begin{pmatrix} \rho_{h-1} \\ \rho(h) \end{pmatrix},$$

which gives us the system:

$$R_{h-1}\phi_{h-1} + \tilde{\rho}_{h-1}\phi_{hh} = \rho_{h-1}, \quad (1)$$
$$\tilde{\rho}_{h-1}^\top \phi_{h-1} + \phi_{hh} = \rho(h). \quad (2)$$

First, solve equation (1) for ϕ_{h-1}:

$$\phi_{h-1} = R_{h-1}^{-1}(\rho_{h-1} - \tilde{\rho}_{h-1}\phi_{hh}).$$

Substitute this expression into equation (2) and solve for ϕ_{hh}:

$$\phi_{hh} = \frac{\rho(h) - \tilde{\rho}_{h-1}^\top R_{h-1}^{-1} \rho_{h-1}}{1 - \tilde{\rho}_{h-1}^\top R_{h-1}^{-1} \tilde{\rho}_{h-1}}. \quad (3)$$

Next, we must show that the PACF, ϕ_{hh}, can be written in the form of equation (3). Define the residuals $\epsilon_t = X_t - \hat{X}_t$ and $\delta_{t-h} = X_{t-h} - \hat{X}_{t-h}$, where $\hat{X}_t$ and $\hat{X}_{t-h}$ are the best linear predictors (BLP) of X_t and X_{t-h}, respectively. The PACF is defined as:

$$\phi_{hh} = \frac{E[\epsilon_t \delta_{t-h}]}{\sqrt{E[\epsilon_t^2] E[\delta_{t-h}^2]}}.$$

Now, we calculate the covariance and variances for the residuals. Using the regression results, we have:

$$\epsilon_t = X_t - \gamma_h^\top \Gamma_h^{-1} X_{t-h}, \quad \delta_{t-h} = X_{t-h} - \tilde{\gamma}_h^\top \Gamma_h^{-1} X_{t-h}.$$

From this, the covariance is:

$$E[\epsilon_t \delta_{t-h}] = \text{cov}(\epsilon_t, \delta_{t-h}) = \gamma(h) - \tilde{\gamma}_{h-1}^\top \Gamma_{h-1}^{-1} \gamma_{h-1}.$$

The variances are calculated similarly:

$$E[\epsilon_t^2] = \gamma(0) - \tilde{\gamma}_{h-1}^\top \Gamma_{h-1}^{-1} \tilde{\gamma}_{h-1}, \quad E[\delta_{t-h}^2] = \gamma(0) - \tilde{\gamma}_{h-1}^\top \Gamma_{h-1}^{-1} \tilde{\gamma}_{h-1}.$$

Substitute these results into the PACF formula to get:

$$\phi_{hh} = \frac{\gamma(h) - \tilde{\gamma}_{h-1}^\top \Gamma_{h-1}^{-1} \gamma_{h-1}}{\sqrt{(\gamma(0) - \tilde{\gamma}_{h-1}^\top \Gamma_{h-1}^{-1} \tilde{\gamma}_{h-1})(\gamma(0) - \tilde{\gamma}_{h-1}^\top \Gamma_{h-1}^{-1} \tilde{\gamma}_{h-1})}}.$$

This proves the result upon factoring out $\gamma(0)$ in the numerator and denominator, giving us:

$$\phi_{hh} = \frac{\rho(h) - \tilde{\rho}_{h-1}^\top R_{h-1}^{-1} \rho_{h-1}}{1 - \tilde{\rho}_{h-1}^\top R_{h-1}^{-1} \tilde{\rho}_{h-1}}.$$

Solution to Q5

We are tasked with finding the linear prediction function $g(x) = a + bx$ that minimizes the mean squared error (MSE):

$$\text{MSE} = E[(Y - g(X))^2],$$

where X and Y are jointly distributed random variables with density function $f(x, y)$. To minimize the MSE, we need to determine the values of a and b that satisfy this condition.

Step 1: Write the prediction equation We assume that $g(x) = a + bx$. Using the prediction equations, we know that $g(x)$ must satisfy:

$$E[Y - g(X)] = 0 \quad \text{and} \quad E[(Y - g(X))X] = 0.$$

These two conditions can be rewritten as:

$$E[Y] = E[a + bx] \quad \text{and} \quad E[XY] = E[X(a + bx)].$$

Step 2: Solve for a and b From the first equation, we have:

$$E[Y] = a + bE[X].$$

Let's assume $E[X] = 0$ and $E[Y] = 1$. Thus, the first equation simplifies to:

$$a = 1.$$

From the second equation, we have:

$$E[XY] = aE[X] + bE[X^2].$$

Given that $E[X] = 0$ and $E[X^2] = 1$, this equation simplifies to:

$$bE[X^2] = E[XY].$$

If we assume $E[XY] = 0$, we get:

$$b = 0.$$

Step 3: Calculate the MSE The prediction function is now $g(x) = 1 + 0 \cdot x = 1$. The MSE becomes:

$$\text{MSE} = E[(Y - g(X))^2] = E[(Y - 1)^2].$$

Now, expanding the square:

$$\text{MSE} = E[Y^2 - 2Y + 1] = E[Y^2] - 2E[Y] + E[1].$$

Given that $E[Y^2] = 4$ and $E[Y] = 1$, we substitute these values into the equation:

$$\text{MSE} = 4 - 2(1) + 1 = 3.$$

Conclusion We have shown that the optimal linear predictor is $g(x) = 1$, with $a = 1$, $b = 0$, and the MSE is 3. This result implies that, in this case, the best linear predictor does not depend on x, and the error associated with the prediction is constant.