Chapter 3, Shumway and Stoffer

3.3 Polynomial zeros

Use R to find the zeros of the polynomial (i.e., use the function polyroot).

(a) \( x_t = 0.8x_{t-1} - 0.15x_{t-2} + w_t - 0.3w_{t-1} \)

\[
\phi(z) = 1 - 0.8z + 0.15z^2 \Rightarrow z_1 = 2, z_2 = 1/0.3
\]
\[
\theta(z) = 1 - 0.3z \quad z_1 = 1/0.3
\]

The process is causal and invertible as the zeros lie outside the unit circle. The zero of the moving average cancels one of those for the AR, reducing to an AR(1) process.

(b) \( x_t = x_{t-1} - 0.5x_{t-2} + w_t - w_{t-1} \)

\[
\phi(z) = 1 - z + 0.5z^2 \Rightarrow z_1 = 1 - \iota, z_2 = 1 + \iota
\]
\[
\theta(z) = 1 - z \quad z_1 = 1
\]

The process is causal, but not invertible.

3.7 ARMA(1,1)

The point of this exercise is to see the qualitative differences among the autocorrelations and partial autocorrelations of the AR(1), MA(1) and ARMA(1,1) processes (i.e., to appreciate what’s in Table 1 on page 109).

In particular, the covariances of the ARMA(1,1) process are very similar to those of an AR(1), decaying geometrically past the initial values: \( \gamma(1), \gamma(k) = \phi\gamma(k-1) \) for \( k = 2, 3, \ldots \). The Yule-Walker equations provide the starting values for this recursion. Multiply both sides of the equation

\[
x_t = \phi_1x_{t-1} + \theta_1w_{t-1} + w_t
\]

by \( x_t, x_{t-1}, \ldots \) and take expectations (use the infinite moving average form \( x_t = \sum \psi_j w_{t-j} \) as well):

\[
\gamma(0) = \phi_1\gamma(1) + (1 + \theta_1\psi_1)\sigma^2
\]
\[
\gamma(1) = \phi_1\gamma(0) + \theta_1\sigma^2
\]
\[
\gamma(2) = \phi_1\gamma(1)
\]
\[
\gamma(h) = \phi_1\gamma(h-1), h = 3, 4, \ldots
\]

3.8 Simulate ARMA(1,1)

The point of this exercise is to notice that the estimated correlations don’t always look very much like the actual autocorrelations. Hence, using Table 1 is a lot harder than you might think unless the series is hundreds of points long.
3.9 Cardio mortality, data analysis

(a) I was curious whether any of you would extend the correlation function far enough to see the clear seasonal oscillation around 52 weeks. The data have a strong annual cycle.

(b) The predictions revert to the mean quickly, with rather wide intervals.

(c) The forecast errors are clearly dependent. In particular, even if we knew the model order, its parameters, and had infinite amounts of data, the errors are
\[
\begin{align*}
    y_{n+1} - \hat{y}_{n+1} &= w_{n+1} \\
    y_{n+2} - \hat{y}_{n+2} &= \psi_1 w_{n+1} + w_{n+2} \\
    y_{n+3} - \hat{y}_{n+3} &= \psi_2 w_{n+1} + \psi_1 w_{n+1} + w_{n+3}
\end{align*}
\]
and so forth. The presence of estimated coefficients adds further dependence, but the order of that effect is dominated by the role of the \(w_t\)s.

(d) Bonferroni adjustment is a start (i.e., set the coverage to \(1-\alpha/4\), say) and works even though the implied tests are dependent. You still get 95% coverage even though the intervals are correlated. Make sure that you aren’t using independence of the coverage of the intervals; as noted in “c”, they are dependent.

3.10 Infinite and finite past predictors

To make life a little easier, I will write the process as \(X_t = w_t - \hat{\theta}w_{t-1}\), where I have flipped the sign \(\hat{\theta} = -\theta\).

(a) Write the mean-zero MA(1) process as the infinite autoregression \((x_t = \theta(B)w_t \text{ implies } \theta^{-1}(B)X_t = w_t)\)
\[
x_t = w_t - \tilde{\theta}x_{t-1} - \tilde{\theta}^2x_{t-2} + \cdots = w_t - \sum_{j=1}^{\infty} \tilde{\theta}^j x_{t-j}
\]
Hence, \(\hat{x}_{n+1} = \mathbb{E}(x_{n+1}|x_n, x_{n-1}, \ldots)\) given the infinite past is the sum on the r.h.s. after dropping \(w_t\), with mean squared error \(\sigma_w^2\)
\[
\hat{x}_{n+1} = -\sum_{j=1}^{\infty} \tilde{\theta}^j x_{n+1-j}
\]

(b) Having access to the finite past means truncating the infinite sum, but those terms are not independent so this does not quite seem right. You can see, though, that simple truncation won’t cost you very much since the weight on \(x_0\), for example, is \(\theta^{n+1}\). The details are a little more messy, but work out nicely thanks to the geometric weights:
\[
\begin{align*}
    \mathbb{E}(x_{n+1} - \hat{x}_{n+1})^2 &= \mathbb{E}(w_{n+1} + \sum_{j=n+1}^{\infty} \tilde{\theta}^j x_{n+1-j})^2 \\
    &= \mathbb{E}(w_{n+1} + \tilde{\theta}^{n+1} \sum_{j=n+1}^{\infty} \tilde{\theta}^{j-(n+1)} x_{n+1-j})^2
\end{align*}
\]
\[
\mathbb{E}(w_{n+1} + \tilde{\theta}^{n+1}w_0)^2 = \sigma_w^2(1 + \tilde{\theta}^{2(n+1)})
\]

Alternatively, direct backsubstitution reveals a telescoping sum:
\[
\tilde{x}_{n+1} = w_{n+1} + \tilde{\theta}w_n + \tilde{\theta}^2w_{n-1} + \cdots + \tilde{\theta}^n w_1
\]
\[
= w_{n+1} + \tilde{\theta}(w_n - \tilde{\theta}w_{n-1}) + \tilde{\theta}^2(w_{n-1} - \tilde{\theta}w_{n-2}) + \cdots + \tilde{\theta}^n(w_1 - \tilde{\theta}w_0)
\]
\[
= w_{n+1} + \tilde{\theta}^{n+1}w_0
\]

That’s a nicer, more constructive analysis — at least for the MA(1) case in which the sum telescopes.

3.11 Nonsingularity of \(\Gamma_n\)

There’s a nice proof based on the spectral density. Heuristically, the eigenvalues of \(\Gamma_n\) are the equally spaced values of the spectral density function, and that function is positive. But that’s not fair.

Suppose that \(\Gamma_n\) is singular for some \(n\). Then for some \(p < n\) there is an AR(\(p\)) model that fits perfectly, \(X_t = \sum_{j=1}^{p} \phi_j X_{t-j}\). Stationarity then means that this expression holds true when shifted, allowing us to backsubstitute, so we have, say,
\[
X_n = \sum_{j=1}^{p} a_j X_j
\]

The Yule-Walker type of equations tell you that (multiply both sides by \(X_n\) and take expectations)
\[
\gamma(0) = \sum_j a_j \gamma(n-j)
\]

Let \(n \to \infty\) and you can see that the variance \(\gamma(0)\) is pushed to zero if the coefficients are bounded. For that, use the sum on the r.h.s. in (1) to show that
\[
\gamma(0) = \sum_j a_j a_k \gamma(j-k) = \phi'\Gamma_p\phi
\]

You can use this to show that the \(\phi_j\) are bounded (for example, plug in the eigendecomposition of \(\Gamma_p\), toss all but one).

3.12 Reflection coefficients

We derived the expression for the reflection coefficient \(\phi_{hh}\) in class. Write down the matrix and scalar equations, eliminate one of the unknowns, then solve for \(\phi_{kk}\). I liked the question because it explains the interpretation of \(\phi_{kk}\) as a partial correlation more clearly, I think, than in the body of the text itself. If you wanted to elaborate, then note that the regression of \(x_t\) on \(x = x_{t-1}, \ldots, x_{t-h}\) gives residuals
\[
\epsilon_t = x_t - \gamma'_h \Gamma^{-1}_h x
\]

3.17 Cardio mortality, data analysis, continued

The AR(2) fits obtained by least squares and Yule-Walker are very similar. You’d need to fit a very high-order model to see a difference.
For the asymptotic variance, use the expression given in the text, \( \text{Var}(\hat{\phi}_1) \approx \text{Var}(\hat{\phi}_2) \approx \frac{1}{n}(1-\phi_2^2) \)

3.19 Hidden white noise

The zeros cancel in the MA and AR components; the data are white noise.

3.31 Global temperature deviations, data analysis

The standard diagnostics (graphics, residual correlations) show that the series ought to be differenced. I let arima fit a model with differencing to build the predictions since this left the messy task of integrating up the predictions to the R software. There are problems, however, with R’s calculations for an integrated model (check out the ranting section of the textbook’s web site).