Readings:
Chapter 3.4 in Chatfield’s *The Analysis of Time Series, An Introduction (6th Edition)*

Homework Problems Supporting Lecture:
3.4, 3.5, 3.9, 3.13

Topics Covered:
Autoregressive Processes AR\((p)\), Duality with AM\((\infty)\)

Review
We have taken some tentative steps in estimation for an MA\((q)\) process and encountered some roadblocks. Along the way, we developed some new notation, the backwards shift operator \(B\), and saw that a new condition, invertibility, was required to ensure a unique MA\((q)\) process for a given ac.f. We even saw an “easy to check” condition for invertibility, namely that the roots of \(P(B) = \beta_0 + \beta_1 B + \cdots + \beta_q B^q\) must all lie outside the unit circle.

Autoregressive Processes
It turns out that, in order to get some traction on estimation, we need to delve a little deeper and explore another special case of stochastic process, the autoregressive process of order \(p\), AR\((p)\). This has nothing to do with retired persons and everything to do with the formula

\[ X_t = Z_t + \text{history} \]

That’s a little vague, so let’s spell it out. The \(Z_t\)s are white noise with \(E[Z_t] = 0\) and \(V[Z_t] = \sigma^2_Z\). By history we mean previous terms in the process

\[ X_t = Z_t + \alpha_1 X_{t-1} + \cdots + \alpha_p X_{t-p} \]

So rather than model explicitly on a finite set of innovations as in an MA\((q)\) process, we build knowledge of prior outcomes into our model. You should be chomping at the bit to express this in backwards shift

\[ X_t = Z_t + \alpha_1 BX_t + \cdots + \alpha_p B^p X_t = Z_t + (\alpha_1 B + \cdots + \alpha_p B^p)X_t \]

Is it clear that we can express \(Z_t\) (the innovation at time \(t\)) as

\[ Z_t = (1 - \alpha_1 B - \cdots - \alpha_p B^p)X_t \]
But using the algebra of shift operators, doesn’t this mean that we can write

\[ X_t = \frac{1}{(1 - \alpha_1 B - \cdots - \alpha_p B^p)} Z_t \]

Good notation, like the backward shift operator, doesn’t just let us write things more compactly, it suggests results and allows us to proceed much faster, and with greater clarity, than we were previously able. Push this result just a little further and we see that an autoregressive process of order p, \( AR(p) \), may be thought of as an (infinite order) moving average process

\[ X_t = (1 + \beta_1 B + \beta_2 B^2 + \cdots) Z_t \]

OK, we will have to find the \( \beta \)'s, but still…proof of concept…examples will be forthcoming. In the meantime, the beauty of treating the \( AR(p) \) process like this is we quickly inherit several results from our work with \( MA(q) \) processes.

Suppose you’d like to find the average of an \( AR(p) \) process. Just recall \( E[Z_t] = 0 \) and take

\[ E[X_t] = E[(1 + \beta_1 B + \beta_2 B^2 + \cdots) Z_t] \]

\[ E[X_t] = E[Z_t] + \beta_1 E[Z_{t-1}] + \cdots + \beta_k E[Z_{t-k}] + \cdots = 0 \]

How about the variance? The \( Z_t \) are independent, so (please use \( V[aX] = a^2 V[X] \))

\[ V[X_t] = V[(1 + \beta_1 B + \beta_2 B^2 + \cdots) Z_t] \]

\[ V[X_t] = V[Z_t] + \beta_1^2 V[Z_{t-1}] + \cdots + \beta_k^2 V[Z_{t-k}] + \cdots \]

\[ V[X_t] = \sigma_Z^2 (1 + \beta_1^2 + \cdots + \beta_k^2 + \cdots) = \sigma_Z^2 \sum_{i=0}^{\infty} \beta_i^2 \]

We obviously took \( \beta_0 = 1 \). Evidently we have a necessary condition for stationarity, that is, we need the \( \beta_i^2 \)'s to converge.

Finally, how about autocorrelation and autocovariance? We saw, for a \( MA(q) \) process,

\[ \gamma(k) = \sigma_Z^2 \cdot \sum_{i=0}^{q-k} \beta_i \beta_{i+k} \] (where appropriate)
So, we have, for an AR(p) process

\[ \gamma(k) = \sigma_Z^2 \cdot \sum_{i=0}^{\infty} \beta_i \beta_{i+k} \] (where appropriate)

Sneaking in a result from real analysis (though maybe not exactly from Calculus II), the series converges when the more basic series is absolutely convergent:

\[ \sum_{i=0}^{\infty} |\beta_i| \]

Time for some examples. We will work in this lecture to exhibit the results we’ve been developing, then get back to estimation in the next lecture. That will allow us to get back to the time series case studies we’ve begun.

**Simulating an AR(p) Process**

We can lightly edit our MA(q=1) code to produce an AR(p=1) simulation.

```r
set.seed(2015); N=1000; alpha = .4; Z = rnorm(N,0,1) X=NULL; X[1] = Z[1];
for (t in 2:N) {
  X[t] = Z[t] + alpha*X[t-1] ;
}
X.ts = ts(X)
plot(X.ts,main="AR(1) Time Series on White Noise, alpha=.4")
(X.acf = acf(X.ts, main="AR(1) Time Series on White Noise, alpha=.4"))
```

It looks to me like the first two or three lag spacings have a significant value. What would we have predicted? Start with
\[ X_t = Z_t + 0.4X_{t-1} = (1 + 0.4B)X_t \]

Walking down the same path again

\[ Z_t = X_t - 0.4BX_t \]
\[ Z_t = (1 - 0.4B)X_t \]
\[ X_t = \left\{ \frac{1}{1 - 0.4B} \right\} Z_t \]

Remembering, yet again, out geometric series

\[ \sum_{k=0}^{\infty} a^k = \frac{1}{1 - a} \]

We write an infinite order \( MA(q) \) as

\[ X_t = \left\{ \sum_{k=0}^{\infty} (0.4B)^k \right\} Z_t \]

Being explicit

\[ X_t = \{1 + 0.4B + (0.4)^2B^2 + (0.4)^3B^3 + \cdots \} Z_t \]

We are trading the \( \alpha \) coefficient: \( \alpha_1 = 0.4 \) for an infinite set of \( \beta \) coefficients

\[ \beta_0 = 1, \ \beta_1 = 0.4, \ \beta_2 = 0.16, \ \beta_3 = 0.064, \ldots, \beta_k = 0.4^k, \ldots \]

Can we work out the autocovariances and autocorrelations?

\[ \gamma(k) = \sigma_Z^2 \cdot \sum_{i=0}^{\infty} \beta_i \beta_{i+k} = 1 \cdot \sum_{i=0}^{\infty} (0.4)^i (0.4)^{i+k} = 0.4^k \sum_{i=0}^{\infty} (0.4^2)^i \]
\[ \gamma(k) = 0.4^k \cdot \frac{1}{1 - 0.16} \]

And now scale for the autocorrelations

\[ \rho(k) = \frac{\sum_{i=0}^{\infty} \beta_i \beta_{i+k}}{\sum_{i=0}^{\infty} \beta_i \beta_i} \]
We worked out the numerator. For the denominator

\[
\sum_{i=0}^{\infty} \beta_i \beta_i = \sum_{i=0}^{\infty} 4^i \cdot 4^i = \sum_{i=0}^{\infty} 16^i = \frac{1}{1 - 0.16}
\]

This leads to a surprisingly simple result

\[
\rho(k) = \frac{0.4^k \cdot \frac{1}{1 - 0.16}}{1} = 0.4^k
\]

Nothing special about \( \alpha_1 = 0.4 \), so we have really shown that, for a first order autoregressive process \( X_t = Z_t + \alpha_1 X_{t-1} \)

\[
\rho(k) = \alpha_1^k
\]

In tabular format

<table>
<thead>
<tr>
<th>( k )</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>...</th>
<th>( k )</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \rho(k) )</td>
<td>0.4</td>
<td>0.16</td>
<td>0.064</td>
<td></td>
<td>0.4^k</td>
<td></td>
</tr>
</tbody>
</table>

How do these compare with our estimates from this realization? A quick call to \( (r.coef = X.acf$acf) \) tells me \( r_1 = 0.4208916633 \), \( r_2 = 0.2211209987 \), and \( r_3 = 0.0905580119 \). Run the code several times without setting the seed to see whether you believe our results are supported. We also made predictions for the mean and variance.

<table>
<thead>
<tr>
<th></th>
<th>predicted</th>
<th>actual</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>( \mu = 0 )</td>
<td>( \bar{x} = -0.105947 )</td>
</tr>
<tr>
<td>variance</td>
<td>( V[X_t] = \sigma_Z^2 \sum_{i=0}^{\infty} \beta_i^2 = \frac{1}{1 - 0.16} = 1.190476 )</td>
<td>( s^2 = 1.216524 )</td>
</tr>
</tbody>
</table>

Now that we know how to proceed, let’s look at how the parameter in a \( AR(p=1) \) process affects the autocorrelation function. Given

\[
\rho(k) = \alpha_1^k
\]
When $\alpha_1 \approx 1$ then we have something close to the random walk we studied previously. Recall that $\alpha = 1$ would give us a nonstationary process (growing variance). As $\alpha_1 \downarrow 0$ the correlations decay more quickly. Note that $\alpha_1 = 0$ would bring us back to white noise. For negative values $\alpha_1 < 0$ we have alternating positive and negative correlations as the terms “flip back and forth”.

Let’s look at a second order process, $AR(p = 2)$ with coefficients $\alpha_1 = 1, \alpha_2 = -0.8$.

$$X_t = Z_t + X_{t-1} - 0.8 X_{t-2}$$

I’ll perform the simulation first, then we’ll apply the theory. The code is changed very little.

```r
alpha1 = 1; alpha2 = -0.8; etc.
for (t in 3:N) {
    X[t] = Z[t] + alpha1*X[t-1] + alpha2*X[t-2];
}
```
Now do the usual analysis by first expressing in terms of the backward shift.

\[ X_t = Z_t + X_{t-1} - .8 X_{t-2} = Z_t + BX_t - .8 B^2 X_t = Z_t + (1 \cdot B - .8 \cdot B^2)X_t \]

Switch onto the \( Z_t \)

\[ Z_t = (1 - 1 \cdot B + .8 \cdot B^2) X_t \]

In principle, we could now write

\[ X_t = \frac{1}{(1 - 1 \cdot B + .8 \cdot B^2)} Z_t \]

I you don’t want to find the Taylor Series directly, you might try factoring, partial fractions, and the sum of two geometric series. Instead, let’s first see if our process is stationary, then explore the Yule-Walker equations to give us the autocorrelation function inductively.

**Stationarity of an AR(p) Process via \( \phi(B) \)**

When we had a MA(q) process we saw that *invertibility* could be demonstrated by showing the roots of

\[ P(B) = \beta_0 + \beta_1 B + \cdots + \beta_q B^q \]

stayed outside of the unit circle.
In much the same spirit (and again without proof - you have to make some choices in an introductory course) we will call the polynomial associated with our $AR(p)$ process $\phi(B)$.

$$\phi(B) = 1 - \alpha_1 B - \alpha_2 B^2 - \cdots - \alpha_p B^p$$

We now claim that if the polynomial associated with the $AR(p)$ process has all of its roots outside of the unit circle, then the process is stationary.

Our process has $\alpha_1 = 1$ and $\alpha_2 = -0.8$, so using the quadratic formula,

$$\frac{\alpha_1 \pm \sqrt{\alpha_1^2 - 4 \cdot (-\alpha_2)(1)}}{2 \cdot (-\alpha_2)} = \frac{1 \pm \sqrt{1 - 4 \cdot (-0.8) \cdot (1)}}{2 \cdot (-0.8)} = 0.625 \pm 0.927025 i$$

Simple complex arithmetic (not a typo!) gives us, for complex variable $z = a + i b$, the modulus $|z| = \sqrt{a^2 + b^2}$. Applied to our roots

$$|z_1| = |z_2| = \sqrt{0.625^2 + 0.927025^2} = 1.25 > 0$$

We are stationary.

**Yule-Walker**

We can derive the Yule Walker equations for the more general case, then apply to our specific example. Start with

$$X_t = Z_t + \alpha_1 X_{t-1} + \cdots + \alpha_p X_{t-p}$$


$$X_t X_{t-k} = Z_t X_{t-k} + \alpha_1 X_{t-1} X_{t-k} + \cdots + \alpha_p X_{t-p} X_{t-k}$$

For an $AR(p)$ process, $E[X_t] = 0$. So $cov[X_{t_1}, X_{t_2}] = E[X_{t_1} \cdot X_{t_2}]$.

Taking expectations:

$$E\left[X_t X_{t-k}\right] = E\left[Z_t X_{t-k}\right] + \alpha_1 E\left[X_{t-1} X_{t-k}\right] + \cdots + \alpha_p E\left[X_{t-p} X_{t-k}\right]$$

Swap expected values of products for covariances
\[ \text{cov}[X_t, X_{t-k}] = \alpha_1 \text{cov}[X_{t-1}, X_{t-k}] + \cdots + \alpha_p \text{cov}[X_{t-p}, X_{t-k}] \]

Again, since the process is stationary, the covariance only depends on the lag spacing

\[ \gamma(k) = \alpha_1 \gamma(1-k) + \cdots + \alpha_p \gamma(p-k) \]

Finally, as our text has frequently pointed out, for a stationary process, \( \gamma(k) = \gamma(-k) \) so write

\[ \gamma(k) = \alpha_1 \gamma(k-1) + \cdots + \alpha_p \gamma(k-p) \]

Scaling to get the autocorrelation gives

\[ \rho(k) = \alpha_1 \rho(k-1) + \cdots + \alpha_p \rho(k-p) \]

We are looking at the second order process \( X_t = Z_t + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} \)

\[ \rho(k) = \alpha_1 \rho(k-1) + \alpha_2 \rho(k-2) \]

\[ \rho(k) = \rho(k-1) - 0.8 \cdot \rho(k-2) \]

How does this move us forward? We know that \( \rho(0) = 1 \). Then for lag spacing \( k = 1 \)

\[ \rho(1) = \alpha_1 \rho(0) + \alpha_2 \rho(-1) = \alpha_1 + \alpha_2 \rho(1) \]

\[ \rho(1) = 1 - 0.8 \rho(1), \quad \rho(1) = \frac{1}{1.8} \approx 0.5555556 \]

Keep pushing this:

\[ \rho(2) = \rho(1) - 0.8 \cdot \rho(0), \quad \rho(2) \approx -0.2444444 \]

\[ \rho(3) = \rho(2) - 0.8 \cdot \rho(1), \quad \rho(3) \approx -0.6888889 \]

And so on. These are pretty good matches.

Take a look at the correlogram produced with the following code:
\( \alpha_1 = 1; \alpha_2 = -0.8 \)

\( \rho = \text{NULL} \)

\( \rho[1] = 1 \)

\( \rho[2] = \frac{\alpha_1}{1-\alpha_2} \)

\[ \text{for} \ (k \ \text{in} \ 3:30) \ {\} \]

\[ \rho[k] = \alpha_1 \times \rho[k-1] + \alpha_2 \times \rho[k-2]; \]

\( \text{plot}(\rho, \text{type} = "h", \text{main} = "\text{Theoretical Autocorrelation for AR(2), } a_1=1, a_2=-0.8") \)