The Moving Average Model and ARMA Models

2.1 White Noise
2.2 The MA(1) model
2.3 The MA(q) model
2.5 Inverting MA models
2.6 What does the PACF look like for MA?
2.7 ARMA(p,q) models.

• The Moving Average (MA) models are very different from the Autoregressive models both in terms of how we write them down and think about them as well as the implied dynamics.
• We begin with the simple idea of a white noise process and then build up to the moving average models.
2.1 White Noise

• A white noise process is a time series that has mean zero, some variance $\sigma^2$, and all autocorrelations equal to zero.

• So it is a series that has no correlation structure.

• Example: $r_t = \epsilon_t$ where $\epsilon_t$ is iid $N(0, \sigma^2)$. This is just the iid Normal time series model that we talked about before.

• We could extend this model by giving it a mean: $r_t = \mu + \epsilon_t$

• Since $\epsilon_t$ is iid $N(0, \sigma^2)$, the returns $r_t$ are iid $N(\mu, \sigma^2)$
White Noise for iid $N(\mu, \sigma^2)$

2.2 The MA(1) model

- We could make this model a little more interesting by saying that $r_t$ is determined by two sequential values of $\varepsilon_t$ as in:

$$r_t = \mu + \theta \varepsilon_{t-1} + \varepsilon_t$$

- The model is clearly different from the AR model since we write the time series for $r_t$ as a combination of two random outcomes from a Normal.
• Notice that both \( r_t \) and \( r_{t-1} \) are determined by \( \varepsilon_{t-1} \)

\[
\begin{align*}
  r_t &= \mu + \theta \varepsilon_{t-1} + \varepsilon_t \\
  r_{t-1} &= \mu + \theta \varepsilon_{t-2} + \varepsilon_{t-1}
\end{align*}
\]

• So if \( \varepsilon_{t-1} \) happens to be very large, and (say) \( \theta \) is positive, then it is more likely that both \( r_t \) and \( r_{t-1} \) will be large.

• Even though \( \varepsilon_t \) is iid, sequential outcomes for \( r_t \) will be correlated!

• Without loss of generality we will restrict \(|\theta| \leq 1\). There is a lack of identification otherwise.

• To better understand how the model works, let’s think about how we would generate data for this model.

• Consider the model

\[
r_t = .5 \varepsilon_{t-1} + \varepsilon_t \text{ where } \varepsilon_t \sim iid \ N(0,1)
\]

<table>
<thead>
<tr>
<th>( t )</th>
<th>( \varepsilon_t )</th>
<th>( \varepsilon_{t-1} )</th>
<th>( r_t )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.51</td>
<td>0.83</td>
<td>0.93</td>
</tr>
<tr>
<td>2</td>
<td>0.90</td>
<td>0.51</td>
<td>1.15</td>
</tr>
<tr>
<td>3</td>
<td>-1.19</td>
<td>0.90</td>
<td>-0.75</td>
</tr>
<tr>
<td>4</td>
<td>1.48</td>
<td>-1.19</td>
<td>0.89</td>
</tr>
<tr>
<td>5</td>
<td>1.50</td>
<td>1.48</td>
<td>2.24</td>
</tr>
<tr>
<td>6</td>
<td>-0.08</td>
<td>1.50</td>
<td>0.67</td>
</tr>
<tr>
<td>7</td>
<td>-0.12</td>
<td>-0.08</td>
<td>-0.16</td>
</tr>
<tr>
<td>8</td>
<td>-0.79</td>
<td>-0.12</td>
<td>-0.79</td>
</tr>
<tr>
<td>9</td>
<td>-0.19</td>
<td>-0.73</td>
<td>-0.56</td>
</tr>
</tbody>
</table>

Generate random outcomes from a \( N(0,1) \)
Let’s look at some simulated data from an MA(1) model

Series 1 $\theta=-.9 \mu=2 \sigma=2$
Series 2 $\theta=.9 \mu=2 \sigma=2$
Series 3 $\theta=.1 \mu=2 \sigma=2$

Note that we are interested in $\theta$ here. $\mu$ just determines the level while $\sigma$ determines how “spread out” the time series will be.
It is a little easier to see the dependence by plotting $r_t$ against $r_{t-1}$

The autocorrelation between $r_t$ and $r_{t-1}$ is about -.5.

Series 2 $\theta = .9 \; \mu = 2 \; \sigma = 2$
The autocorrelation between $r_t$ and $r_{t-1}$ is about .5.

Series 3 $\theta=.1$ $\mu=2$ $\sigma=2$
The autocorrelation between $r_t$ and $r_{t-1}$ is about .2.

**MA(1) model summary**

- The MA(1) model is given by:

$$r_t = \mu + \theta \varepsilon_{t-1} + \varepsilon_t \text{ where } \varepsilon_t \sim iid \ N(0, \sigma^2)$$

- There are three parameters that we need to know, $\mu$, $\theta$, and $\sigma$. 


2.3 The MA(q) model

- We already saw that for an MA(1) model $r_t$ is correlated with $r_{t-1}$.

- What do you expect for the correlation between $r_t$ and $r_{t-2}$?

- Well, $r_t$ depends on $\varepsilon_t$ and $\varepsilon_{t-1}$

- $r_{t-2}$ depends on $\varepsilon_{t-2}$ and $\varepsilon_{t-3}$

- Since $\varepsilon_t$ is iid we should expect that the correlation between $r_t$ and $r_{t-j}$ should be zero for $j>1$.

Series 2 $\theta=.9 \mu=2 \sigma=2$

The autocorrelation between $r_t$ and $r_{t-2}$ is about 0.
• We can allow for richer dependence in $r_t$ by allowing $r_t$ to depend on more values of $\varepsilon_t$.

• The MA($q$) model says:

$$r_t = \mu + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t$$

so the value of $r$ depends on $\varepsilon_t$ and $q$ past values of $\varepsilon$.

• An MA(2) model looks like this:

$$r_t = \mu + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \varepsilon_t$$

where $\varepsilon_t \sim iid \ N(0, \sigma^2)$

• Now $r_t$ and $r_{t-2}$ will be correlated since they both depend on $\varepsilon_{t-2}$.

(Notice that $r_{t-2} = \mu + \theta_1 \varepsilon_{t-3} + \theta_2 \varepsilon_{t-4} + \varepsilon_{t-2}$)
Fact: For an MA(q) model, the first q autocorrelations of $r_t$ will be non-zero while the $q+j^{th}$ autocorrelations will be (about) zero for all $j>0.$

2.4 Moving average properties

- We saw that the Autoregressive model implies slow decay of the autocorrelations. This means that values today are correlated with future values in a slowly declining fashion.
- The MA model says that what happens this period is correlated with what happens in the next period but not beyond next period.
- A simple intuitive idea is alike a hangover model. If you drank a lot last night, you are not likely to drink a lot tonight, but tomorrow night you are feeling fine and might drink again!
2.5 Inverting an MA model

• Consider the MA(1) model

\[ y_t = \mu + \theta \varepsilon_{t-1} + \varepsilon_t \]

• Using our lag operator we can write:

\[ y_t = \mu + (1 + \theta L) \varepsilon_t \]

• If \(|\theta| < 1\) then an inverse of the lag polynomial exists and the inverse is given by:

\[(1 + \theta L)^{-1} = (1 + (-\theta L) + (-\theta L)^2 + (-\theta L)^3 \ldots)\]

• You can verify that

\[(1 + \theta L)^{-1} (1 + \theta L) = 1\]

• Multiplying both sides of the MA(1) model by \((1 + \theta L)^{-1}\) we get:

\[ \sum_{j=0}^{\infty} (-\theta)^j y_{t-j} = \mu \sum_{j=0}^{\infty} (-\theta)^j + \varepsilon_t = \frac{\mu}{1 + \theta} + \varepsilon_t \]

or

\[ y_t = \frac{\mu}{1 + \theta} - \sum_{j=1}^{\infty} (-\theta)^j y_{t-j} + \varepsilon_t = \beta_0 + \sum_{j=1}^{\infty} \beta_j y_{t-j} + \varepsilon_t \]

where \(\beta_0 = \frac{\mu}{1 + \theta}\) and \(\beta_j = -(-\theta)^j\)

So, if \(|\theta| < 1\) we say that the MA(1) mode is invertible and can therefore be written as an infinite order AR model.
2.6 What does the PACF look like for an MA model?

- Since we can write an invertible MA model as an infinite order AR model, the PACF function for an MA model should not be truncated, but decline roughly in powers of \((-\theta)^j\).

Here is the ACF and PACF for the first MA model:

<table>
<thead>
<tr>
<th>Auto.correlation</th>
<th>Partial Correlation</th>
<th>AC</th>
<th>PAC</th>
<th>Q-Stat</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>1</td>
<td>-0.473</td>
<td>-0.473</td>
<td>23.084</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>2</td>
<td>-0.048</td>
<td>-0.346</td>
<td>33.309</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>3</td>
<td>-0.035</td>
<td>-0.234</td>
<td>23.439</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>4</td>
<td>-0.032</td>
<td>-0.121</td>
<td>23.537</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>5</td>
<td>-0.115</td>
<td>-0.328</td>
<td>34.587</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>6</td>
<td>-0.336</td>
<td>-0.888</td>
<td>30.949</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>7</td>
<td>-0.277</td>
<td>-0.184</td>
<td>36.308</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>8</td>
<td>-0.175</td>
<td>-0.014</td>
<td>42.555</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>9</td>
<td>-0.033</td>
<td>-0.087</td>
<td>42.674</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>10</td>
<td>-0.085</td>
<td>-0.211</td>
<td>43.687</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>11</td>
<td>-0.074</td>
<td>-0.114</td>
<td>44.916</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>12</td>
<td>-0.109</td>
<td>-0.140</td>
<td>44.484</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>13</td>
<td>-0.039</td>
<td>-0.064</td>
<td>44.692</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>14</td>
<td>-0.072</td>
<td>-0.025</td>
<td>46.272</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>15</td>
<td>-0.084</td>
<td>0.200</td>
<td>40.116</td>
</tr>
</tbody>
</table>
Inverting an MA(q) model

- Consider the finite order MA(q) model:
  \[ y_i = \mu + \left( 1 + \theta_1 L + \theta_2 L^2 + \ldots + \theta_q L^q \right) \varepsilon_i \]

- We factor the polynomial into:
  \[ (1 + \theta_1 L + \theta_2 L^2 + \ldots + \theta_q L^q) = (1 - \lambda_1 L)(1 - \lambda_2 L)\ldots(1 - \lambda_q L) \]

- If all of the values of \( \lambda_j \) lie inside the unit circle the polynomial is invertible. We write the inverse for the \( j^{th} \) term as:
  \[ (1 - \lambda_j L)^{-1} \]
  \[ (1 + \theta_1 L + \theta_2 L^2 + \ldots + \theta_q L^q)^{-1} = (1 - \lambda_1 L)^{-1} (1 - \lambda_2 L)^{-1} \ldots (1 - \lambda_q L)^{-1} \]
\[
\left(1 + \theta_1 L + \theta_2 L^2 + \ldots + \theta_q L^q\right)^{-1}\left(1 + \theta_1 L + \theta_2 L^2 + \ldots + \theta_q L^q\right)
\]
\[
= (1 - \lambda_1 L)^{-1} (1 - \lambda_2 L)^{-1} \ldots (1 - \lambda_q L)^{-1} \left[ (1 - \lambda_1 L)(1 - \lambda_2 L) \ldots (1 - \lambda_q L) \right]
\]
\[
= (1 - \lambda_1 L)^{-1} (1 - \lambda_2 L)(1 - \lambda_2 L)^{-1} \ldots (1 - \lambda_q L)^{-1} (1 - \lambda_q L)
\]
\[
= 1
\]

- Multiplying both sides by the inverses we get:

\[
\left(1 + \theta_1 L + \theta_2 L^2 + \ldots + \theta_q L^q\right)^{-1} y_i
\]
\[
= (1 + \theta_1 L + \theta_2 L^2 + \ldots + \theta_q L^q)^{-1} \left[ \mu + (1 + \theta_1 L + \theta_2 L^2 + \ldots + \theta_q L^q) \varepsilon_i \right]
\]

or

\[
(1 - \lambda_1 L)^{-1} (1 - \lambda_2 L)^{-1} \ldots (1 - \lambda_q L)^{-1} y_i
\]
\[
= (1 - \lambda_1 L)^{-1} (1 - \lambda_2 L)^{-1} \ldots (1 - \lambda_q L)^{-1} \mu + \varepsilon_i
\]
so any invertible MA(q) model can be written as an infinite order AR model.

\[
(1 - \lambda_1 L)^{-1} (1 - \lambda_2 L)^{-1} ... (1 - \lambda_q L)^{-1} y_t = (1 - \lambda_1 L)^{-1} (1 - \lambda_2 L)^{-1} ... (1 - \lambda_q L)^{-1} \mu + \varepsilon_t \\
y_t + \sum_{j=1}^{\infty} \beta_j y_{t-j} = \beta_0 + \varepsilon_t \\
y_t = \beta_0 + \sum_{j=1}^{\infty} \beta_j y_{t-j} + \varepsilon_t
\]

Summary of ACF and PACF for AR(p) and MA(q) models

<table>
<thead>
<tr>
<th></th>
<th>AR(p) Model</th>
<th>MA(q) Model</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ACF</strong></td>
<td>Slowly decaying</td>
<td>Non-zero for first q and (about) zero thereafter</td>
</tr>
<tr>
<td><strong>PACF</strong></td>
<td>Non-zero for first p and (about) zero thereafter</td>
<td>Slowly decaying</td>
</tr>
</tbody>
</table>
2.7 ARMA(p,q) models

- The world does not always conform to the simple dichotomy of AR and MA models.
- In reality, the world is complex and we more complex models are often needed.
- There is nothing to prevent us from combining AR and MA models.
- This seems to work very well in practice.

ARMA(1,1)

\[ r_t = \beta_0 + \beta_1 r_{t-1} + \theta \varepsilon_{t-1} + \varepsilon_t \]

where \( \varepsilon_t \sim iid \ N(0,\sigma^2) \)

and \( \varepsilon_t \) independent of \( r_{t-j}, j > 0 \)
• An ARMA(p,q) model will have p lags of r and q lags of $\varepsilon$.

$$r_i = \beta_0 + \beta_1 r_{i-1} + \cdots + \beta_p r_{i-p} + \theta_1 \varepsilon_{i-1} + \cdots + \theta_q \varepsilon_{i-q} + \varepsilon_i$$

• The model becomes difficult to interpret as the dynamics are much more complex.
• It is also difficult to infer $p$ and $q$ from the ACF and PACF.
• However, we can still use the same approach as before.
  – Look at the ACF and PACF (see if its pure AR or MA).
  – If it is not obviously pure AR or MA fit a low order ARMA model (I like to start with ARMA(1,1)).
  – Check the residual series to see if it is roughly uncorrelated.
  – Add additional AR and MA terms if necessary.