Handout 7

1   Invertibility of MA processes

In Handout 6 we considered the problem of expressing an AR process as an MA process of infinite order. Here, we are concerned with the related problem of whether or not an MA process can be represented as an AR process (possibly of infinite order). If so, we say that the MA process is invertible. What are the conditions that an MA process has to satisfy for it to be invertible?

To fix ideas, consider and MA(1) process:

$$y_t = \epsilon_t + \rho\epsilon_{t-1} \tag{1}$$

The process also implies that

$$y_{t-1} = \epsilon_{t-1} + \rho\epsilon_{t-2} \tag{2}$$

Multiplying equation (2) by ρ and subtracting it from equation (1), one obtains

$$y_{t-1} - \rho y_{t-1} = \epsilon_t - \rho^2\epsilon_{t-2} \tag{3}$$

Lagging equation (2) once more, multiplying it by ρ and adding it to equation (3) yields:

$$y_{t-1} - \rho y_{t-1} + \rho^2 y_{t-2} = \epsilon_t + \rho^3\epsilon_{t-3}$$

Continuing in this fashion, the question of invertibility for this process boils down to whether or not the innovation ϵ_t can be recovered from the current and lagged observed variables y

without any reference to other innovations. As suggested by the substitutions above, this will be the case when $|\rho| < 1$.

The similarities with the invertibility of the AR process into an MA process extend to the less restrictive MA(q) process. Consider now

$$y_t = \epsilon_t + \rho_1 \epsilon_{t-1} + \rho_2 \epsilon_{t-2} + \ldots + \rho_q \epsilon_{t-q}. \tag{4}$$

The process above can be expressed in vector form as:

$$
\begin{pmatrix} \epsilon_t \\ \epsilon_{t-1} \\ . \\ . \\ . \\ \epsilon_{t-q} \end{pmatrix} = A \begin{pmatrix} \epsilon_{t-1} \\ \epsilon_{t-2} \\ . \\ . \\ . \\ \epsilon_{t-q-1} \end{pmatrix} + \begin{pmatrix} y_t \\ 0 \\ . \\ . \\ . \\ 0 \end{pmatrix} \tag{5}
$$

where

$$
A = \begin{pmatrix} -\rho_1 & -\rho_2 & . & . & . & -\rho_q \\ 1 & 0 & . & . & . & 0 \\ 0 & 1 & . & . & . & 0 \\ 0 & 0 & . & & & \\ 0 & 0 & 0 & . & & \\ 0 & 0 & 0 & 0 & . & 0 \end{pmatrix}
$$

By the same argument as outlined for the solution of an AR(n) process, the MA(q) process above will be invertible if all the eigenvalues of A are strictly less than 1 in modulus.

1.1   Implications for estimation

Estimating MA processes can be a tricky issue. We shall explore how to estimate MA processes directly later on in the course, with the aid of the Kalman filter. In the meantime, if an MA process is invertible, the section above suggests that we could recover the parameters of

interest through its AR representation. If the eigenvalues of the matrix A above for the specific process of interest are well within the unit circle, then an AR process of small order might very well provide a good approximation to the exact $AR(\infty)$ process. Furthermore, as seen in the previous handout, the parameters of interest could be estimated through a simple OLS regression.

1.2 Invertible and non-invertible representations

Consider again the MA(1) process

$$y_t = \epsilon_t + \rho\epsilon_{t-1},$$

where ϵ_t is governed by a white noise process with variance σ^2. Let $\rho > 1$. As we saw in the preceding handout, y_t will still be covariance stationary, in this case. Consider also an alternative process

$$\tilde{y}_t = \tilde{\epsilon}_t + \tilde{\rho}\tilde{\epsilon}_{t-1},$$

where $\tilde{rho}$ is equal to $\frac{1}{\rho}$, therefore the process is invertible. Let $\tilde{\epsilon}$ be governed by a white noise process with variance $\theta^2\sigma^2$.

Proposition 1 y_t and $\tilde{y}_t$ will have the same first and second moments.

This proposition is easy to prove. Just look at the formulas for the variance and autocovariance at lag 1 from Handout 6.

Proposition 2 Even if the data-generating process is y_t the process governing $\tilde{y}_t$ will give a valid representation of the process with the implied series for $\tilde{epsilon}_t$ being white noise. The innovations associated with this representation are called "fundamental."

This proposition is a little more tricky to prove. Instead, verify it with Monte Carlo experiments.

Notice that a corollary of the two propositions above is that any MA(1) process with $|\rho| > 1$ has an equivalent invertible representation. If $|rho| = 1$ the process only has one representation and it is non-invertible.

2 Tests for serial correlation

After estimating the parameters governing a certain process, one may wonder if the residuals actually conform with a white noise process. Several tests are routinely run to test whether or not the residuals are white noise.

Portmanteau test.

The Q statistic of the Portmanteau test is routinely made available by time series packages such as EViews. The statistic can be motivated as appropriate for carrying out a Lagrangian multiplier test with the null hypothesis of white noise, against the alternative of an AR(K) or MA(K) process. The statistic is computed as:

$$Q = N \sum_{k=1}^{K} r_k^2$$

where N is the number of observations in the sample, and r_k is the sample autocorrelation at lag k. This statistic is approximately distributed as χ^2 with $K - p - q$ degrees of freedom, where p and q are the number of AR and MA terms in the fitted model.

The modified Ljung-Box-Pierce statistic, given by

$$Q^* = N(N+2) \sum_{k=1}^{K} \frac{r_k^2}{N-k}$$

is also frequently used. It has the same approximate distribution as the Q statistic above.

Durbin-Watson statistic. The Durbin-Watson d statistic is given by

$$d = \sum_{t=2}^{N} \frac{(\hat{z}_t - \hat{z}_{t-1})^2}{\hat{z}_t^2}$$

The d statistic can be used to test the null hypothesis that the residuals z_t are white noise, against the alternative hypothesis that the residuals are governed by an AR(1) process. Notice that if the model under the null has been fitted, we expect $d \approx 2$. You can see that by noticing that:

$$\sum_{t=2}^{N} (\hat{z}_t - \hat{z}_{t-1}) \approx 2 \sum_{t=2}^{N} \hat{z}_t - 2 \sum_{t=2}^{N} \hat{z}_t \hat{z}_{t-1}$$

where $\frac{\sum \hat{z}_t \hat{z}_{t-1}}{\hat{z}_t^2}$ is the autocorrelation at lag 1 for the residuals. Tabulations of the cut off points for the d statistic are commonly reproduced in econometrics textbooks. For example, see Greene.

3   Detrending

Unfortunately, not all of the time series that might be of interest appear to be stationary. A first reasonable way of handling this issue is to check whether the observations available for the series of interest appear to have an upward trend visually.

If an upward trend appears obvious to visual inspection, consider a few alternatives:

1. Difference Consider the process

   $$y_t - y_{t-1} = \rho(y_{t-1} - y_{t-2}) + \epsilon_t$$

   The process described above is said to be integrated of order 1 and autoregressive of order 1. Taking $\rho > 0$, the observations drawn from such a process will indeed exhibit an upward trend. For the process above, differencing the observations before fitting an AR process seems like a very sensible strategy. So what are potential pitfalls?

   Well, a simple autoregressive process (not an integrated one) with a high autoregressive coefficient might also appear to have a spurious upward trend if only few observations are available. In that case differencing would introduce a non-invertible MA component into the process. Furthermore, the growth rate for the process might be deterministic, rather than stochastic, but observational noise might make the deterministic trend of difficult detection.

   Tests such as the one proposed by Dickey and Fuller, can take some of the guess-work away from the visual inspection of the data.

2. Removing a time trend Consider the process

   $$y_t = \gamma t + \rho y_{t-1} + \epsilon_t$$

   If the data generating process conforms to the hypothesis above, then removing a deterministic trend from the observations prior to fitting the AR(1) component of the process would be a sensible strategy.

3. Accounting for cointegrating relationships

Remember the stochastic growth model outlined in Handout 4.

There we showed that the model's consumption series c_t could be made stationary when divided by $a_t q_t$, which were both taken to be unobserved integrated processes. Alternatively, subtracting the log of $a_t q_t$ from the log of c_t would also produce a stationary series. When a simple linear relationship among integrated variables produces a result that is stationary, there is said to be a cointegrating relationship between the original series.

The problem with exploiting this cointegrating relationship involving c_t, a_t and q_t for estimation purposes is that the term a_t and q_t are unobserved.

More practically, remember that the theoretical model also implied that y_t could be made stationary by the same normalization involving $a_t q_t$. By implication, c_t/y_t would also be stationary. Chapters 19 and 20 of Hamilton "Time Series Analysis" (listed in the syllabus) provide a good treatment of many issues surrounding cointegration. One of the tests often used in practice to test the hypothesis of cointegration among time series observations is the Johansen test.

In sum, notice that detrending is a tricky issue and that it involves taking a stand on the form of the data-generating process. Moreover, when conducting hypothesis testing on estimated parameters, the detrending method will also influence the properties of those tests.

4   Forecasting: Box Jenkins

This section outlines the methodology outlined by Box and Jenkins to set up a forecasting model. For more details see the classic textbook by Box and Jenkins: "Time Series Analysis, Forecasting and Control", Holden Day, 1970.

1. Model identification Examine the data to determine which member of the ARIMA class of models appears to be the most appropriate.

The examination of the data my rely on plots, the correlogram, tests for stationarity and cointegration.

2. Estimation

Use an estimator appropriate to the model chosen to estimate the relevant parameters.

3. Diagnostic checking

Examine the residuals from the fitted model to see if they conform with the model's hypotheses. Use tests such as the Portmanteau test outlined above.

4. Consideration of alternative models if necessary

If the diagnostic check passes, you are done. Otherwise modify the model until the diagnostic tests become satisfactory.

Once a suitable model is chosen, out-of-sample forecasts can be obtained by taking conditional expectations.

The methodology above is widely used in practice. The defense for any one forecastic methodology rests on the forecast performance. For a systematic review of alternatives, see for example Chatfield "Time-Series Forecasting", Chapman and Hall, 2001.

We shall return to the topic of comparing forecasts later on in the course.