Lectures for STAT280/380 (continued)

Nalini Ravishanker

Fall 2005
Forecasting ARMA Processes

Consider a stationary and invertible ARMA(p,q) process $\phi(B)X_t = \theta(B)Z_t$ with MA representation given by

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$$

where, $\{Z_t\} \sim WN(0, \sigma^2)$ and we recall that the $\psi$ weights are given by

$$\psi(B) = \phi^{-1}(B)\theta(B).$$
Suppose we wish to forecast $X_{n+l}$ for $l = 1, 2, \cdots, L$. Denote the forecast for $X_{n+l}$ by $X_n(l)$. We can write the forecast for $X_{n+l}$ as a linear combination of past random shocks:

$$X_n(l) = \xi_0 Z_n + \xi_1 Z_{n-1} + \xi_2 Z_{n-2} + \cdots$$

The coefficients $\xi_j$ must be determined such that the mean squared forecast error

$$E[(X_{n+l} - X_n(l))^2]$$

is minimized. The resulting forecasts are the Minimum Mean Squared Error (MMSE) Forecasts.
The mean square error is

\[ E[X_{n+l} - X_n(l)]^2 = \]
\[ = E[Z_{n+l} + \psi_1 Z_{n+l-1} + \cdots + \psi_{l-1} Z_{n+1} + (\psi_l - \xi) Z_n + (\psi_{l+1} - \xi_1) Z_{n-1} + \cdots]^2 \]
\[ = (1 + \psi_1^2 + \cdots + \psi_{l-1}^2) \sigma^2 + \sum_{j=0}^{\infty} (\psi_{l+j} - \xi_j)^2 \sigma^2 \]

Clearly, this is minimized if \( \xi_j = \psi_{l+j} \) for \( j = 0, 1, 2, \ldots \).

Hence the MMSE forecast for \( X_{n+l} \) is given by

\[ X_n(l) = \psi_l Z_n + \psi_{l+1} Z_{n-1} + \cdots. \]
We can also express the MMSE forecast as the conditional expectation of $X_{n+l}$ given $X_n, X_{n-1}, \cdots$ (which is equivalent to $Z_n, Z_{n-1}, \cdots$):

$$E[X_{n+l}|X_n, X_{n-1}, \cdots] = E[(Z_{n+l} + \psi_1 Z_{n+l-1} + \cdots + \psi_{l-1} Z_{n+1} + \psi_l Z_n + \psi_{l+1} Z_{n-1} + \cdots)|X_n, X_{n-1}, \cdots] = \psi_l Z_n + \psi_{l+1} Z_{n-1} + \cdots$$

since we have seen that

$$E[Z_{n+h}|X_n, X_{n-1} \cdots] = \begin{cases} Z_{n+h} & h \leq 0 \\ 0 & h > 0 \end{cases}.$$
Example: AR(1) process: Let

\[ X_t - \mu = \phi(X_{t-1} - \mu) + Z_t \]

with \(|\phi| < 1\). Note that

\[
E(X_n|X_n, X_{n-1}, \cdots) = X_n, \\
E(Z_{n+1}|X_n, X_{n-1}, \cdots) = 0
\]
The MMSE forecast for $X_{n+1}$ for lead time $l = 1$ is

\[ X_n(1) = E[X_{n+1}|X_n, X_{n-1}, \ldots] \]
\[ = E[\{\mu + \phi(X_n - \mu) + Z_{n+1}\}|X_n, X_{n-1}, \ldots] \]
\[ = \mu + \phi(X_n - \mu) \]

The MMSE forecast for $X_{n+2}$ for lead time $l = 2$ is

\[ X_n(2) = E[X_{n+2}|X_n, X_{n-1}, \ldots] \]
\[ = E[\{\mu + \phi(X_{n+1} - \mu) + Z_{n+2}\}|X_n, X_{n-1}, \ldots] \]
\[ = \mu + \phi(X_n(1) - \mu) = \mu + \phi^2(X_n - \mu) \]

The $l$-step ahead forecast has the form

\[ X_n(l) = \mu + \phi^l(X_n - \mu) \]

Note that eventually (as $l \to \infty$), the MMSE forecast from an AR(1) process is the process mean $\mu$. 

We will wish to study the accuracy of the forecasts and to construct $100(1 - \alpha)\%$ prediction intervals for the unknown $X_{n+l}$ for $l = 1, \cdots, L$.

The $l$-step-ahead forecast error corresponding to the MMSE forecast from an ARMA process is

$$e_n(l) = X_{n+l} - X_n(l) = Z_{n+l} + \psi_1 Z_{n+l-1} + \cdots + \psi_{l-1} Z_{n+1}$$
The variance of the $l$-step ahead forecast error is

$$Var[e_n(l)] = \sigma^2[1 + \psi_1^2 + \cdots + \psi_{l-1}^2]$$

For $l = 1, \cdots, L$, the $100(1 - \alpha)\%$ prediction interval for $X_{n+l}$ is given by

$$X_n(l \pm z_{\alpha/2}[Var(e_n(l))]^{1/2}$$

The one-step-ahead forecast for $X_{n+1}$ from origin $n$ is then $X_n(1)$, and the one-step-ahead forecast error is $e_n(1)$.

B&D (section 2.5) refer to the MMSE forecast for $X_{n+l}$ from origin $n$ by $P_nX_{n+l}$. 
Note: These are the forms of the forecasts, forecast errors and forecast error variances. In practice, we will substitute estimated values of the process parameters.

Also note that since an ARMA process admits an MA representation and an AR representation, we can express the forecast $X_n(l)$ in terms of the $\psi$ weights and past random shocks or in terms of the $\pi$ weights and past observations.
It is useful to have formulas for updating forecasts as new observations become available over time. We know that the MMSE forecast for $X_{n+l+1}$ standing at origin $n$ and with forecast lead time $l + 1$ is

$$X_n(l + 1) = \psi_{l+1}Z_n + \psi_{l+2}Z_{n-1} + \cdots.$$ 

Suppose the next observation $X_{n+1}$ becomes available. We can update the prediction for $X_{n+l+1}$ via

$$X_{n+1}(l) = \psi_lZ_{n+1} + \psi_{l+1}Z_n + \psi_{l+2}Z_{n-1} + \cdots.$$ 

That is,

$$X_{n+1}(l) = X_n(l + 1) + \psi_lZ_{n+1} = X_n(l + 1) + \psi_l[X_{n+1} - X_n(1)]$$ 

by the definition of the one-step ahead forecast errors.
The updated forecast for $X_{n+l+1}$ is thus a linear combination of the $l + 1$-step-ahead forecast for $X_{n+l+1}$ made from origin $n$ and the one-step-ahead forecast error $e_n(1) = X_{n+1} - X_n(1) = Z_{n+1}$. 
Example: AR(1) Process (continued): The forecast errors are

\[ e_n(1) = X_{n+1} - X_n(1) \]
\[ = \mu + \phi(X_n - \mu) + Z_{n+1} - \mu - \phi(X_n - \mu) \]
\[ = Z_{n+1}; \]
\[ e_n(2) = X_{n+2} - X_n(2) \]
\[ = \mu + \phi(X_{n+1} - \mu) + Z_{n+2} - \mu - \phi^2(X_n - \mu) \]
\[ = Z_{n+2} + \phi[X_{n+1} - \mu - \phi(X_n - \mu)] \]
\[ = Z_{n+2} + \phi Z_{n+1}; \text{ etc.} \]

or in general,

\[ e_n(l) = Z_{n+l} + \phi Z_{n+l-1} + \cdots + \phi^{l-1} Z_{n+1} \]
The variance of the \( l \)-step ahead forecast error is

\[
\text{Var}[e_n(l)] = \sigma^2[1 + \phi^2 + \cdots + \phi^{2(l-1)}] = \sigma^2 \frac{1 - \phi^{2l}}{1 - \phi^2}
\]

which can be derived directly or by substituting \( \psi_j = \phi^j \). Once we estimate the process parameters (Chapter 5 in B&D), we will substitute the estimates in place of the parameters \( \mu \), \( \phi \) and \( \sigma^2 \).
Example: AR(2) process: Suppose $\mu = 0$. The process is

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + Z_t.$$ 

The one-step-ahead forecast for $X_{n+1}$ from origin $n$ is

$$X_n(1) = E[(\phi_1 X_n + \phi_2 X_{n-1} + Z_{n+1})|X_n, X_{n-1}, \cdots]$$

$$= \phi_1 X_n + \phi_2 X_{n-1}.$$ 

while the two-step-ahead forecast for $X_{n+2}$ from origin $n$ is

$$X_n(2) = E[(\phi_1 X_{n+1} + \phi_2 X_n + Z_{n+2})|X_n, X_{n-1}, \cdots]$$

$$= \phi_1 X_n(1) + \phi_2 X_n.$$ 

The general form for the $l$-step-ahead forecast for $X_{n+l}$ from origin $n$ is the linear difference equation

$$X_n(l) = \phi_1 X_n(l-1) + \phi_2 X_n(l-2).$$
Since we know that for the AR(2) model, the \( \psi \) weights are

\[
\psi_1 = \phi_1; \psi_2 = \phi_1^2 + \phi_2; \psi_j = \phi_1 \psi_{j-1} + \phi_2 \psi_{j-2}
\]

for \( j \geq 2 \),

we can compute the \( l \)-step-ahead forecast error and the corresponding forecast error variance using the general formula.
Innovations: Suppose the observations are \( X_1, X_2, \ldots, X_n \). Suppose the one-step-ahead forecasts for \( X_1, X_2, \ldots, X_n \) are written as \( X_0(1), X_1(1), X_2(1), \ldots, X_{n-1}(1) \), with \( X_0(1) = 0 \).

The one-step-ahead forecast errors are then \( e_0(1), e_1(1), e_2(1), \ldots, e_{n-1}(1) \), with \( e_0(1) = X_1 \). These one-step-ahead forecast errors are called **innovations** and may be denoted by \( U_1, \ldots, U_n \) (see B&D, page 71).