Integrated Moving Averages

The Integrated Moving Average (IMA) is often a useful model for economic time series. It is related to "exponential smoothing", a simple method for forecasting time series, which will be discussed later in more detail. An integrated moving average is simply an ARIMA model with $p=0$. That is, the $IMA(d,q)$ model is the same as the $ARIMA(0,d,q)$. The reason for the name "integrated moving average" should be clear: The $IMA(d,q)$ is a moving average which has been integrated d times. Here, we will study the simplest case, the $IMA(1,1)$, also known as $ARIMA(0,1,1)$. The model can be written as

$$x_t - x_{t-1} = \varepsilon_t - a\,\varepsilon_{t-1} \ ,$$

where a is between -1 and 1 (because of the invertibility condition). Since $d=1$, the series $\{x_t\}$ is nonstationary. So strictly speaking, the series has no mean. Nevertheless, it is useful to think of $\{x_t\}$ as fluctuating about a **local mean,** $\bar{x}_t$ which changes with t. If we define $\alpha = 1-a$, then it can be shown that $\{x_t\}$ has the $AR(\infty)$ representation

$$x_t = \alpha \sum_{k=1}^{\infty} (1-\alpha)^{k-1} x_{t-k} + \varepsilon_t \ , \tag{1}$$

which is the same as saying that

$$x_t = \bar{x}_{t-1} + \varepsilon_t \ , \tag{2}$$

where

$$\bar{x}_{t-1} = \alpha \sum_{k=1}^{\infty} (1-\alpha)^{k-1} x_{t-k} \tag{3}$$

is the **local mean** at time $t-1$. We see from (3) that the local mean, $\bar{x}_{t-1}$ is an **Exponentially Weighted "Moving Average" (EWMA)** of previous values of x_t with weights

$$\alpha \ , \ \alpha(1-\alpha) \ , \ \alpha(1-\alpha)^2 \ , \ \alpha(1-\alpha)^3 \ \cdots ,$$

which decay towards zero geometrically, that is, exponentially fast. It is also interesting to note that these weights sum to 1, since

$$\alpha[1 + (1-\alpha) + (1-\alpha)^2 + \cdots] = \alpha\left(\frac{1}{1-(1-\alpha)}\right) = \frac{\alpha}{\alpha} = 1 \ ,$$

where we have used the formula for the sum of a geometric series.

It can also be shown that $\{x_t\}$ has the $MA(\infty)$ representation

$$x_t = \varepsilon_t + \alpha \sum_{k=1}^{\infty} \varepsilon_{t-k} \quad . \tag{4}$$

Since from (2) we know that $x_t = \bar{x}_{t-1} + \varepsilon_t$, it follows that

$$\bar{x}_{t-1} = \alpha \sum_{k=1}^{\infty} \varepsilon_{t-k} \quad . \tag{5}$$

Forecasting

Since $x_{t+1} = \bar{x}_t + \varepsilon_{t+1}$ from Equation (2), the best one-step forecast is $\bar{x}_t$, the current local mean. This can be contrasted with the case of the random walk, where the best forecast is the most recent *obserevation*, x_t. It can be shown that

$$\bar{x}_t = \alpha x_t + (1-\alpha)\,\bar{x}_{t-1} \quad . \tag{6}$$

Thus, each new local mean is a compromise (weighted average) of the previous local mean and the most recent observation. Formula (6) shows how the new observation x_t influences the value of the local mean, and is very useful for forming forecasts recursively in "real time": As new observations become available, we can simply update our local mean, and thereby obtain the new one-step forecast, without doing any long calculations.

To obtain h-step forecasts, we note from the $MA(\infty)$ representation (4) that

$$x_{t+h} = \varepsilon_{t+h} + \alpha \sum_{k=1}^{\infty} \varepsilon_{t+h-k} \quad ,$$

so the best h-step forecast is

$$f_{t,h} = \alpha(\varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2} + \cdots) = \alpha \sum_{k=1}^{\infty} \varepsilon_{t+1-k} = \bar{x}_t \quad ,$$

where we have used equation (5) for the last step. Thus, we have shown that

$$f_{t,h} = \bar{x}_t \quad ,$$

so that the best forecast for *any* lead time is just the current local mean, $\bar{x}_t$.

To see this in another way, note from Equations (4) and (5) that

$$x_{t+h} - \bar{x}_t = \varepsilon_{t+h} + \alpha \sum_{k=1}^{\infty} \varepsilon_{t+h-k} - \alpha \sum_{k=1}^{\infty} \varepsilon_{t+1-k}$$

$$= \varepsilon_{t+h} + \alpha [\varepsilon_{t+h-1} + \varepsilon_{t+h-2} + \cdots + \varepsilon_{t+1}]$$

$$= \varepsilon_{t+h} + \alpha \sum_{k=1}^{h-1} \varepsilon_{t+k} \quad .$$

Thus,

$$x_{t+h} = \bar{x}_t + \alpha \sum_{k=1}^{h-1} \varepsilon_{t+k} + \varepsilon_{t+h} \quad ,$$

so that as we move into the future from time t (by letting h increase), the process will diverge from the current local mean $\bar{x}_t$ according to the "random walk"

$$\alpha \sum_{k=1}^{h-1} \varepsilon_{t+k} + \varepsilon_{t+h} \quad ,$$

which has zero mean and is not forecastable.