Introduction to Time Series Analysis. Lecture 10.

Peter Bartlett

Forecasting.

1. The innovations representation.

2. Recursive method: Innovations algorithm.

Review: One-step-ahead linear prediction

$$X_{n+1}^n = \phi_{n1} X_n + \phi_{n2} X_{n-1} + \cdots + \phi_{nn} X_1$$

$$\Gamma_n \phi_n = \gamma_n,$$

$$P_{n+1}^n = \mathrm{E}\left(X_{n+1} - X_{n+1}^n\right)^2 = \gamma(0) - \gamma_n' \Gamma_n^{-1} \gamma_n,$$

$$\Gamma_n = \begin{bmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(n-1) \\ \gamma(1) & \gamma(0) & & \gamma(n-2) \\ \vdots & & \ddots & \vdots \\ \gamma(n-1) & \gamma(n-2) & \cdots & \gamma(0) \end{bmatrix},$$

$$\phi_n = (\phi_{n1}, \phi_{n2}, \ldots, \phi_{nn})', \quad \gamma_n = (\gamma(1), \gamma(2), \ldots, \gamma(n))'.$$

The innovations representation

Instead of writing the best linear predictor as

$$X_{n+1}^{n} = \phi_{n1} X_n + \phi_{n2} X_{n-1} + \cdots + \phi_{nn} X_1,$$

we can write

$$X_{n+1}^{n} = \theta_{n1} \underbrace{\left(X_n - X_n^{n-1}\right)}_{\text{innovation}} + \theta_{n2} \left(X_{n-1} - X_{n-1}^{n-2}\right) + \cdots + \theta_{nn} \left(X_1 - X_1^{0}\right).$$

This is still linear in $X_1, \ldots, X_n$.

The innovations are uncorrelated:
$\mathrm{Cov}(X_j - X_j^{j-1}, X_i - X_i^{i-1}) = 0$ for $i \neq j$.

Comparing representations: $U_n = X_n - X_n^{n-1}$ versus X_n

$$\begin{pmatrix} U_1 \\ U_2 \\ \vdots \\ U_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ -\phi_{11} & 1 & & 0 \\ \vdots & & \ddots & \\ -\phi_{n-1,n-1} & -\phi_{n-1,n-2} & \cdots & 1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}$$

$$\begin{pmatrix} X_1^0 \\ X_2^1 \\ \vdots \\ X_n^{n-1} \end{pmatrix} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \theta_{11} & 0 & & 0 \\ \vdots & & \ddots & \\ \theta_{n-1,n-1} & \theta_{n-1,n-2} & \cdots & 0 \end{pmatrix} \begin{pmatrix} U_1 \\ U_2 \\ \vdots \\ U_n \end{pmatrix}$$

Innovations Algorithm

$$X_1^0 = 0, \qquad X_{n+1}^n = \sum_{i=1}^{n} \theta_{ni} \left(X_{n+1-i} - X_{n+1-i}^{n-i} \right).$$

NB: error in text.

$$\theta_{n,n-i} = \frac{1}{P_{i+1}^i} \left(\gamma(n-i) - \sum_{j=0}^{i-1} \theta_{i,i-j} \theta_{n,n-j} P_{j+1}^j \right).$$

$$P_1^0 = \gamma(0) \qquad P_{n+1}^n = \gamma(0) - \sum_{i=0}^{n-1} \theta_{n,n-i}^2 P_{i+1}^i.$$

Innovations Algorithm: Example

$$\theta_{n,n-i} = \frac{1}{P_{i+1}^i}\left(\gamma(n-i) - \sum_{j=0}^{i-1}\theta_{i,i-j}\theta_{n,n-j}P_{j+1}^j\right).$$

$$P_1^0 = \gamma(0) \qquad P_{n+1}^n = \gamma(0) - \sum_{i=0}^{n-1}\theta_{n,n-i}^2 P_{i+1}^i.$$

$$\theta_{1,1} = \gamma(1)/P_1^0, \qquad P_2^1 = \gamma(0) - \theta_{1,1}^2 P_1^0$$

$$\theta_{2,2} = \gamma(2)/P_1^0, \quad \theta_{2,1} = \left(\gamma(1) - \theta_{1,1}\theta_{2,2}P_1^0\right)/P_2^1,$$

$$P_3^2 = \gamma(0) - \left(\theta_{2,2}^2 P_1^0 + \theta_{2,1}^2 P_2^1\right)$$

$$\theta_{3,3}, \quad \theta_{3,2}, \quad \theta_{3,1}, \quad P_4^3, \ldots$$

Predicting h steps ahead using innovations

The innovations representation for the one-step-ahead forecast is

$$P(X_{n+1}|X_1, \ldots, X_n) = \sum_{i=1}^{n} \theta_{ni} \left(X_{n+1-i} - X_{n+1-i}^{n-i}\right),$$

What is the innovations representation for $P(X_{n+h}|X_1, \ldots, X_n)$?

Fact: If $h \geq 1$ and $1 \leq i \leq n$, we have
$\mathrm{Cov}(X_{n+h} - P(X_{n+h}|X_1, \ldots, X_{n+h-1}), X_i) = 0.$

Thus, $P(X_{n+h} - P(X_{n+h}|X_1, \ldots, X_{n+h-1})|X_1, \ldots, X_n) = 0.$
That is, the best prediction of X_{n+h} is the
best prediction of the one-step-ahead forecast of X_{n+h}.

Predicting h steps ahead using innovations

$$
\begin{aligned}
& P(X_{n+h}|X_1,\ldots,X_n) \\
&= P\left(P(X_{n+h}|X_1,\ldots,X_{n+h-1})|X_1,\ldots,X_n\right) \\
&= P\left(\sum_{i=1}^{n+h-1} \theta_{n+h-1,i}\left(X_{n+h-i} - X_{n+h-i}^{n+h-i+1}\right)|X_1,\ldots,X_n\right) \\
&= \sum_{i=1}^{n+h-1} \theta_{n+h-1,i} P\left(\left(X_{n+h-i} - X_{n+h-i}^{n+h-i+1}\right)|X_1,\ldots,X_n\right) \\
&= \sum_{i=h}^{n+h-1} \theta_{n+h-1,i} P\left(\left(X_{n+h-i} - X_{n+h-i}^{n+h-i+1}\right)|X_1,\ldots,X_n\right) \\
&= \sum_{i=h}^{n+h-1} \theta_{n+h-1,i}\left(X_{n+h-i} - X_{n+h-i}^{n+h-i+1}\right)
\end{aligned}
$$

Predicting h steps ahead using innovations

$$P(X_{n+1}|X_1,\ldots,X_n) = \sum_{i=1}^{n} \theta_{ni} \left(X_{n+1-i} - X_{n+1-i}^{n-i}\right)$$

$$P(X_{n+h}|X_1,\ldots,X_n) = \sum_{j=h}^{n+h-1} \theta_{n+h-1,j} \left(X_{n+h-j} - X_{n+h-j}^{n+h-j+1}\right)$$

$$= \sum_{i=1}^{n} \theta_{n+h-1,h-1+i} \left(X_{n+1-i} - X_{n+1-i}^{n-i}\right)$$

$(j = i + h - 1)$

Mean squared error of h-step-ahead forecasts

From orthogonality of the predictors and the error,

$\mathrm{E}\left((X_{n+h} - P(X_{n+h}|X_1,\ldots,X_n))\,P(X_{n+h}|X_1,\ldots,X_n)\right) = 0.$

That is, $\mathrm{E}\left(X_{n+h}P(X_{n+h}|X_1,\ldots,X_n)\right) = \mathrm{E}\left(P(X_{n+h}|X_1,\ldots,X_n)^2\right)$.

Hence, we can express the mean squared error as

$$
\begin{aligned}
P_{n+h}^{n} &= \mathrm{E}\left(X_{n+h} - P(X_{n+h}|X_1,\ldots,X_n)\right)^2 \\
&= \gamma(0) + \mathrm{E}\left(P(X_{n+h}|X_1,\ldots,X_n)\right)^2 \\
&\qquad - 2\mathrm{E}\left(X_{n+h}P(X_{n+h}|X_1,\ldots,X_n)\right) \\
&= \gamma(0) - \mathrm{E}\left(P(X_{n+h}|X_1,\ldots,X_n)\right)^2 .
\end{aligned}
$$

Mean squared error of h-step-ahead forecasts

But the innovations are uncorrelated, so

$$\begin{aligned}
P_{n+h}^{n} &= \gamma(0) - \mathrm{E}\left(P(X_{n+h}|X_1,\ldots,X_n)\right)^2 \\
&= \gamma(0) - \mathrm{E}\left(\sum_{j=h}^{n+h-1} \theta_{n+h-1,j}\left(X_{n+h-j} - X_{n+h-j}^{n+h-j-1}\right)\right)^2 \\
&= \gamma(0) - \sum_{j=h}^{n+h-1} \theta_{n+h-1,j}^2 \ \mathrm{E}\left(X_{n+h-j} - X_{n+h-j}^{n+h-j-1}\right)^2 \\
&= \gamma(0) - \sum_{j=h}^{n+h-1} \theta_{n+h-1,j}^2 \ P_{n+h-j}^{n+h-j-1}.
\end{aligned}$$