Univariate time series

Optimal linear forecast

V. Lefieux

Table of contents

Table of contents

Optimal linear forecast

Let $(X_t)_{t\in\mathbb{Z}}$ be a stationary process.
Let $\mathcal{H}_{-\infty}^{t}(X)$ be the space be the space spanned by linear combinations of $(X_i)_{i\leq t}$ and 1.
The (one step) optimal linear forecast de X_t given its past is :

$$\widehat{X}_t = \mathbb{E}\left(X_t \,/ \mathcal{H}_{-\infty}^{t-1}(X)\right).$$

The one step optimal linear forecast errors $\varepsilon_t = X_t - \widehat{X}_t$ are called innovations.

Proposition

The innovations process is a white noise.

Table of contents

Method

For a stationary process $(X_t)_{t\in\mathbb{N}}$, we use recursive algorithms in order calculate the forecast of X_{T+1} based on $(X_1, \ldots, X_T)$.

With the Durbin-Levinson algorithm : one step forecast 1/3

The one step forecast of X_{T+1} based on $(X_1, \ldots, X_T)$ is :

$$\widehat{X}_T(1) = \widehat{X}_{T+1} = \sum_{i=1}^{T} a_i(T) X_{T+1-i}$$

where :

$$\begin{pmatrix} a_1(T) \\ \vdots \\ a_T(T) \end{pmatrix} = R_T^{-1} \begin{pmatrix} \rho(1) \\ \vdots \\ \rho(T) \end{pmatrix}.$$

With the Durbin-Levinson algorithm : one step forecast 2/3

The one step forecast of X_{T+2} based on $(X_1, \ldots, X_{T+1})$ is :

$$\widehat{X}_{T+2} = \sum_{i=1}^{T+1} a_i\,(T+1)\,X_{T+2-i}$$

where :

$$\begin{pmatrix} a_1\,(T+1) \\ \vdots \\ a_{T+1}\,(T+1) \end{pmatrix} = R_{T+1}^{-1} \begin{pmatrix} \rho\,(1) \\ \vdots \\ \rho\,(T+1) \end{pmatrix}.$$

With the Durbin-Levinson algorithm : one step forecast 3/3

We can use the Durbin-Levinson algorithm in order to obtain coefficients $\left(a_1\left(T+1\right),\ldots,a_{T+1}\left(T+1\right)\right)$ based on coefficients $\left(a_1\left(h\right),\ldots,a_h\left(h\right)\right)_{h\in\{1,\ldots,T\}}$.
Moreover there is a relationship between mean squared errors. With :

$$\begin{aligned} v_T &= \mathbb{E}\left[\left(X_{T+1}-\widehat{X}_{T+1}\right)^2\right] \\ &= \mathbb{E}\left[\left(X_{T+1}-\sum_{i=1}^{T} a_i(T)X_{T+1-i}\right)^2\right] \end{aligned}$$

we have :

$$v_{T+1} = v_T\left[1-\left(a_{T+1}\left(T+1\right)\right)^2\right] = v_T\left[1-r^2\left(T+1\right)\right].$$

With the Durbin-Levinson algorithm : multiple step forecast

The one step forecast of X_{T+k}, with $k \in \{2, \ldots, T-1\}$, based on $(X_1, \ldots, X_T)$ is :

$$\widehat{X}_T(k) = \sum_{i=1}^{k-1} a_i(T)\widehat{X}_T(k-i) + \sum_{i=k}^{T} a_i(T) X_{T+k-i}.$$

Notice that one must have $k \ll T$.

Introducing the innovation algorithm 1/4

The one step forecast of X_{T+1} based on $(X_1, \ldots, X_T)$ is :

$$\widehat{X}_T(1) = \widehat{X}_{T+1} = \sum_{i=1}^{T} a_i(T) X_{T+1-i}.$$

Innovations are :

$$\varepsilon_{T+1} = X_{T+1} - \widehat{X}_{T+1} = X_{T+1} - \sum_{i=1}^{T} a_i(T) X_{T+1-i}.$$

Introducing the innovation algorithm 2/4

We have :

$$\begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_{T+1} \end{pmatrix} = A_{T+1} \begin{pmatrix} X_1 \\ \vdots \\ X_{T+1} \end{pmatrix}$$

with :

$$A_{T+1} = \begin{bmatrix} 1 & 0 & \ldots & \ldots & 0 \\ -a_1(1) & 1 & 0 & \ldots & 0 \\ -a_2(2) & -a_1(2) & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ -a_t(T) & -a_{T-1}(T) & \ldots & -a_1(T) & 1 \end{bmatrix}.$$

Introducing the innovation algorithm 3/4

So :

$$\begin{pmatrix} X_1 \\ \vdots \\ X_{T+1} \end{pmatrix} = A_{T+1}^{-1} \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_{T+1} \end{pmatrix} = C_{T+1} \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_{T+1} \end{pmatrix}$$

where :

$$C_{T+1} = \begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ \theta_1(1) & 1 & 0 & \dots & 0 \\ \theta_2(2) & \theta_1(2) & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \theta_t(T) & \theta_{T-1}(T) & \dots & \theta_1(T) & 1 \end{bmatrix}.$$

Introducing the innovation algorithm 4/4

Finaly :

$$\begin{pmatrix} \widehat{X}_1 \\ \vdots \\ \widehat{X}_{T+1} \end{pmatrix} = \begin{pmatrix} X_1 \\ \vdots \\ X_{T+1} \end{pmatrix} - \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_{T+1} \end{pmatrix}$$

$$= [C_{T+1} - I_{T+1}] \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_{T+1} \end{pmatrix}$$

$$= [C_{T+1} - I_{T+1}] \left[\begin{pmatrix} X_1 \\ \vdots \\ X_{T+1} \end{pmatrix} - \begin{pmatrix} \widehat{X}_1 \\ \vdots \\ \widehat{X}_{T+1} \end{pmatrix} \right].$$

The innovation algorithm

With the notation : $v_h = \mathbb{E}\left[\left(X_{h+1} - \widehat{X}_{h+1}\right)^2\right]$, we have :

- $v_0 = \gamma(0)$,
- $\forall i \in \{1, \ldots, h-1\}$:

$$\theta_{h-i}(h) = \frac{1}{v_i}\left[\gamma(h-i) - \sum_{j=0}^{i-1} \theta_{i-j}(i)\,\theta_{h-j}(h)\,v_j\right],$$

- $v_h = \gamma(h+1) - \sum_{i=0}^{h-1} \theta_{h-i}^2(h)\,v_i$.

Table of contents

Definition

Let $(X_t)_{t\in\mathbb{Z}}$ be a stationary process.
$(X_t)_{t\in\mathbb{Z}}$ is a regular (or non deterministic) process if :

$$\mathbb{E}\left[\left(X_t - \widehat{X}_t\right)^2\right] > 0.$$

$(X_t)_{t\in\mathbb{Z}}$ is a singular (or deterministic) process if :

$$\mathbb{E}\left[\left(X_t - \widehat{X}_t\right)^2\right] = 0.$$

Wold decomposition

If $(X_t)_{t\in\mathbb{Z}}$ is a non deterministic stationary process then there exist a unique decomposition (named Wold decomposition) :

$$\forall t \in \mathbb{Z} : X_t = \varepsilon_t + \sum_{i=1}^{+\infty} \psi_i \varepsilon_{t-i} + V_t$$

where :

- $(\varepsilon_t)_{t\in\mathbb{Z}}$ is a white noise,
- $(V_t)_{t\in\mathbb{Z}}$ is a deterministic process, uncorrelated with $(X_t)_{t\in\mathbb{Z}}$,
- $\sum_{i=1}^{+\infty} \psi_i^2 < +\infty$.