Financial Time Series – Forecasting

Andreas Petersson

TMS088/MSA410 – April 2020

Mathematical Sciences, Chalmers University of Technology & University of Gothenburg, Sweden

Forecasting

Predict $(X_{n+h}, h > 0)$ in terms of $(X_t, t = 1, \dots, n)$. What is a good prediction?

Definition

Let X and Y be random variables and let Y be an approximation of X. The *mean squared error* of Y is defined by

$$\mathrm{MSE}(Y, X) := \mathbb{E}((Y - X)^2).$$

The best predictor

Definition

Let $(X_t, t \in \mathbb{Z})$ be a time series with $\text{Var}(X_t) < \infty$ for all $t \in \mathbb{Z}$ and $X^n := (X_{t_1}, \ldots, X_{t_n})$ a collection of random variables of the time series at n different times. Then the function of X^n denoted by $b_t(X^n)$ is called a *best predictor* of X_t for some $t \in \mathbb{Z}$, if it minimizes the mean squared error, i.e.,

$$b_t(X^n) := \underset{g(X^n)}{\arg\min}\, \text{MSE}(g(X^n), X_t) = \underset{g(X^n)}{\arg\min}\, \mathbb{E}((g(X^n) - X_t)^2),$$

where the minimum is taken over all measurable functions $g : \mathbb{R}^n \to \mathbb{R}$.

The best predictor is the conditional expectation

Proposition

Let $(X_t, t \in \mathbb{Z})$ *be a time series with* $\mathsf{Var}(X_t) < \infty$ *for all* $t \in \mathbb{Z}$ *and* $X^n := (X_{t_1}, \ldots, X_{t_n})$ *a collection of random variables of the time series at* n *different times. Then the best predictor of* X_t *for some* $t \in \mathbb{Z}$ *is the conditional expectation of* X_t *given* X^n*, i.e.,*

$$b_t(X^n) = \mathbb{E}(X_t | X^n).$$

Proof.

$\square$

The best predictor is the conditional expectation

Proposition

Let $(X_t, t \in \mathbb{Z})$ *be a time series with* $\mathrm{Var}(X_t) < \infty$ *for all* $t \in \mathbb{Z}$ *and* $X^n := (X_{t_1}, \ldots, X_{t_n})$ *a collection of random variables of the time series at* n *different times. Then the best predictor of* X_t *for some* $t \in \mathbb{Z}$ *is the conditional expectation of* X_t *given* X^n*, i.e.,*

$$b_t(X^n) = \mathbb{E}(X_t | X^n).$$

Proof.

$$\begin{aligned}
\mathbb{E}((g(X^n) - X_t)^2) &= \mathbb{E}((g(X^n) - \mathbb{E}(X_t|X^n) + \mathbb{E}(X_t|X^n) - X_t)^2) \\
&= \mathbb{E}((g(X^n) - \mathbb{E}(X_t|X^n))^2) + \mathbb{E}((\mathbb{E}(X_t|X^n) - X_t)^2) \\
&\quad + 2\,\mathbb{E}((g(X^n) - \mathbb{E}(X_t|X^n))(\mathbb{E}(X_t|X^n) - X_t)).
\end{aligned}$$

$\square$

The best predictor is the conditional expectation

Proposition

Let $(X_t, t \in \mathbb{Z})$ *be a time series with* $\mathsf{Var}(X_t) < \infty$ *for all* $t \in \mathbb{Z}$ *and* $X^n := (X_{t_1}, \ldots, X_{t_n})$ *a collection of random variables of the time series at* n *different times. Then the best predictor of* X_t *for some* $t \in \mathbb{Z}$ *is the conditional expectation of* X_t *given* X^n*, i.e.,*

$$b_t(X^n) = \mathbb{E}(X_t | X^n).$$

Proof.

$$\mathbb{E}((g(X^n) - \mathbb{E}(X_t|X^n))(\mathbb{E}(X_t|X^n) - X_t)) =$$

$\square$

The best predictor is the conditional expectation

Proposition

Let $(X_t, t \in \mathbb{Z})$ be a time series with $\mathsf{Var}(X_t) < \infty$ for all $t \in \mathbb{Z}$ and $X^n := (X_{t_1}, \ldots, X_{t_n})$ a collection of random variables of the time series at n different times. Then the best predictor of X_t for some $t \in \mathbb{Z}$ is the conditional expectation of X_t given X^n, i.e.,

$$b_t(X^n) = \mathbb{E}(X_t | X^n).$$

Proof.

$$\begin{aligned}
\min_{g(X^n)} \mathbb{E}((g(X^n) - X_t)^2) &= \min_{g(X^n)} \left(\mathbb{E}((g(X^n) - \mathbb{E}(X_t|X^n))^2) + \mathbb{E}((\mathbb{E}(X_t|X^n) - X_t)^2)\right) \\
&= \mathbb{E}((\mathbb{E}(X_t|X^n) - X_t)^2) + \min_{g(X^n)} \mathbb{E}((g(X^n) - \mathbb{E}(X_t|X^n))^2).
\end{aligned}$$

□

The best linear predictor

Definition

Let $(X_t, t \in \mathbb{Z})$ be a time series with $\mathrm{Var}(X_t) < \infty$ for all $t \in \mathbb{Z}$ and $X^n := (X_{t_1}, \ldots, X_{t_n})$ a collection of random variables of the time series at n different times. Then the linear function of 1 and X^n denoted by $b_t^l(X^n)$ is called a *best linear predictor* of X_t for some $t \in \mathbb{Z}$ if it minimizes the mean squared error, i.e.,

$$b_t^l(X^n) := \underset{g(X^n)}{\arg\min}\, \mathrm{MSE}(g(X^n), X_t) = \underset{g(X^n)}{\arg\min}\, \mathbb{E}((g(X^n) - X_t)^2),$$

where the minimum is taken over all linear functions g of 1 and X^n, i.e., for all functions g such that there exist $a_0, \ldots, a_n \in \mathbb{R}$ such that $g(X^n) := a_0 + a_1 X_{t_n} + a_2 X_{t_{n-1}} + \cdots + a_n X_{t_1}$.

Note: In BD, when $X^n := (X_1, \ldots, X_n)$, $b_{n+h}^l(X^n)$ is denoted by $P_n X_{n+h}$ and $b_{n+1}^l(X^n)$ is denoted by $\hat{X}_{n+1}$

Existence of $b_t^l(X^n)$

- There exists a minimum $a = (a_0, a_1, \ldots, a_n) \in \mathbb{R}^{n+1}$ to S, where $S(a) := \mathbb{E}((a_0 + a_1 X_{t_n} + \cdots + a_n X_{t_1} - X_t)^2)$

Existence of $b_t^l(X^n)$

- There exists a minimum $a = (a_0, a_1, \ldots, a_n) \in \mathbb{R}^{n+1}$ to S, where $S(a) := \mathbb{E}((a_0 + a_1 X_{t_n} + \cdots + a_n X_{t_1} - X_t)^2)$
- The prediction equations:
  - $\mathbb{E}(X_t - b_t^l(X^n)) = 0$

Existence of $b_t^l(X^n)$

- There exists a minimum $a = (a_0, a_1, \ldots, a_n) \in \mathbb{R}^{n+1}$ to S, where $S(a) := \mathbb{E}((a_0 + a_1 X_{t_n} + \cdots + a_n X_{t_1} - X_t)^2)$
- The prediction equations:
  - $\mathbb{E}(X_t - b_t^l(X^n)) = 0$
  - $\mathbb{E}\left(\left(X_t - b_t^l(X^n)\right) X_{t_{n+1-j}}\right) = 0$ for $j = 1, \ldots, n$

Existence of $b_t^l(X^n)$

- There exists a minimum $a = (a_0, a_1, \ldots, a_n) \in \mathbb{R}^{n+1}$ to S, where $S(a) := \mathbb{E}((a_0 + a_1 X_{t_n} + \cdots + a_n X_{t_1} - X_t)^2)$
- The prediction equations:
  - $\mathbb{E}(X_t - b_t^l(X^n)) = 0$
  - $\mathbb{E}\left(\left(X_t - b_t^l(X^n)\right) X_{t_{n+1-j}}\right) = 0$ for $j = 1, \ldots, n$
- The predictor is unique: Let $a^{(1)}$ and $a^{(2)}$ be two minima with corresponding predictors $b_t^{(l,1)}(X^n)$ and $b_t^{(l,2)}(X^n)$. Set $Z := b_t^{(l,1)}(X^n) - b_t^{(l,2)}(X^n)$

Prediction proposition

Proposition

Let $(X_t, t \in \mathbb{Z})$ *be a time series with* $\mathsf{Var}(X_t) < \infty$ *for all* $t \in \mathbb{Z}$ *and* $X^n := (X_{t_1}, \ldots, X_{t_n})$ *a collection of random variables of the time series at* n *different times. Then the best linear predictor of* X_t *is given by*

$$b_t^l(X^n) = a_0 + a_1 X_{t_n} + a_2 X_{t_{n-1}} + \cdots + a_n X_{t_1},$$

where the coefficients $(a_i, i = 0, \ldots, n)$ *are determined by the linear equations*

1. $\mathbb{E}(X_t - b_t^l(X^n)) = 0$,
2. $\mathbb{E}(X_{t_j}(X_t - b_t^l(X^n))) = 0$ *for all* $j = 1, \ldots, n$.

Corollary

Let $X = (X_t, t \in \mathbb{Z})$ and X^n be as in Proposition 7 and assume in addition that X is stationary with mean μ and autocovariance function γ. Then the coefficients $(a_i, i = 0, \ldots, n)$ of $b_t^l(X^n)$ are determined by the linear equations

$$a_0 = \mu \left(1 - \sum_{i=1}^{n} a_i \right)$$

and

$$\Gamma_n (a_1, \ldots, a_n)' = (\gamma(t - t_n), \ldots, \gamma(t - t_1))'$$

with

$$\Gamma_n = (\gamma(t_{n+1-j} - t_{n+1-i}))_{i,j=1}^{n}.$$

Moreover,

$$\mathrm{MSE}(b_t^l(X^n), X_t) = \gamma(0) - (a_1, \ldots, a_n)(\gamma(t - t_n), \ldots, \gamma(t - t_1))'.$$

Prediction of stationary time series

Prediction of stationary time series

Note 1: When $X^n := (X_1, \ldots, X_n)$ the equations to derive the coefficients $(a_0, \ldots, a_n)$ for $b^l_{n+h}(X^n)$ simplify to

$$a_0 = \mu \left(1 - \sum_{i=1}^{n} a_i \right)$$

and

$$(\gamma(i-j))_{i,j=1}^{n} (a_1, \ldots, a_n)' = (\gamma(h), \ldots, \gamma(h+n-1))'.$$

Prediction of stationary time series

Note 1: When $X^n := (X_1, \ldots, X_n)$ the equations to derive the coefficients $(a_0, \ldots, a_n)$ for $b^l_{n+h}(X^n)$ simplify to

$$a_0 = \mu \left(1 - \sum_{i=1}^{n} a_i \right)$$

and

$$(\gamma(i-j))_{i,j=1}^{n} (a_1, \ldots, a_n)' = (\gamma(h), \ldots, \gamma(h+n-1))'.$$

Note 2: Changing μ affects a_0 only $\rightarrow$ $\mu = 0$ w.l.o.g.

Prediction of stationary time series

Note 1: When $X^n := (X_1, \dots, X_n)$ the equations to derive the coefficients $(a_0, \dots, a_n)$ for $b^l_{n+h}(X^n)$ simplify to

$$a_0 = \mu \left(1 - \sum_{i=1}^{n} a_i \right)$$

and

$$(\gamma(i-j))_{i,j=1}^{n} (a_1, \dots, a_n)' = (\gamma(h), \dots, \gamma(h+n-1))'.$$

Note 2: Changing μ affects a_0 only $\rightarrow \mu = 0$ w.l.o.g.
Note 3: The coefficients of $b^l_{n+h}(X^n)$ are the same as those of $b^l_{t+n+h}((X_{t+1}, \dots, X_{t+n}))$ for all $t \in \mathbb{Z}$

Example: An AR(1) process

Let a stationary time series model be given by

$$X_t - \phi_1 X_{t-1} = Z_t,$$

where $(Z_t, t \in \mathbb{Z}) \sim \mathrm{WN}(0, \sigma^2)$. Let $|\phi_1| < 1 \implies$ existence and that $\mathbb{E}[Z_t X_{t-j}] = 0$ for $j > 0$.

Example: An AR(1) process

Let a stationary time series model be given by

$$X_t - \phi_1 X_{t-1} = Z_t,$$

where $(Z_t, t \in \mathbb{Z}) \sim \mathrm{WN}(0, \sigma^2)$. Let $|\phi_1| < 1 \implies$ existence and that $\mathbb{E}[Z_t X_{t-j}] = 0$ for $j > 0$.

$$\begin{pmatrix} 1 & \phi_1 & \phi_1^2 & \cdots & \phi_1^{n-1} \\ \phi_1 & 1 & \phi_1 & \cdots & \phi_1^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi_1^{n-1} & \phi_1^{n-2} & \phi_1^{n-3} & \cdots & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \phi_1 \\ \phi_1^2 \\ \vdots \\ \phi_1^n \end{pmatrix}$$

Example: An MA(1) process

Let a stationary time series model be given by

$$X_t = Z_t + \theta_1 Z_{t-1},$$

where $(Z_t, t \in \mathbb{Z}) \sim \mathrm{WN}(0, \sigma^2)$.

Example: An MA(1) process

Let a stationary time series model be given by

$$X_t = Z_t + \theta_1 Z_{t-1},$$

where $(Z_t, t \in \mathbb{Z}) \sim \mathrm{WN}(0, \sigma^2)$. $b^l_{n+1}(X^n) = \sum_{i=1}^n a_i X_{n+1-i}$ and $\mathbb{E}((X_{n+1} - \sum_{i=1}^n a_i X_{n+1-i})X_j) = 0$ for $j = 1, \ldots, n$.

Example: AR(1) with a missing value

Let a stationary time series model be given by

$$X_t - \phi_1 X_{t-1} = Z_t,$$

where $(Z_t, t \in \mathbb{Z}) \sim \mathrm{WN}(0, \sigma^2)$ and $|\phi_1| < 1$. We observe X_1 and X_3 but miss X_2. Compute $b_2^l((X_1, X_3)) = a_1 X_3 + a_2 X_1$.

Example: AR(1) with a missing value

Let a stationary time series model be given by

$$X_t - \phi_1 X_{t-1} = Z_t,$$

where $(Z_t, t \in \mathbb{Z}) \sim \mathrm{WN}(0, \sigma^2)$ and $|\phi_1| < 1$. We observe X_1 and X_3 but miss X_2. Compute $b_2^l((X_1, X_3)) = a_1 X_3 + a_2 X_1$.

$$\gamma_X(h) = \frac{\sigma^2 \phi_1^{|h|}}{1 - \phi_1^2}$$

Example: AR(1) with a missing value

Let a stationary time series model be given by

$$X_t - \phi_1 X_{t-1} = Z_t,$$

where $(Z_t, t \in \mathbb{Z}) \sim \mathrm{WN}(0, \sigma^2)$ and $|\phi_1| < 1$. We observe X_1 and X_3 but miss X_2. Compute $b_2^l((X_1, X_3)) = a_1 X_3 + a_2 X_1$.

$$\gamma_X(h) = \frac{\sigma^2 \phi_1^{|h|}}{1 - \phi_1^2}$$

$$\begin{pmatrix} 1 & \phi_1^2 \\ \phi_1^2 & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} \phi_1 \\ \phi_1 \end{pmatrix}$$