Lecture 20

Forecasting Part 1

Arnab Hazra

Minimum MSE predictor

- Suppose we wish to find a prediction function $g(X)$ that minimizes

$$\mathrm{MSE} = E[(Y - g(X))^2],$$

where X and Y are jointly distributed random variables with density function $f(x, y)$. Show that MSE is minimized by the choice

$$g(X) = E(Y|X).$$

- How does the result change when X is a random vector $\boldsymbol{X} = (X_1, \ldots, X_T)$?

Minimum MSE predictor for time series

- In forecasting, the goal is to predict future values of a time series, $X_{T+m}, m = 1, 2, \ldots$, based on the data collected to the present, $\mathcal{X}_{1:T} = \{X_1, X_2, \ldots, X_T\}$.

- Until we discuss estimation, we will assume X_t is stationary and the model parameters are known.

- The minimum mean square error predictor of X_{T+m} is

$$X_{T+m}^T = E(X_{T+m}|\mathcal{X}_{1:T}).$$

Linear predictors

- We will restrict attention to linear predictors of the form

$$X_{T+m}^{T} = \alpha_0 + \sum_{k=1}^{T} \alpha_k X_k.$$

- Linear predictors that minimize the mean square prediction error are called best linear predictors (BLPs).
- If the process is Gaussian, minimum mean square error predictors and best linear predictors are the same.
- BLP for Stationary Processes: Given data $\mathcal{X}_{1:T}$, the best linear predictor is found by solving

$$E[(X_{T+m} - X_{T+m}^{T})X_k] = 0, \quad k = 0, 1, \ldots, T,$$

for $\alpha_0, \alpha_1, \ldots, \alpha_T$. Here $X_0 = 1$.

Nonzero mean case

- If $E(X_t) = \mu$, for $k = 0$, $E(X_{T+m}) = E(X_{T+m}^T) = \mu$.
- Taking expectation in $X_{T+m}^T = \alpha_0 + \sum_{k=1}^T \alpha_k X_k$, we have

$$\mu = \alpha_0 + \sum_{k=1}^{K} \alpha_k \mu.$$

- Thus, $\alpha_0 = \mu(1 - \sum_{k=1}^T \alpha_k)$.
- Hence, the form of the BLP is

$$X_{T+m}^T = \mu + \sum_{k=1}^{T} \alpha_k (X_k - \mu).$$

- WLOG, we will assume $\mu = 0$, in which case, $\alpha_0 = 0$.

One-step-ahead prediction

- The BLP of X_{T+1} is of the form

$$X_{T+1}^{T} = \phi_{T,1} X_T + \phi_{T,2} X_{T-1} + \ldots + \phi_{T,T} X_1 = \sum_{j=1}^{T} \phi_{T,j} X_{T+1-j}.$$

- Here $\alpha_k = \phi_{T,T-k+1}$.
- Thus, the coefficients $\{\phi_{T,1}, \phi_{T,2}, \ldots, \phi_{T,T}\}$ satisfy

$$E\left[\left(X_{T+1} - \sum_{j=1}^{T} \phi_{T,j} X_{T+1-j}\right) X_{T+1-k}\right] = 0, \quad k = 1, \ldots, T.$$

- This implies $\sum_{j=1}^{T} \phi_{T,j}\, \gamma(k-j) = \gamma(k), \quad k = 1, \ldots, T.$
- In matrix notations, $\mathbf{\Gamma}_T\, \boldsymbol{\phi}_T = \boldsymbol{\gamma}_T$ and hence, $\boldsymbol{\phi}_T = \mathbf{\Gamma}_T^{-1} \boldsymbol{\gamma}_T$ and $X_{T+1}^{T} = \boldsymbol{\phi}'_T \tilde{\boldsymbol{X}}$, where $\tilde{\boldsymbol{X}} = (X_T, X_{T-1}, \ldots, X_1)$.

Mean square one-step-ahead prediction error

- Show that the mean square one-step-ahead prediction error is

$$P_{T+1}^T = E[(X_{T+1} - X_{T+1}^T)^2] = \gamma(0) - \gamma_T' \mathbf{\Gamma}_T^{-1} \gamma_T.$$

- Here $P_{T+1}^T = E[(X_{T+1} - \gamma_T' \mathbf{\Gamma}_T^{-1} \tilde{\boldsymbol{X}})^2]$.

- For an AR(2) model $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + W_t$, show that $\phi_{T,1} = \phi_1$ and $\phi_{T,2} = \phi_2$.

Thank you!