Assignment 3: Forecasting

DE09: Time Series Analysis,Srikanth Pai, MSE

1. Try to write clear proofs for "Show that.." questions. Your proofs must start by stating definitions, followed by a sequence of logical inferences and the conclusion must be at the end.
2. For "Give an example...." questions, state the example clearly at the start and then justify why it works afterwards.
3. I do not condone using AI tools blindly. Use it to learn intermediate steps and you will be prepared for the exam.Note that the AI based derivations are often erroneous on a conceptual level. So follow the notes from class.

Maths practice

1. Find the minimum value of $(x-1)^2+(y-2)^2+(z-3)^2+(w-4)^2$ if $x+y+z+w=0$ and $2x+3y+4z+w=0$. Solve it using the vector space method.
2. Define a *generating function* of a sequence $\{a_t\}$ as a power series
$$A(x)=\sum_{t=0}^{\infty} a_t x^t.$$
If $Y_t=\epsilon_t+2\epsilon_{t-1}$ and $Y(x), E(x)$ are the generating functions of sequence Y_t, ϵ_t respectively, show that
$$Y(x)=(1+2x)E(x).$$
You may assume $\epsilon_{-1}=0$.
3. What is Cholesky decomposition? Prove the uniqueness and compute the decomposition for the matrix
$$\begin{bmatrix} 1 & 1 \\ 1 & 4 \end{bmatrix}$$

Forecasting theory

1. Prove that the conditional expectation is the MMSE estimator.
2. Using vector space methods, show that if the MMSE estimator $\hat{Y}$, is a function of the observed data $\mathbf{X}$, then
$$\mathbb{E}[\hat{Y}(Y-\hat{Y})]=0.$$
Note that in vector space parlance, this means optimal forecast error is independent of the optimal forecast.
3. For a jointly normal random vector $(Y, X_1, X_2, ...X_n)$, show that the conditional expectation $\mathbb{E}(Y|\mathbf{X})$ is a linear function of the vector $\mathbf{X}.$
4. Show that if the data for multilinear regression is written in the form
$$\mathbf{y}=\mathbf{X}\beta+\epsilon,$$
then prove that the MMSE estimator (assuming ϵ is white noise) is of the form
$$\hat{\beta}=(\mathbf{X^T X})^{-1}\mathbf{X^T y}.$$

5. Let $\hat{\mathbf{y}} = H\mathbf{y}$ where $H = \mathbf{X}(\mathbf{X^T X})^{-1}\mathbf{X^T}$.
   (a) Show that H is symmetric and idempotent. Conclude $H(I-H) = (I-H)H = 0$.
   (b) Show that $y - \hat{y}$ is perpendicular to $\hat{y}$. Again we see that the error due to optimal forecast is perpendicular to the optimal forecast.

6. Derive the optimal forecast coefficients for time series data in matrix form as

$$\underline{\alpha} = \mathbb{E}(Y_{t+1}\mathbf{X}_t^T)\mathbb{E}(\mathbf{X}_t\mathbf{X}_t^T)^{-1}$$

   [Note: This was done in class and is from Hamilton.]

7. Write a note on Wiener-Kolmogorov representation and also indicate when it is used.

8. Use the Wiener-Kolmogorov representation and compute the infinite sample s-period ahead forecast for AR(1), MA(1) and ARMA(1, 1) process.[Its worked out in Hamilton.]

9. Show that the infinite sample s-period ahead forecast of AR(p) process only depends on p latest values. [This leads to PACF vanishing for AR(p) process after p values.]

10. Write the formula for sample autocorrelation $\hat{\rho}_j$ in terms of given data $x_1, x_2, ..., x_n$. Now suppose $\hat{\rho}_1 = 0.3, \hat{\rho}_2 = 0.1, \hat{\rho}_j = 0, j \geq 3$, and the data is assumed to follow AR(2) model

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \epsilon_t$$

    Using the Yule-Walker equations, estimate ϕ_1, ϕ_2.

Problems

1. Work out the coefficients of the optimal forecast for a $MA(1)$ process based on the latest two samples. Use the formula from previous problem for the derivation.

2. Using Cholesky decomposition, work out the updated linear projection formula. [See Hamilton 4.5.20]

3. Find the optimal forecast $\hat{Y}_{t+2}$ in each of the cases where the known data $\mathbf{X}_t$ is also specified.
   (a) $Y_t = \epsilon_t + 2\epsilon_{t-1}$ where $\mathbf{X}_t = \{..., \epsilon_{t-1}, \epsilon_t\}$ for all t.
   (b) $Y_t = \epsilon_t + 2\epsilon_{t-1}$ where $\mathbf{X}_t = \{..., \epsilon_{-1}, \epsilon_0\}$ for all t.
   (c) $Y_t = \epsilon_t + 2\epsilon_{t-1}$ where $\mathbf{X}_t = \{\epsilon_4\}$ for all t.
   (d) $Y_t = 0.5Y_{t-1} + 2\epsilon_t$ where $\mathbf{X}_t = \{Y_t, Y_{t-1}, ...\}$ for all t.
   (e) $Y_t = 0.5Y_{t-1} + \epsilon_t + 2\epsilon_{t-1}$ where $\mathbf{X}_t = \{..., \epsilon_{t-1}, \epsilon_t\}$ for all t.

4. Consider a time series given by

$$Y_t = 1 + \epsilon_t + 0.5\epsilon_{t-2}$$

   where $\{\epsilon_t\}$ is white noise with unit variance. Work out the optimal linear forecast $\hat{Y}_{t+2|t,t-1,t-2}$ based on three latest samples.