Time Series Models

Jerry Peng

1 Autocorrelation Functions (ACF)

This function measures the correlation between observations in a time series and observations at previous time steps, called lags. For instance, the correlation between the observations at time t and t-1 (one lag), t-2 (two lags), and so on. The ACF is a way to identify repeating patterns or serial correlation in a time series. When plotted, it provides a graph that shows these correlations for different lags. If the ACF shows a gradual decline, it indicates a high level of autocorrelation. The autocorrelatio function at lag k is defined as

$$ACF(k) = \frac{Cov(X_t, X_{t-k})}{Var(X_t)} = \frac{\sum_{t=1}^{N-k}(X_t - \bar{X})(X_{t-k} - \bar{X})}{\sum_{t=1}^{N}(X_t - \bar{X})^2}$$

where

- X_t is the value of the series at time t
- $\bar{X}$ is the mean of the series
- N is the total number of observations

2 Autoregressive Model

Definition An autoregressive model (AR) is a linear model where the current value of the series is regressed on its previous values. The idea is that past values have an influence on current values. This is particularly useful in time series data where there's a natural ordering of data points in time.

Mathematical Formulation An AR model of order p is denoted as $AR(p)$. The expression is as the following:

$$X_t = c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + \epsilon_t$$

where

- X_t is the current value of the series
- c is the constant (intercept term)
- $\phi_1, \phi_2, \ldots, \phi_p$ are the parameters of the model
- $X_{t-1}, X_{t-2}, \ldots, X_{t-p}$ are the past values of the series
- ϵ_t is white noise error term

AR models are widely used in economics, finance, physics, and environmental studies for forecasting future values based on past behavior. AR models assume a **linear relationship** and are not suitable for time series data that exhibit non-linear patterns. They also require the series to be **stationary**

Autocorrelation Functions (ACF) For time series $\{x_t, t \in T\}$, and $t, s \in T$, $\gamma_{t,s}$ is the autocoveriance function for sequence $\{x_t\}$

$$\gamma_{t,s} = E(x_t - \mu_t)(x_s - \mu_s)$$

Hence, the autocorrelation function for autoregressive model is

$$\rho_k = \frac{\gamma_{t,s}}{\sqrt{var(x_t)var(x_{t-k})}}$$

Which is similar to Pearson's correlation function

Partial Autocorrelation Functions (PACF) For stationary series $\{x_t\}$, PACF on lag k is to measure the impact of x_{t-k} to x_t in the middle of given time series $k-1$ on random variables $x_{t-1}, x_{t-2}, \ldots, x_{t-k+1}$

$$\rho_{x_t, x_{t-k}|x_{t-1},\ldots,x_{t-k+1}} = \frac{E[(x_t - \hat{E}x_t)(x_{t-k} - \hat{E}x_{t-k})]}{E[(x_{t-k} - \hat{E}x_{t-k})^2]}$$

In order to calculate PACF function, we need to use past k period sequence values $x_{t-1}, x_{t-2}, \ldots, x_{t-k}$ to make a autoregressive fitting, which is

$$x_t = \phi_{k1}x_{t-1} + \phi_{k2}x_{t-2} + \cdots + \phi_{kk}x_{t-1k} + \epsilon_k$$

The expectations are

$$\hat{E}x_t = \sum_{i=1}^{k-1} \phi_{ki}x_{t-i} + \phi_{kk}\hat{E}(x_{t-k})$$

$$x_t - \hat{E}x_t = \phi_{kk}(x_{t-k} - \hat{E}(x_{t-k}) + \epsilon_t$$

Then we multiple both side with $(x_{t-k} - \hat{E}(x_{t-k}))$, then we get

$$E[(x_t - \hat{E}x_t)(x_{t-k} - \hat{E}(x_{t-k})] = \phi_{kk}E[(x_{t-k} - \hat{E}(x_{t-k}))^2]$$

$$\phi_{kk} = \frac{E[(x_t - \hat{E}x_t)(x_{t-k} - \hat{E}(x_{t-k})]}{E[(x_{t-k} - \hat{E}(x_{t-k}))^2]}$$

Hence, the PACF at lag k is the coefficient of X_{t-k} in the $AR(k)$ model

Make Predictions Assume the original point is h, and F_h means the information at point h, $\hat{x}_h(l)$ is the prediction of a sequence $\{x_t\}$ starting from x_h and take l forward steps.

$$x_{h+l} = \phi_0 + \sum_{i=1}^{p} \phi_i x_{h+l-i} + \epsilon_{h+l}$$

$$\hat{x}_h(l) = E(x_{h+l}|F_h) = \phi_0 + \sum_{i=1}^{l-1} \phi_i \hat{x}_h(i) + \sum_{i=l}^{p} \phi_i x_{h+l-i}$$

3 Moving Average Model

Definition The Moving Average model is based on the assumption that the future value of a series is a function of past errors (or shocks) rather than past values. The errors/shocks represent the differences between the observed values and the predicted values (from a simple mean or another model).

Mathematical Formulation The MA model is usually represented as $MA(q)$, where q is the order of the moving average. The mathematical formula is:

$$x_t = \mu + \epsilon_t + \sum_{i=1}^{q} \theta_i \epsilon_{t-i}$$

where

- x_t: Current value of the series at time t.
- μ: Mean of the series.
- ϵ_t: Error term (or shock) at time t
- θ_i: Parameters of the model which need to be estimated.
- q: Order of the moving average model.

Make predictions The prediction for future value x_{t+1} is

$$\hat{x}_{t+1} = \mu + \epsilon_t + \sum_{i=1}^{q} \theta_i \epsilon_{t+1-i}$$

4 ARMA Model

Definition The ARMA (Autoregressive Moving Average) model combines both Autoregressive (AR) and Moving Average (MA) models to analyze and forecast time series data. This model is particularly useful for understanding and predicting time series data that shows both autoregression and moving average characteristics.

Mathematical Representation An ARMA model is generally represented as $ARMA(p, q)$, where p is the order of the AR part, and q is the order of the MA part. The model is expressed as:

$$x_t = \alpha + \sum_{i=1}^{p} \phi_i x_{t-i} + \epsilon_t + \sum_{j=1}^{q} \theta_j \epsilon_{t-j}$$

where

- x_t: Current value of the series at time t.
- α: Constant term.
- ϕ_i: Coefficients of the AR part.
- ϵ_t: Error term at time t.
- θ_j: Coefficients of the MA part.
- p: Order of the AR part.
- q: Order of the MA part.

5 ARIMA Model

Definition ARIMA (Autoregressive Integrated Moving Average) is an extension of the ARMA (Autoregressive Moving Average) model. It is designed to model a wider range of time series data, including non-stationary series. ARIMA models are particularly useful for understanding and forecasting time series data.

For more detailed information, please refer to this link: https://zhuanlan.zhihu.com/p/634120397

References