Time Series Analysis

Model Diagnostics

Hossein Moradi Rekabdarkolaee

South Dakota State University

email: *hossein.moradirekabdarkolaee@sdstate.edu*.

Office: CAME Building, *Room: 240*.

Introduction

- So far, we have learned
  - Different components of time series
    - Trend,
    - Seasonality,
    - Error.
  - Different type of time series
    - Stationary (weak and strong)
    - Non-stationary.
  - Different models for time series
    - Autoregressive,
    - Moving Average,
    - ARMA,
    - ARIMA.

Introduction

- Different way to find the order of time series
  - ACF,
  - PACF,
  - EACF,
  - Information Criteria-based method (AIC, BIC, AICc, etc.)
- Different approaches to estimate the parameters of the model we choose to fit the time series.
  - Method of Moment,
  - Least Squares,
  - Maximum Likelihood.

Introduction

- In this Section, we are going to talk about
  - Model diagnostics or model criticism,
  - What does it mean
  - How to do it.
  - Do some examples.

Diagnostics

- Model diagnostics, or model criticism, is concerned with testing the goodness of fit of a model and, if the fit is poor, suggesting appropriate modifications.
- The main objective is to check the adequacy of the selected model.
- After fitting an ARIMA model to data, to evaluate the fit, diagnostics checking of the residuals need to be performed.

Diagnostics

- If fact all of the relevant information from the data should be extracted by the model.
- For instance, the part of the data unexplained by the model (residuals) should be small and no systematic patterns should be left in the residuals.
- residuals = actual - predicted
- Standardization allows us to see residuals of unusual size much more easily.

Residual Analysis

- Diagnostic checking essentially involves the statistical properties of the residuals of the fitted model (normality assumptions, weak white noise assumptions, etc.)
- The standardized residuals can be calculated as follows

$$e_t = \frac{Y_t - \hat{Y}_t}{s}, \qquad \text{where } s^2 = \frac{1}{n-1}\sum_{t=1}^{n}(Y_t - \hat{Y}_t)^2$$

- Then, e_t should follow standard Normal White Noise.
- So residuals should be plotted in order to check for
  - substantial residual autocorrelation (Sample ACF, PACF),
  - normality (Q-Q-plots, histogram, normality test),
  - constant variance.

Residual Analysis

- For instance, consider in particular an AR(2) model with a constant term:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \theta_0 + \epsilon_t,$$

then, estimated ϕ_1, ϕ_2 and θ_0, the residuals are defined as

$$e_t = Y_t - \hat{\phi}_1 Y_{t-1} - \hat{\phi}_2 Y_{t-2} - \hat{\theta}_0.$$

- For general invertible ARMA model, when $\theta_0 = 0$, we have

$$Y_t = \pi_1 Y_{t-1} + \pi_2 Y_{t-2} + \pi_3 Y_{t-3} + \ldots + \epsilon_t,$$

- Thus, the estimated residuals are defined as

$$e_t = Y_t - \hat{\pi}_1 Y_{t-1} - \hat{\pi}_2 Y_{t-2} - \hat{\pi}_3 Y_{t-3} - \ldots$$

Diagnostics

- To check against changes in mean and/or variance level, we need to plot the e_t.
- Let us look at some examples.
- Things that we want to check:
  - We want to make sure that the residuals do not have significant ACF and PACF.
  - We want to check if the residuals are normally distributed.
  - Is the transformation necessary?

Autocorrelation of the Residuals

- To check the independence of the noise terms in the model, we examine the sample ACF of the residuals
- The sample ACF of the residuals are denoted with $\hat{r}_k$
- We know that for true white noise and large n, the sample ACF are approximately uncorrelated and normally distributed with zero means and variance $\frac{1}{n}$.
- Generally, the residuals are approximately normally distributed with zero means; however, for small lags k and j, the variance of $\hat{r}_k$ can be substantially less than $\frac{1}{n}$,
- and the estimates $\hat{r}_k$ and $\hat{r}_j$ can be highly correlated.

Autocorrelation of the Residuals - AR(1)

- For instance for AR(1) model it can be shown that

$$Var(\hat{r}_1) \approx \frac{\phi^2}{n}$$
$$Var(\hat{r}_k) \approx \frac{1-(1-\phi^2)\phi^{2k-2}}{n} \quad \text{for} \;\; k > 1$$
$$corr(\hat{r}_1, \hat{r}_k) \approx -sign(\phi)\frac{(1-\phi^2)\phi^{k-2}}{1-(1-\phi^2)\phi^{2k-2}} \quad \text{for} \;\; k > 1$$

where

$$sign(\phi) = \begin{cases} 1 & \text{if } \phi > 0 \\ 0 & \text{if } \phi = 0 \\ -1 & \text{if } \phi < 0 \end{cases}$$

- Simulation results are showing that $Var(\hat{r}_1) \approx \frac{1}{n}$ is a reasonable approximation for $k \geq 2$ over a wide range of ϕ-values.

Autocorrelation of the Residuals - AR(2)

- For instance for AR(2) model it can be shown that

$$Var(\hat{r}_1) \approx \frac{\phi_2^2}{n}, \qquad Var(\hat{r}_2) \approx \frac{\phi_2^2 + \phi_1^2(1+\phi_2)^2}{n}$$

- If the AR(2) parameters are not too close to the stationarity boundary, then

$$Var(\hat{r}_k) \approx \frac{1}{n} \qquad \text{for } \; k \geq 3$$

- It can be shown that results analogous to those for AR models hold for MA models.
- In particular, replacing ϕ by θ gives the results for the MA(1) case.
- Similarly, results for the MA(2) case can be stated by replacing ϕ_1 and ϕ_2 by θ_1 and θ_2, respectively.

The Ljung-Box Test

- In addition to looking at residual correlations at individual lags, it is useful to have a test that of their magnitudes as a group.
- A common method for testing against residual auto-correlation is Ljung-Box test:

$$\begin{cases} H_0: & \text{The error terms are uncorrelated,} \\ H_a: & \text{the error terms are correlated.} \end{cases}$$

- Ljung-Box statistic:

$$Q = n(n+2)\sum_{k=1}^{K} \frac{\hat{r}_k^2}{n-k}$$

The Ljung-Box Test

- If $e_t \sim WN$ then $Q \sim \chi_{(K-p-q)}$, as $n \to \infty$
- Here the maximum lag K is selected somewhat arbitrarily but large enough that the weights are negligible for $j > K$.
- The test would reject the ARMA(p, q) model if the observed value of Q exceeded an appropriate critical value in a chi-square distribution with $K - p - q$ degrees of freedom.

Overfitting and Parameter Redundancy

- After specifying and fitting what we believe to be an adequate model, we fit a slightly more general model; that is, a model “close by” that contains the original model as a special case.

- For example, if an AR(2) model seems appropriate, we might overfit with an AR(3) model

- then, the original AR(2) model would be confirmed if:

  1. the estimate of the additional parameter, ϕ_3, is not significantly different from zero, and
  2. the estimates for the parameters in common, ϕ_1 and ϕ_2, do not change significantly from their original estimates
  3. Notice that, using higher order (p,q) leads to suboptimal estimation (overfitting).

Overfitting and Parameter Redundancy

- The implications for fitting and overfitting models are as follows:

  1. Specify the original model carefully. If a simple model seems at all promising, check it out before trying a more complicated model.
  2. When overfitting, do not increase the orders of both the AR and MA parts of the model simultaneously.
  3. Extend the model in directions suggested by the analysis of the residuals.
  4. For example, if after fitting an MA(1) model, substantial correlation remains at lag 2 in the residuals, try an MA(2), not an ARMA$(1, 1)$.

Summary

- The parsimony principle, Keep it simple!
- If two models performs equally well in terms of residual variance or likelihood, pick the one with fewer parameters
- Including too many parameters can lead to overfitting
- This applies to p and q, but also to d.
- Too high d leads overdifferencing