Time Series Regression

Regression based methods are important in time series analysis. Assume that the response x_t $t = 1, \ldots, n$ is connected to a collection of inputs $z_t = (z_{t1}, \ldots, z_{tq})', \; t = 1, \ldots, n$ by the linear regression model

$$x_t = \beta_0 + \beta_1 z_{t1} + \cdots + \beta_q z_{tq} + w_t \tag{1}$$

where $w_t \sim wn(0, \sigma_w^2)$ and $\beta = (\beta_1, \ldots, \beta_q)'$ is a vector of parameters.
For the moment the inputs z_t , $t = 1, \ldots, n$ are assumed to be fixed. The equation (1) can be written on vector form as

$$x_t = \beta z' + w_t.$$

The parameters $\beta_1, \ldots, \beta_q$ is found by minimizing the sum of squares

$$Q = \Sigma_{t=1}^{n}(x_t - \beta_0 + \beta_1 z_{t1} - \cdots - \beta_q z_{tq})^2 = \Sigma_{t=1}^{n}(x_t - \beta z')^2.$$

The solution is, when the matrix $\Sigma_{t=1}^{n} z_t z_t'$ is non-singular,

$$\hat{\beta} = (\Sigma_{t=1}^{n} z_t z_t')^{-1} \Sigma_{t=1}^{n} z_t x_t.$$

and the minimal value of Q is the *sum of squared errors*

$$SSE = \Sigma_{t=1}^{n}(x_t - \hat{\beta} z')^2.$$

When $w_t \sim wn(0, \sigma_w^2)$,

- $E(\hat{\beta}) = \beta$
- If $s_w^2 = \frac{SSE}{n-(q+1)}$, $E(s_w^2) = \sigma_w^2$.
- $cov(\hat{\beta}) = \sigma_w^2 (\Sigma_{t=1}^n z_t z_t')^{-1}$

When, in addition, $w_t \sim iidN(0, \sigma_w^2)$,

- $\hat{\beta}$ multinormal
- t$=\frac{(\hat{\beta}_i - \beta_i)}{s_w \sqrt{c_{ii}}} \sim t_{n-(q+1)}$

Reduced or sub model

$$x_t = \beta_0 + \beta_1 z_{t1} + \cdots + \beta_q z_{tr} + w_t \tag{2}$$

i.e. $\beta_{r+1} = \cdots = \beta_q = 0$. Let SSE_r be the sum of squared errors in the reduced model.

Test for reduced model against full model, i.e. $H_0 : \beta_{r+1} = \cdots = \beta_q = 0$:

$$F = \frac{(SSE_r - SSE)/(q-r)}{SSE/(n-q-1)} = \frac{MSR}{MSE}$$

is under H_0 Fisher distributed with $q - r$ and $n - q - 1$ degrees of freedom.

ANOVA-tables useful

Special case (r=0):

Model 0 $\beta_1 = \cdots = \beta_q = 0$

$$x_t = \beta_0 + w_t \tag{3}$$

Coefficient of determination

$$R^2 = \frac{SSE_0 - SSE}{SSE_0}$$

where

$$SSE_0 = \Sigma_{t=1}^{n}(x_t - \bar{x})^2.$$

Model selection

Stepwise procedures

Information criteria, k=r+1 regressors $\hat{\sigma}_k^2 = \frac{SSE(k)}{n}$

- Akaike's Information Criterion (AIC):
  $AIC = \log \hat{\sigma}_k^2 + \frac{n+2k}{n} \log \hat{\sigma}_k^2 + \frac{n+2k}{n}$
- AIC, Bias Corrected (AICc):
  $AICc = \log \hat{\sigma}_k^2 + \frac{n+k}{n-k-2}$
- Bayesian Information Criterion (BIC):
  $BIC = \log \hat{\sigma}_k^2 + \frac{k+\log n}{n}$

Exploratory data analysis

Stationarity is a regularity condition that allows averaging over time to estimate expectation and autocovariance/autocorrelation. Often the assumption is violated in the raw data, and some preliminary transformations to obtain stationarity is necessary.

- Detrending: $x_t = \mu_t + y_t$, y_t stationary, e.g. $\mu_t = \beta_0 + \beta_1 t$.
- Differencing: $\bigtriangledown x_t = x_t - x_{t-1}$, appropriate if a random walk is a candidate model.

Notation:

- Backshift operator $Bx_t = x_{t-1}$, $B^k x_t = x_{t-k}$
- Difference order d: $\bigtriangledown^d = (1-B)^d$

Transformation can be useful, $y_t = \log x_t$ or more generally

$$y_t = \begin{cases} \frac{(x_t^\lambda - 1)}{\lambda} & \lambda \neq 0 \\ \log x_t & \lambda = 0 \end{cases}$$

Scatterplot matrices visualize relations between different lags of a series and different lags of several series.

Regression analysis can also be used to assess periodic behaviour in series.

Smoothing time series

Smoothing can be useful to reveal patterns in time series as trends and periodic behavior. Several available procedures

- Moving average
- Kernel smoothing, use a kernel to average over the observations
- Lowess, based on nearest neigbourhood regression
- Splines, piecewise polynomial regressions

It is also possible to smooth the scatterplot of the observations from one time series against observations from another.

Special care needed at start and end of series.