Based on the data from homework 1, build a time series regression model for the variable CPI and produce forecasts for time points 81-90. The considerations that you should take into account are (not necessarily in that order)

- Do you need to transform the series CPI to have stationarity/normallity/constant variance etc. You can consider box-cox type transformation and/or differencing the series.

- What regressors to put in the model? You can put either of the regressors available to you, EMP, GNPPERCA, INTR or additionally you can transform the regressors to create new one. For example, you may want to include in the regression the lagged value of EMP (\(EMP_{t-1}\)) rather than \(EMP_t\), i.e.;

\[
CPI_t = \beta_0 + \beta_1 EMP_{t-1} + \ldots + \epsilon_t, \quad t = 2, 3, \ldots, 80.
\]

- As you have already realized, an ordinary multiple regression model will not be an efficient model. You need to model the errors, \(\epsilon_t\) as time series. In homework 3, you have looked at the properties of the residuals from one particular multiple regression model

\[
CPI_t = \beta_0 + \beta_1 EMP_t + \beta_2 GNPPERCA_t + \beta_3 INTR_t + \epsilon_t.
\]

Do the same analysis (looking at sample autocorrelation and partial autocorrelation) for the residuals from the multiple regression model you are fitting. Identify possible ARMA\((p, q)\) models \((p \leq 3, q \leq 3)\) for the residuals or the differenced residuals \((\epsilon_t - \epsilon_{t-1})\) in the case the residuals look nonstationary. Use the AIC (Akaieke’s Information Criterion) for deciding which ARMA model to choose.

- Verify that the model you have fitted is stationary.

- Look at the Box-Ljung white-noise test value of the residuals from the ARMA model to check that the assumptions about the errors in the ARMA model are met.

- Produce you forecast and the forecast variance of the transformed model.