

Solutions to Homework Assignment #3

1. Weekly mortgage rate series.

One can identify a model for the series by using either (a) ACF and PACF of the data and its first differenced series or (b) the “ar” command. The ACF of the observed mortgage rates indicates the need to take the first difference. The PACF or ACF of the differenced data has only a significant value at lag 1. Thus, an ARIMA(1,1,0) or ARIMA(0,1,1) is specified. For the “ar” command, an AR(2) is identified. Since an ARIMA(1,1,0) model can be regarded as an AR(2) model with a unit root, I use the ARIMA(1,1,0) model for simplicity. [You may use other models so long as (1) the estimates of the parameters are significant and (2) model checking confirms that the fitted model is adequate. In fact, an ARIMA(0,1,1) model has a smaller AIC value for the mortgage rate series.]

The fitted model is

$$(1 - 0.21B)(1 - B)r_t = a_t, \quad \sigma_a^2 = 0.0063,$$

where r_t denotes the weekly mortgage rate. The standard error of the AR(1) coefficient is 0.07 so that the estimate is significant at the 5% level.

For model checking, we have $Q(12) = 4.42$ with p-value 0.97, indicating no residual serial correlations. The model is adequate. [Note that you can also check the mean of the residuals. The mean is not different from zero at the 5% level.]

The 1-step to 4-step ahead forecasts at April 3, 2008 are 5.886, 5.888, 5.888, and 5.888, respectively. The associated standard errors are 0.080, 0.125, 0.160, and 0.189, respectively.

2. Weekly mortgage rate and interest rate series. Let y_t and x_t be the weekly mortgage rate and interest rate, respectively. The simple linear regression is

$$y_t = 5.035 + 0.277x_t + e_t,$$

where the estimate 0.277 is highly significant. However, the Ljung-Box statistics of the residuals show $Q(12) = 1221.84$ with p-value close to zero. Thus, the model is inadequate. The residuals of the prior model have high serial correlations so that we take the first difference of the two series. Denote the differenced series by dy_t and dx_t , respectively.

The linear regression model is

$$dy_t = 0.0009 + 0.316dx_t + e_t, \quad \sigma_e = 0.075.$$

The coefficient 0.316 is highly significant. The Ljung-Box statistics of the residuals fail to indicate any serial correlations in the residuals with $Q(12) = 7.88$ with p-value 0.79. Thus, the above model is adequate for the two series.

For the model, since the t -ratio of the estimate 0.316 is 6.36, which is highly significant. Thus, we reject the null hypothesis of $\beta = 0$, where β is the regression coefficient. Consequently, the mortgage rate depends on the interest rate.

Remark: The ACF of the residuals shows two spikes at lags 1 and 2, even though they are not significant at the 5% level. If one considers an MA(2) model for the residuals, then the fitted model is

$$dy_t = 0.001 + 0.386dx_t + (1 - 0.162B + 0.141B^2)a_t, \quad \sigma_a^2 = 0.0054.$$

The two MA coefficients have t -ratio very close to -2 and 2, respectively. This model is also adequate. The difference between this model and the one without MA part is small. Perhaps, it is due to the fact that the sample size is 222, which is not large in time series analysis. Either model confirms that mortgage rate depends on the interest rate.

3. Decile 1 return series. The regression model is

$$r_t = 0.0065 + 0.095Jan_t + e_t,$$

where r_t is the monthly regression of Decile 1 portfolio and Jan_t is the January dummy variable. The t -ratio of 0.095 coefficient is 7.43, indicating highly significant of the estimate. However, the residual ACF shows a significant serial correlation at lag 1, and we entertain a regression model with MA(1) errors. The fitted model is

$$r_t = 0.007 + 0.088Jan_t + a_t + 0.182a_{t-1}, \quad \sigma_a^2 = 0.0029.$$

The standard error of the regression coefficient is 0.012 whereas that of the MA(1) coefficient is 0.065. Thus, the estimates are statistically significant. In other words, the January effect is statistically significant at the 5% level. The Ljung-Box statistics of the residuals shows $Q(24) = 15.49$ with p-value 0.91. Thus the model is adequate.

The 1- to 12-step ahead forecasts are 0.094 0.007 0.0072 0.0072 0.0072 0.0072 0.0072 0.0072 0.0072 0.0072 0.0072 0.0072.

4. The fitted model is

$$(1 - 0.999B^{12})(r_t - 0.0148) = (1 - 0.193B)(1 - 0.976B^{12})a_t, \quad \sigma_a^2 = 0.0030.$$

The Ljung-Box statistics of the residuals show $Q(24) = 15.13$ with p-value 0.92. Thus, the residuals have no serial correlations. The model is also adequate.

The 1-step to 12-step ahead forecasts are 0.078 0.017 0.006 0.005 0.017 0.007 0.005 -0.003 0.006 -0.001 0.005 0.028.

These forecasts are similar to that of Problem 3, especially in view of the standard errors of the forecasts.

5. FedFx earnings data. The ACF of the differenced data (regular and seasonal differences) shows that the airlines model is reasonable for the log series. The fitted model is

$$(1 - B)(1 - B^4)r_t = (1 - 0.53B)(1 - 0.82B^4)a_t, \quad \sigma_a^2 = 0.0078.$$

The standard errors of the two estimates are 0.103 and 0.223, respectively. The two coefficients are significant. Residuals of the model show $Q(12) = 4.72$ with p-value 0.97. Thus, the model is adequate.

The 1-step to 4-step ahead forecasts are 1.44, 2.17, 1.77, 2.16, respectively. [This is obtained by taking anti-log transformation, then subtract 0.5.]

Note that if one use $\exp[r_n(\ell) + \frac{1}{2}\sigma_a^2] - 0.5$, then the forecasts are 1.45, 2.18, 1.78, 2.17, respectively. These numbers are close to what we had above.