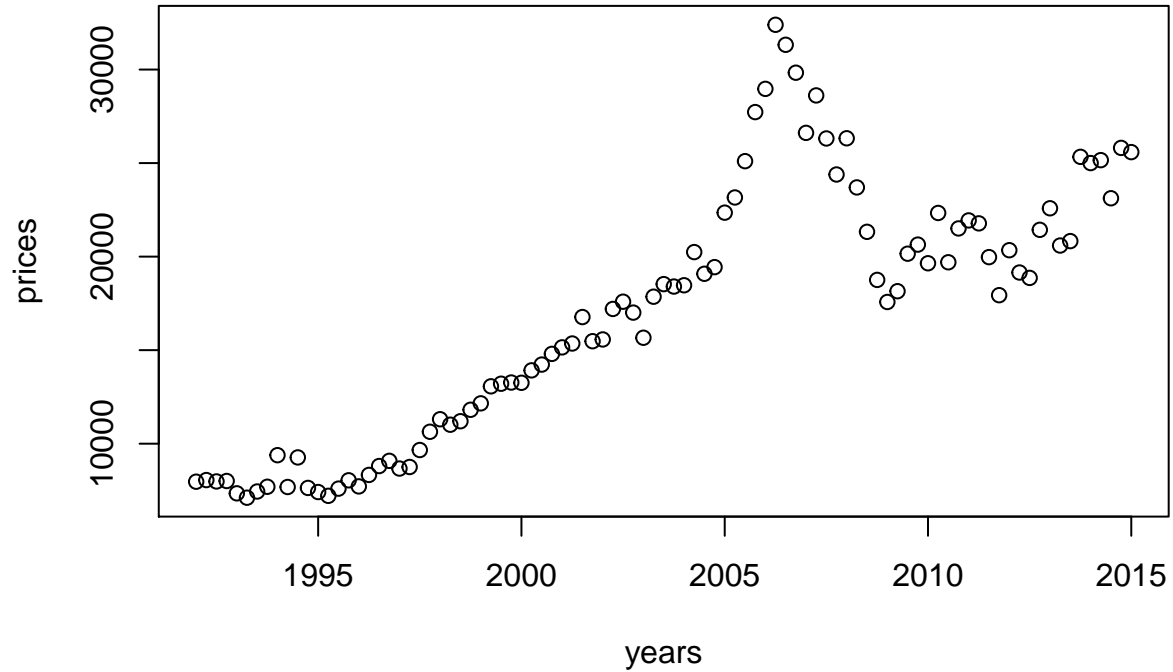# Assignment 1

Arturo Arranz Mateo in colaboration with Dominik Dominković
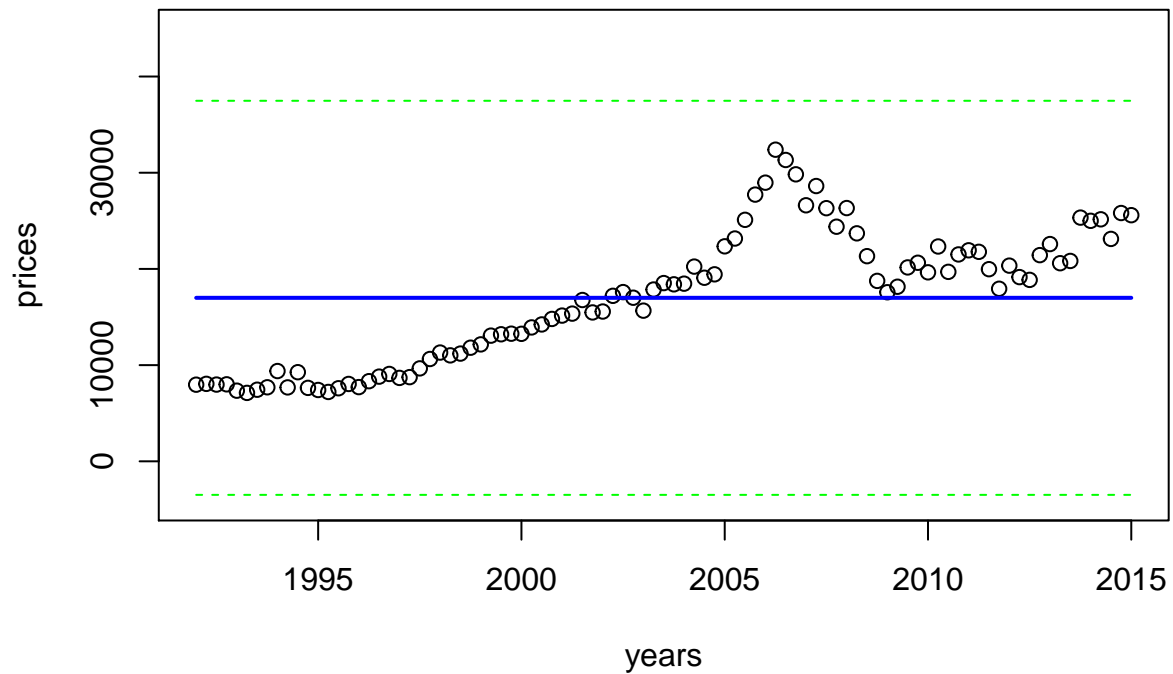
March 2, 2016

Question 1

*Plot the apartment price as a function of time.*

Question 2

Does the global mean value and standard deviation of the prices give a reasonable representation of the data? (The answer should be elaborated.)

As we can see in the graphic, the mean and the standard deviation by itself, do not give us any information about the trend. It would be useful in the case, we were dealing with white noise, where the data is normally distributed. Here clearly, we have a trend. In the Q-Q plot we can see clearly how the data is not normally distributed.

Normal Q−Q Plot
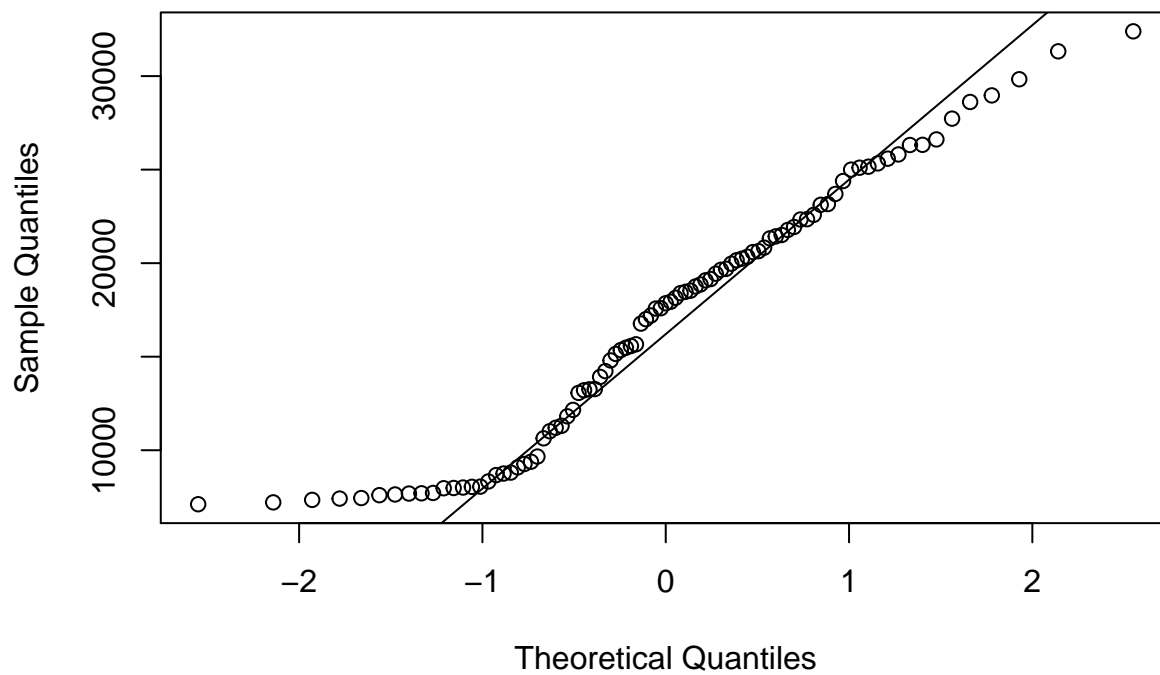
Question 3

Let us consider a model for all the data. Formulate a GLM model in form of a simple linear regression model for all the data. Estimate the model parameters. Would this model be useful for making predictions of future apartment prices?
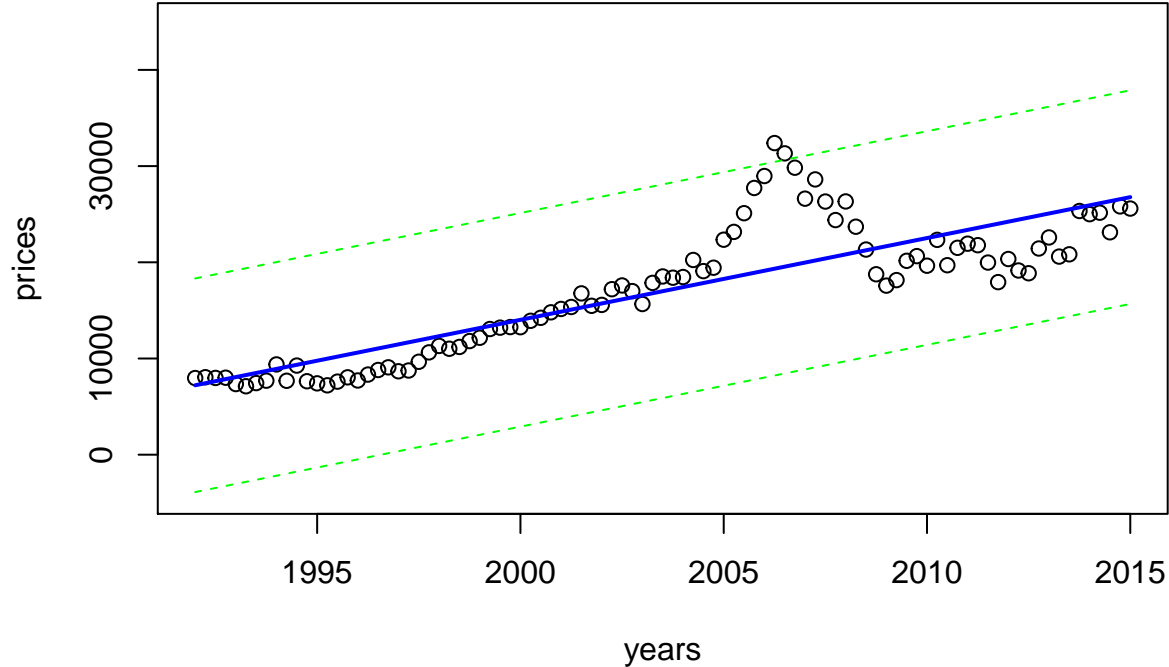
As we can see in the plot, we can make predictions based on the global trend. However, around 2006, the data is more than 3 times bigger than the standard deviation respect our prediction. This is not a very robust model for rapid changes on the trend. In additio, the spike from 2006, have a big impact on the model, which is carried during the next predictions.

Question 4

Now we will consider methods which considers data more locally. Use simple exponential smoothing to predict the apartment prices for the remaining part of 2015 and 2016. (Just choose a value for the smoothing constant, eg. $\lambda = 0.9$). Comment on the results.

The simple exponential smoothing can be written as:

$$\hat{Y}_{N+l|N} = c \sum_{j=0}^{N-1} \lambda^j Y_{N-j} = c[Y_N + \lambda Y_{N-1} + ... + \lambda^{N-1} Y_1]$$

Where c is the normalizing constant which is choosen such that the sum of the weights is 1. If N is large we obtain $c \approx (1 - \lambda)$. With this assumption we can construct the ons-step prediction:

$$\hat{Y}_{N+2|N+1} = (1 - \lambda) \sum_{j=0} \lambda^j Y_{N+1-j} = (1 - \lambda)Y_{N+1} + \lambda \hat{Y}_{N+1|N}$$
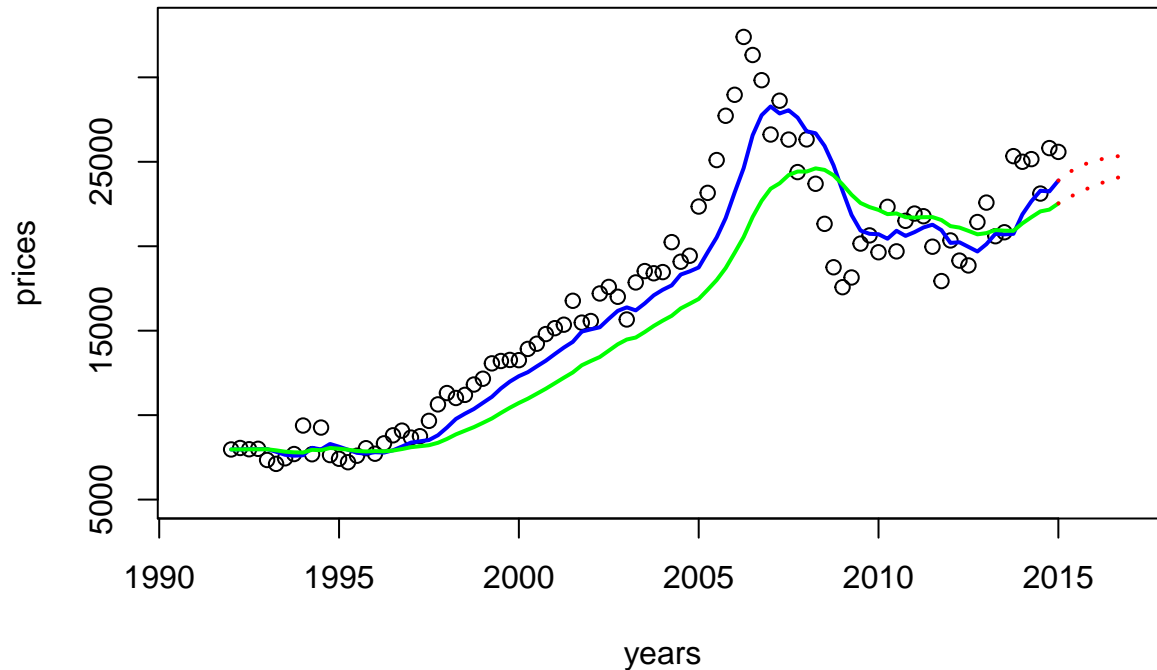
The predictions for lambda = 0.9:

2015	2015.25	2015.5	2015.75	2016	2016.25	2016.5	2016.75	2017
22533.25	22838.52	23113.27	23360.54	23583.09	23783.38	23963.64	24125.88	24271.89

In the graph is represented a simple model with lambda **0.9 in green** and **0.75 in blue.**

We can notice several things from the graph. First, we can notice the effect on choosing a lower or higher forgetting factor. A larger lambda **react slower to changes** as we can see from the green line. Also we can see how a lower lambda **is more sensitive** to drastic changes in the trend. We can see how the blue line follow the spike from the year 2006.

Secondly, at the predictions we can see how we can not infer a trend with a constant **locally mean model**. This model is useful for one-step predictions, since for calculating each new predictions, we need the measurement from current state.

Note that for calculating the predictions I assumed that each one-step predictions would be the same as the last measurment.

$$\hat{Y}_{94|93} = (1 - \lambda)\hat{Y}_{93} + \lambda\hat{Y}_{93|92}$$

$$\hat{Y}_{95|93} = (1 - \lambda)\hat{Y}_{93} + \lambda\hat{Y}_{94|93}$$

$$\hat{Y}_{95|93} = (1 - \lambda)\hat{Y}_{93} + \lambda\hat{Y}_{95|94}$$

We can see how slowly the predictions reach the last measurment. When this value is reached, the predictions are constant.

Question 5

Use a local linear trend model to predict the apartment prices for the remaining part of 2015 and 2016. State the uncertainty of the predictions. Comment on the results.

In the local trend model we create a linear model which iteratively update the parameters, θ, and also include a forgetting factor, λ, which give more value to recent data. We can write this case as:

$$\hat{\theta}_N = arg\min_{\theta} S(\theta; N)$$

$$S(\theta; N) = \sum_{j=0}^{N-1} \lambda^j [Y_{N-j} - f^T(-j)\theta]^2$$

Before solving the minimizing problem we have to work on our model. We need to create a global model. We aim to create a linear model:

$$Y_{N+j} = \theta_0 + \theta_1 j + \epsilon_{N+j} \quad \Rightarrow \quad Y_{N+j} = f^T(j)\theta + \epsilon_{N+j}$$

$$\text{where} \quad f(j) = \begin{pmatrix} 1 & j \end{pmatrix}^T$$

For multivariate:

$$Y_N = x_N \theta_N + \epsilon_N$$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{bmatrix} = \begin{bmatrix} f^T(-N+1) \\ f^T(-N+2) \\ \vdots \\ f^T(0) \end{bmatrix} \theta_N + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_N \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} 7968 \\ 8055 \\ \vdots \\ 25586 \end{bmatrix} = \begin{bmatrix} f^T(92) \\ f^T(91) \\ \vdots \\ f^T(0) \end{bmatrix} \theta_N + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_N \end{bmatrix}$$

Now, solving the minimizing problem we get a WLS:

$$\hat{\theta}_N = ((x_N)^T \Sigma^{-1} x_N)^{(} - 1)(x_N)^T \Sigma^{-1} Y$$

where the weight matrix is:

$$\begin{bmatrix} \lambda^{N-1} & 0 & \cdots & 0 \\ 0 & \lambda^{N-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = \begin{bmatrix} 0.9^{92} & 0 & \cdots & 0 \\ 0 & 0.9^{91} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

Our estimated parameters are:

$$\hat{\theta}_N = F^{-1} hN = \begin{bmatrix} 23957 & 124 \end{bmatrix}^T$$

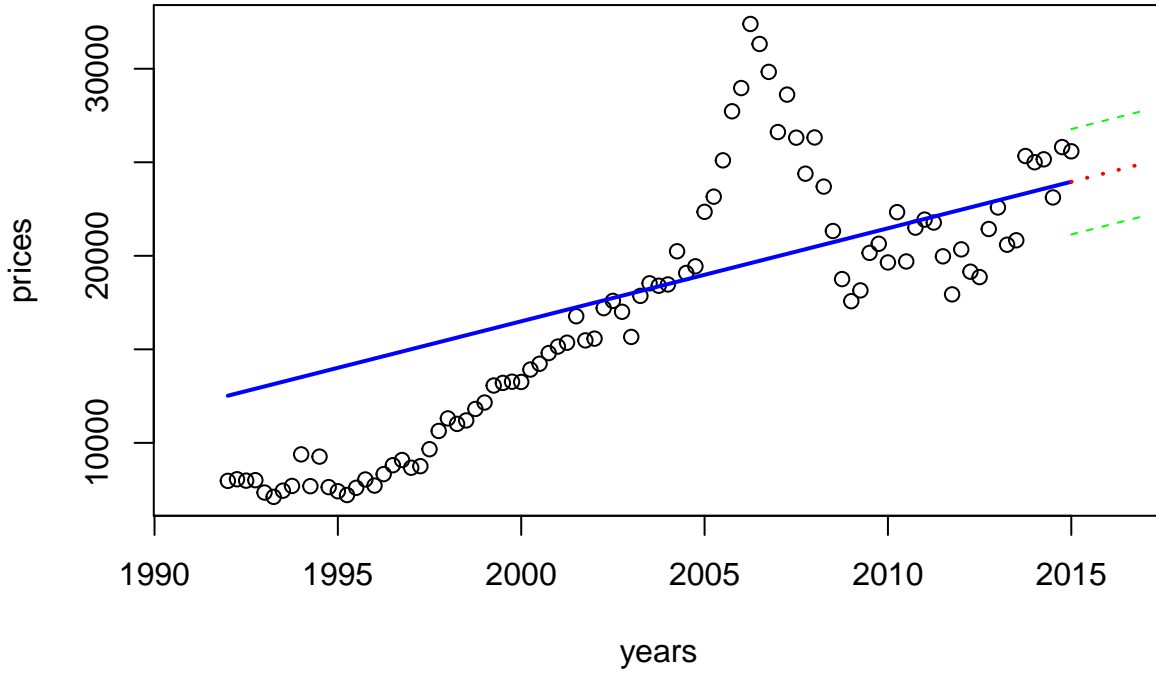$$F^{-1} = ((x_N)^T \Sigma^{-1} x_N)^{-1}$$

$$hN = (x_N)^T \Sigma^{-1} Y$$

Now we can do one step predictions:

$$Y_{N+j} = f(j+1)^T \theta$$

We can transit from j to j+1 with:

$$f(j+1) = Lf(j)$$

$$L = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad f(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

The predictions are:

2015	2015.25	2015.5	2015.75	2016	2016.25	2016.5	2016.75	2017
23957.58	24081.9	24206.21	24330.53	24454.85	24579.17	24703.49	24827.81	24952.13

Now we can estimate sigma and the variance of the error:

$$\hat{\sigma^2} = (Y - x_N \hat{\theta}_6)^T \Sigma_{-1} (Y - x_N \hat{\theta}_6)/(T - p) = 7935751$$

$$Var(\epsilon_{93}) = \hat{\sigma^2}(1 + f^T(l)(F_6)^{-1}f(l)) = 12183640025$$

In this model we can see how the last measurments have more impact on the trend than the model build in the question 1.3. This result in more accurate predictions.

Question 7