

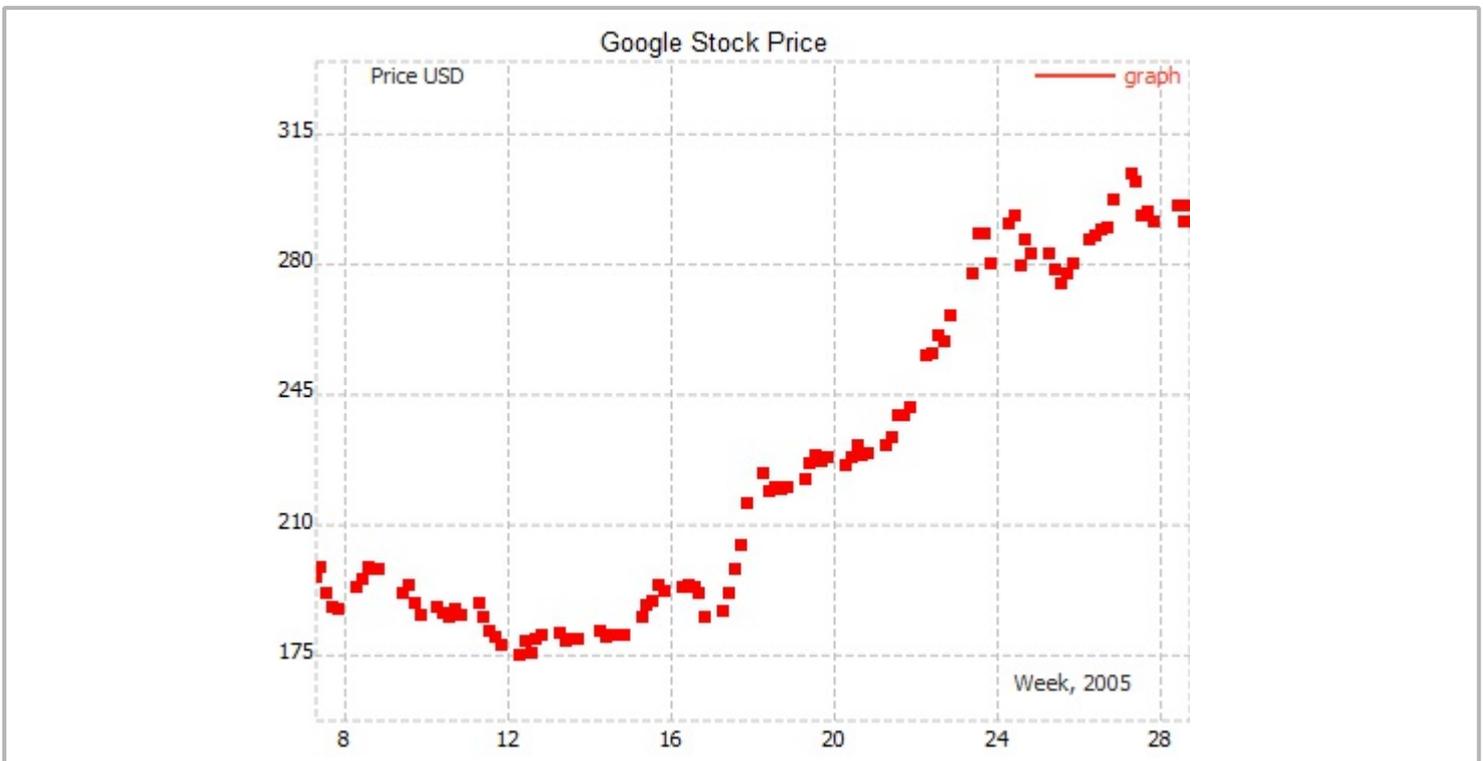
Model Identification in Economic Forecasting

A time series is a sequence of measurements of the same variable made over time spaced in regular intervals, for example, daily, monthly or yearly. As an example, we can have a measure of stock price, with measurements observed daily. An autoregressive model - AR(p) is when a current value from a time series is regressed on previous values from that same time series. An order of autoregressive model p tells how many previous measurements are taken to calculate the current measurement. For example AR(2) model is given below, where n_t is Additive White Gaussian Noise - AWGN of variance σ^2 .

$$x_t = f_0 + f_1 x_{t-1} + f_2 x_{t-2} + n_t$$

As an example for our analysis we take the data set which consists of $n = 105$ values which are the closing stock price of a share of Google stock between 7-2-2005 to 7-7-2005. The value against time is illustrated below.

```
1 time := excel read("googledata.xlsx", "Sheet1", "C2:C106", false)
2 price := excel read("googledata.xlsx", "Sheet1", "B2:B106", false)
3 graph := join mat cols(time, price)
```



Autocorrelation and Partial Autocorrelation

The coefficient of correlation between two values separated by k lags in a time series is called the autocorrelation function (ACF). The ACF is a way to measure the linear relationship between an observation at our current time and observations at previous times. If we assume an AR(k) model, then we may wish to only measure the association between values separated by k . In this case we calculate the correlation of the transformed time series using **partial autocorrelation function (PACF)**.

The PACF is the most useful option and is commonly used for identifying the order of an autoregressive model. The most common is a graphical approach to estimate the lag of an autoregressive model by inspecting the ACF and PACF values against the lag. In MatDeck, there are functions `pacf()` and `correlation()` which can be used for model identification.

```

4  pcftest := pacf(price, 10)
5  acftest := acf(price, 10)
6  lagp := ynodes(x, 1, 26, 10)
7  laga := ynodes(x, 1, 26, 11)

```

Here, we use the acf() function to prepare the autocorrelation according to the procedure explained above.

```

8  acf(dta, n)
9  {
10  len := size(dta)
11  xx := vector create(n + 1, true, 0)
12  ro := vector create(n + 1, false, 0)
13  for(i := 0; i <= n; i = i + 1)
14  {
15  xx[i] = subset(dta, n - i, 0, len - i - 1, 0)
16  ro[i] = correlation(xx[0], xx[i])
17  }
18  return(ro)
19  }

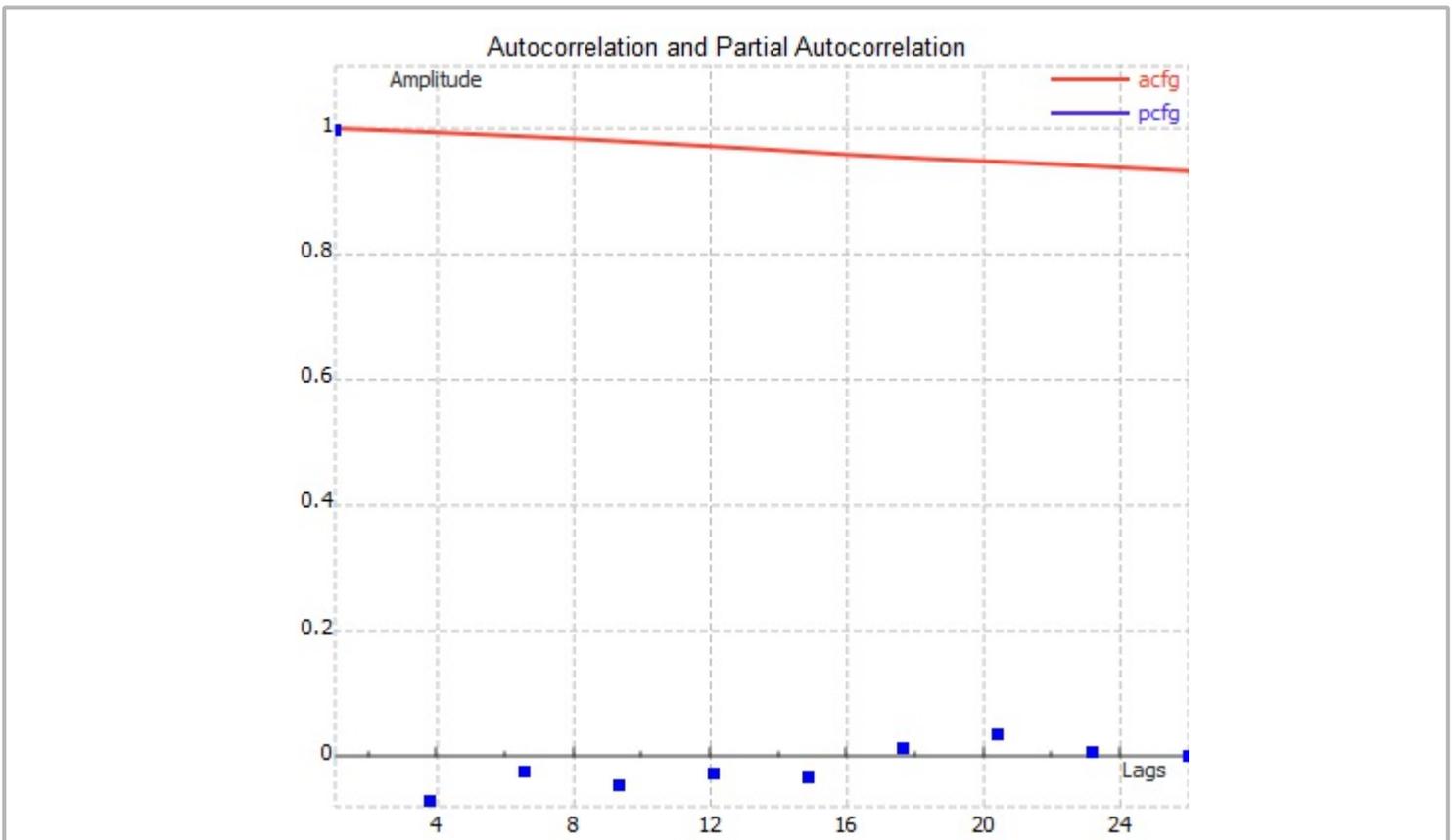
```

We next look at a plot of partial autocorrelations for the data:

```

20  pcfg := join mat cols(lagp, pcftest)
21  acfg := join mat cols(laga, acftest)

```



Here we notice that there is a significant pacf sample at a lag of 1 and much lower samples for the subsequent lags, therefore we can conclude that an AR(1) model would likely be feasible for given data

$$x_t = f_0 + f_1 x_{t-1} + n_t$$