4.4 Stock Market Prediction

In [17]:

```
# read data from txt files
with open('nasdaq00.txt','r') as f:
    contents = f.readlines()
    f.close()
with open('nasdaq01.txt','r') as f:
    contents2 = f.readlines()
    f.close()
```

In [18]:

```
# data management
import numpy as np
for i in range(len(contents)):
    temp = contents[i].strip()
    contents[i] = float(temp)
contents = np.array(contents)
```

In [19]:

```
# seperate today's data and last three days' data
x = contents[3:]
x_prev = np.array([contents[2:len(contents)-1],contents[1:len(contents)-2],contents[
```

In [20]:

```
# using MLE to estimate coefficients
from numpy.linalg import inv
A = np.dot(x_prev,x_prev.T)
b = np.dot(x_prev,x[:,np.newaxis])
a_mle = np.dot(inv(A),b)
print(a_mle)
```

```
[[0.95067337]
 [0.01560133]
 [0.03189569]]
```

In [21]:

```
# data management
for i in range(len(contents2)):
    temp = contents2[i].strip()
    contents2[i] = float(temp)
contents2 = np.array(contents2)
```

In [22]:

```
# seperate today's data and last three days' data
y_test = contents2[3:]
x_test = np.array([contents2[2:len(contents2)-1],contents2[1:len(contents2)-2],conte
```

In [23]:

```
print('MSE for 2000 ',((x - np.dot(a_mle.T,x_prev)) ** 2).mean())
print('MSE for 2001 ',((y_test - np.dot(a_mle.T,x_test)) ** 2).mean())
```

```
MSE for 2000  13902.401076367898
MSE for 2001  2985.09792411156
```