Bayesian Curve Fitting

Although we have included a prior distribution p(w|α), we are so far still making a point estimate of w and so this does not yet amount to a Bayesian treatment. In a fully Bayesian approach, we should consistently apply the sum and product rules of probability, which requires, as we shall see shortly, that we integrate over all values of w. Such marginalizations lie at the heart of Bayesian methods for pattern recognition. Stock Prediction has become an important requirement for the investors and the traders in order to invest their money in the right place and incur minimal losses. There are prediction models which take care of this. Bayesian-curve fitting is one of them. It fetches the history of the values and self-learn the pattern followed by these values. It then can be used to predict a future value by the pattern with which it was trained.

To test the Bayesian curve fitting, we take a set of values on the X-axis as the Time frames around which the price values of the stocks have been plotted on the Y-axis.

X-axis (Unit:Days): plot 1 to N values (here N=10 i.e X-axis: 1 to 10)

Y-axis (Unit:Price($)): plot values of the prices at the time on X-axis of the stocks.

We fetch the data from the CSV files which are obtained from yahoofinance.com website.These files are in the form of Comma-separated-values.

The figure below shows an example of a CSV file of Apple with the stock price history of 24th January to 23rd February. We extract the value of the 'Close' key from each line as the Price on the Y-axis

A1 fx Date

	A	B	C	D	E	F	G	H
1	Date	Open	High	Low	Close	Volume	Adj Close	
2	########	137.38	137.48	136.3	136.53	20704100	136.53	
3	########	136.43	137.12	136.11	137.11	20745300	137.11	
4	########	136.23	136.75	135.98	136.7	24265100	136.7	
5	########	135.1	135.83	135.1	135.72	22084500	135.72	
6	########	135.67	135.9	134.84	135.35	22118000	135.35	
7	########	135.52	136.27	134.62	135.51	35501600	135.51	
8	########	133.47	135.09	133.25	135.02	32815500	135.02	
9	########	133.08	133.82	132.75	133.29	23035400	133.29	
10	########	132.46	132.94	132.05	132.12	20065500	132.12	
11	2/9/2017	131.65	132.45	131.12	132.42	28349900	132.42	
12	2/8/2017	131.35	132.22	131.22	132.04	23004100	131.47	
13	2/7/2017	130.54	132.09	130.45	131.53	38183800	130.9622	
14	2/6/2017	129.13	130.5	128.9	130.29	26845900	129.7275	
15	2/3/2017	128.31	129.19	128.16	129.08	24507300	128.5228	
16	2/2/2017	127.98	129.39	127.78	128.53	33710400	127.9752	
17	2/1/2017	127.03	130.49	127.01	128.75	1.12E+08	128.1942	
18	########	121.15	121.39	120.62	121.35	49201000	120.8261	
19	########	120.93	121.63	120.66	121.63	30377500	121.1049	
20	########	122.14	122.35	121.6	121.95	20562900	121.4236	
21	########	121.67	122.44	121.6	121.94	26337600	121.4136	
22	########	120.42	122.1	120.28	121.88	32377600	121.3539	
23	########	119.55	120.1	119.5	119.97	23211000	119.4521	
24								
25								
26								

The figure shown below are taken from Pattern Recognition and Machine Learning by Christopher M. Bishop. These formulas have been executed in the algorithm for Bayesian Curve Fitting.

$$p(t|x, \mathbf{x}, \mathbf{t}) = \mathcal{N}\left(t|m(x), s^2(x)\right) \tag{1.69}$$

where the mean and variance are given by

$$m(x) = \beta\boldsymbol{\phi}(x)^{\mathrm{T}}\mathbf{S}\sum_{n=1}^{N}\boldsymbol{\phi}(x_n)t_n \tag{1.70}$$

$$s^2(x) = \beta^{-1} + \boldsymbol{\phi}(x)^{\mathrm{T}}\mathbf{S}\boldsymbol{\phi}(x). \tag{1.71}$$

Here the matrix $\mathbf{S}$ is given by

$$\mathbf{S}^{-1} = \alpha\mathbf{I} + \beta\sum_{n=1}^{N}\boldsymbol{\phi}(x_n)\boldsymbol{\phi}(x)^{\mathrm{T}} \tag{1.72}$$

The figure below shows all the values calculated using formulas 1.69-1.72. We have executed the Bayesian Algorithm and the mean ‘m(x)’, variance ‘s^2(x)’ and the Prediction ‘predicted’ through Gaussian Normal distribution.

```
>
> ##to calculate absolute mean error
> abs_mean_error=abs(predicted-actual)
> ##to calculate average relative error
> avg_rel_error=(mean_abs_error)/actual
> variance
          phix
phix 0.1295969
> mean
         [,1]
phix 133.2575
> predicted
[1] 132.8314
> abs_mean_error
[1] 0.7914096
> avg_rel_error
[1] 0.01881279
```

Few of the values have been taken as constants:

Alpha=5*10^-3

Beta=11.1

M=4(which the order of the x vector)

N=10 i.e the number of values given to the Bayesian curve to train itself.

x=11 I,e the 11th day for which we want to predict the price. Thus we feed the history (of 10 days price v/s time values from 24th Jan-23rd Feb 2016) of 10 different stocks.

The figure shows the predicted, absolute mean error and average relative error of the stock of companies mentioned as follows:

AAPL	MSFT
GOOG	M
AMZN	TSLA
BABA	NFLX
MSFT	TWTR

AAPL.csv
Predicted price at time= 132.8339
Absolute Mean Error for this stock= 0.7939095
Average Relative Error for this stock= 0.01881279

AMZN.csv
Predicted price at time= 830.3857
Absolute Mean Error for this stock= 10.67571
Average Relative Error for this stock= 0.00303039

BABA.csv
Predicted price at time= 102.8071
Absolute Mean Error for this stock= 0.7628813
Average Relative Error for this stock= 0.02398417

FB.csv
Predicted price at time= 134.2775
Absolute Mean Error for this stock= 0.01911318
Average Relative Error for this stock= 0.01850999

GOOG.csv
Predicted price at time= 815.0638
Absolute Mean Error for this stock= 6.68384
Average Relative Error for this stock= 0.003072863

M.csv
Predicted price at time= 32.31045
Absolute Mean Error for this stock= 0.3304483
Average Relative Error for this stock= 0.07767482

MSFT.csv
Predicted price at time= 64.38453
Absolute Mean Error for this stock= 0.8365418
Average Relative Error for this stock= 0.03921757

NFLX.csv
Predicted price at time= 143.9418
Absolute Mean Error for this stock=0.7981618
Average Relative Error for this stock= 0.01716209

TSLA.csv
Predicted price at time= 272.5878
Absolute Mean Error for this stock= 10.50778
Average Relative Error for this stock= 0.009478178

TWTR.csv
Predicted price at time= 16.09678
Absolute Mean Error for this stock= 2.623215
Average Relative Error for this stock=0.1326945