**MATLAB PACKAGE**

**TEAM MATES:**

1. **Kalyani Menon (15PT14)**
2. **Prasanth S (15PT25)**
3. **Shaalini S (15PT31)**
4. **Shubhashri A G (15PT32)**
5. **Surya S (15PT35)**

**RAINFALL PREDICTION (Weather Forecasting)**

The world has come to a point where every single moment, investment and resource has so much of meaning and to predict the trend before its occurrence has become a thing of today. We have done an analysis on weather forecasting, so as to predict the amount of rainfall in the near future.(Data collected –past one and a half months :meteorological dept.(m.accuweather.com))

**CONCEPTS USED:**

**MULTIPLE REGRESSION ANALYSIS :**

Regression analysis with two or more independent variables, or at least one non-linear predictor is referred to as multiple regression analysis.

**Y(i)=b(0) + b(1)x(1) + b(2)x(2) +…………………….+ b(k)x(k)+e(i)**

In a multiple regression setting because of the potentially large number of predictors, it more efficient to use matrices to define the regression model and the subsequent analyses.

Consider the following simple linear regression function:

Y(i)=b(0) + b(1)x(1) + b(2)x(2) +…………………….+ b(k)x(k)+e(i) for i=1,2,…..,n

If we actually let i=1,…..,n,we see that we obtain n equations,

Y1=bo+b1x1+e1

Y2=b0+b1x2+e2

.

.

.

Yn=b0+b1xn+en

[y1 [ 1 x1 [e1

Y2 1 x2 [b0 e2

Y3 = 1 x3 b1] + e3

. . .

. . .

Yn] 1 xn] en]

Using the matrix notation, our simple linear regression function, reduces to a short statement.

Y=Xb+e

* X is an n x 2 matrix
* Y is an nx1 column vector, b is a 2x1 column vector and e is an nx1 column vector
* The matrix X and the vector b are multiplied together

**LEAST SQUARE ESTIMATES IN MATRIX NOTATION**

The px1 vector containing the estimates of the p parameters of the regression function can be shown to equal

b= [ b0

b1

b2 = **INV((X’X)) X’Y**

.

.

b(p-1)]

where inv(X’X) is the inverse of the X’X matrix, and X’is the transpose of the matrix.

EXPONENTIAL SMOOTHING METHOD:

Exponential smoothing methods weigh data from the previous time period with exponentially decreasing importance in the forecast. This method focuses upon the most recent data. Secondly, during forecasting, this method takes into account all the observed values because each smoothing value is based upon the values observed previously. In this manner, the values observed most recently receive the highest weight; the previously observed value receives the second highest weight and so on.

In exponential smoothing method, forecasting is carried out by multiplying the actual value for the present time period(t),by a value between 0 and 1(the exponential smoothing constant).This exponential constant is referred to as alpha. So ,the resultant value is alpha\*X(t).This value is added to the product of the present time period forecast F(t) and (1-alpha).Algebraically,

**F(t+1)=alpha\*X(t)+(1-alpha)\*F(t).**

Where F(t+1) is the forecast for the next time period (t+1),F(t) the forecast for the present time period (t),alpha the exponential smoothing constant(0<=alpha<=1) and X(t) the actual value for the present time period(t).

PROBLEM STATEMENT :

Based on the data collected from the metrological department on the amount of rainfall that occurred in Coimbatore for the months April and May 2016,the following analysis is done.

The dependent variables are,

1. temperature
2. humidity
3. pressure
4. wind velocity.

The independent variable is

a)amount of rainfall

The amount of rainfall is also collected. The regression coefficients are found and using that regression equation is formed (least square method).Similarly using back forecasting technique the forecast values are obtained and graphs are plotted to find and compare the accuracy of prediction.

CODE:

%%MULTIPLE LINEAR REGRESSION MODEL,EXPONENTIAL SMOOTHING FACTOR...

%%Based on the data collected from the meteorological department on the amount of rainfall and the factors that affect it,the following analysis is done.

%%The data pertains to the region of Coimbatore through the past one and a half months.

%%The dependent variable is amount of rainfall and the independent variables are TEMPERATURE,HUMIDITY,PRESSURE,WIND VELOCITY.

%%VARIABLES USED:

%%Y =DEPENDENT VARIABLE(amount of rainfall)

%%Xi=INDEPENDENT VARIABLE (X1=TEMPERATURE(CELCIUS),X2=HUMIDITY(PERCENTAGE),X3=PRESSURE(mb),X4=WIND VELOCITY(km/hr))

%%bi=LINEAR REGRESSION PARAMETERS

Y= [ 0

0

0

0

0

0

0

41

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

2

0

1

18

0

2]

X= [ 37 75 1005 1

37 68 1005 13

37 61 1007 18

37 44 1008 11

37 37 1008 8

37 51 1010 9

37 43 1009 7

37 38 1006 5

37 39 1006 18

37 37 1009 20

37 43 1008 11

37 52 1008 8

37 62 1010 9

37 71 1009 7

37 76 1006 5

37 81 1006 18

37 84 1009 24

37 87 1006 11

37 88 1008 9

37 89 1008 10

37 90 1010 9

37 94 1009 5

37 97 1005 17

37 99 1005 20

37 90 1007 14

37 75 1006 12

37 59 1006 10

36 48 1008 13

36 41 1007 13

36 36 1004 14

36 33 1003 15

36 34 1005 6

36 37 1002 10

36 39 1005 8

36 49 1005 5

36 57 1002 8

36 64 1002 9

36 76 1004 8

36 77 1002 12

36 82 1002 8 ]

fprintf('\nTHE REGRESSION COEFFICIENTS ARE:');

%%by least square method we obtain this formula,

B=inv(X'\*X)\*X'\*Y

fprintf('THE REGRESSION EQUATION IS:Y= %6.2f x1+ %6.2f x2 + %6.2f x3 + %6.2f x4',B(1),B(2),B(3),B(4));

f(1)=Y(1);

%%alpha (exponential smoothing factor =0.9)

%% forecast(t+1)=alpha\*amount of rainfall(t)+(1-alpha)\*forecast(t);

fprintf('\nTHE PREDICTED VALUE BY BACK FORECASTING(exponential smoothing) IS:');

for i=2:40

f(i)=(0.9)\*Y(i-1)+(0.1)\*f(i-1);

end

f

plot(Y,'-b')

xlabel('Days');

ylabel('Amount Of Rainfall');

text(14,700,'Observed Amount','EdgeColor','r','LineWidth',2);

hold on;

grid on;

plot(f,'-r')

legend('observed','forecast',0);

hold off;

RESULTS AND OUTPUT:

Y =

0

0

0

0

0

0

0

41

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

0

2

0

1

18

0

2

X =

37 75 1005 1

37 68 1005 13

37 61 1007 18

37 44 1008 11

37 37 1008 8

37 51 1010 9

37 43 1009 7

37 38 1006 5

37 39 1006 18

37 37 1009 20

37 43 1008 11

37 52 1008 8

37 62 1010 9

37 71 1009 7

37 76 1006 5

37 81 1006 18

37 84 1009 24

37 87 1006 11

37 88 1008 9

37 89 1008 10

37 90 1010 9

37 94 1009 5

37 97 1005 17

37 99 1005 20

37 90 1007 14

37 75 1006 12

37 59 1006 10

36 48 1008 13

36 41 1007 13

36 36 1004 14

36 33 1003 15

36 34 1005 6

36 37 1002 10

36 39 1005 8

36 49 1005 5

36 57 1002 8

36 64 1002 9

36 76 1004 8

36 77 1002 12

36 82 1002 8

THE REGRESSION COEFFICIENTS ARE:

B =

1.5380

-0.0468

-0.0479

-0.3355

THE REGRESSION EQUATION IS:Y= 1.54 x1+ -0.05 x2 + -0.05 x3 + -0.34 x4

THE PREDICTED VALUE BY BACK FORECASTING(exponential smoothing) IS:

f =

Columns 1 through 9

0 0 0 0 0 0 0 0 36.9000

Columns 10 through 18

3.6900 0.3690 0.0369 0.0037 0.0004 0.0000 0.0000 0.0000 0.0000

Columns 19 through 27

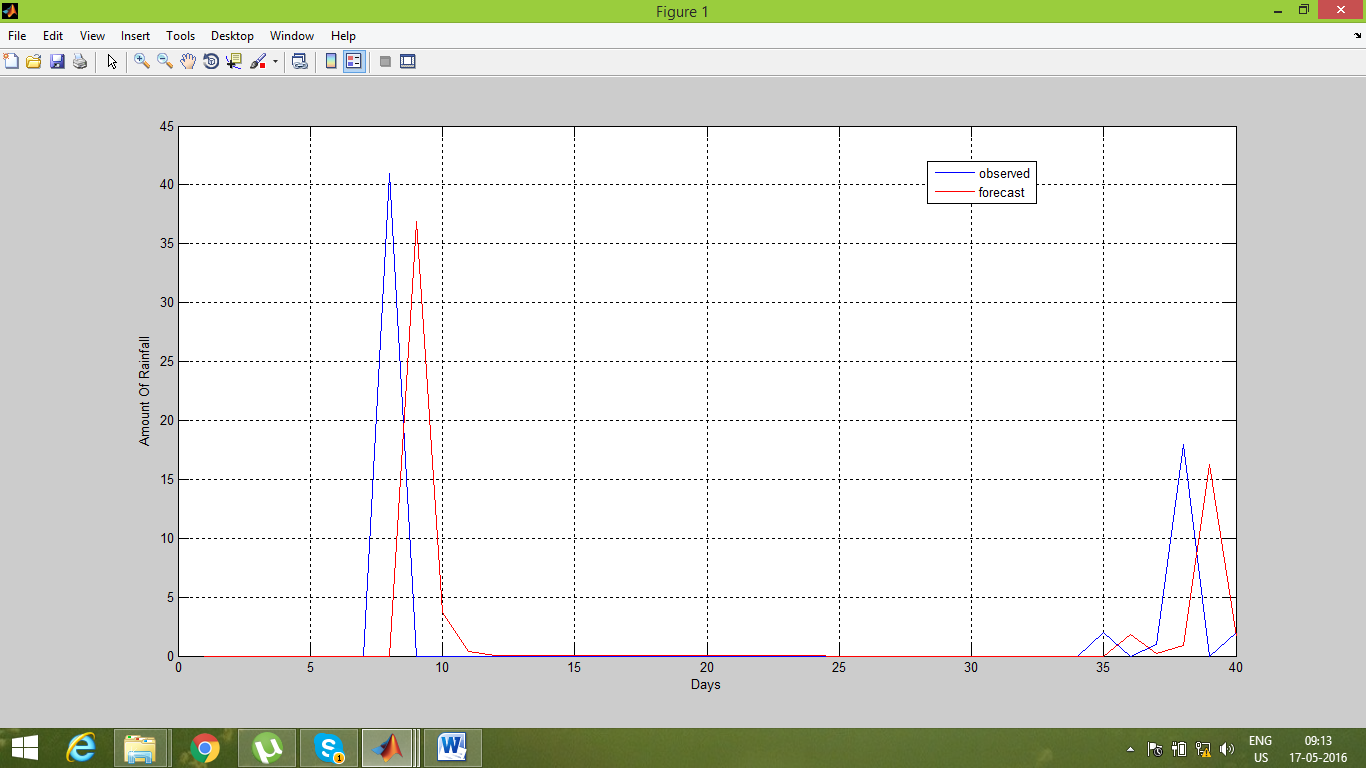
0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000

Columns 28 through 36

0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 1.8000

Columns 37 through 40

0.1800 0.9180 16.2918 1.6292



%%conclusion:The amount of rainfall likely to occur by the forecast

%%shows that it is almost accurate with the data present with the

%%metereological department(increasing trend)

%%The lower regression coefficients are due to less number of

%%observations which can be increased by having more number of

%%observations.

%%Interpretation:Keeping the humidity,pressure and wind velocity

%%constant for one unit increse in temperature,there is an increase

%%of 1.5380 units in the amount of rainfall.

%%Keeping the temperature,pressure and wind velocity constant,we

%%find that for an increse of 1 unit of humidity,there is a

%%decrease of 0.0468 units of rainfall.Similarly,the amount of

%%rainfall is inversely proportional to wind velocity and pressure.

%%The alpha smoothing factor is 0.9 which means it takes the

%%present data chareacteristics more into account than the past

%%statistics.

%%The graph shows the accuracy of the backforecasting method for

%%the next period of time(recent).