**Contents**

Abstract ------------------------------------------------------------------------------------

Acknowledgement ----------------------------------------------------------------------

Introduction ------------------------------------------------------------------------------

* 1. Background------------------------------------------------------------------
  2. Problem Statement-------------------------------------------------------
  3. Introduction to Data Used for this Research-----------------------

Literature Review----------------------------------------------------------------------

System Model--------------------------------------------------------------------------

3.1 Supervised Learning---------------------------------------------------

3.2 Model Representation------------------------------------------------

3.3 Cost function intuition ------------------------------------------------------------

3.4 Gradient descent intuition-----------------------------------------------

3.5 Linear Regression with multiple variables-----------------------------------

3.6 Feature Scaling---------------------------------------------------------------------

3.7 Learning rate----------------------------------------------------------------------

Implementation----------------------------------------------------------------------------

4.1 Input data format-----------------------------------------------------------

4.2 Coding snap shot-----------------------------------------------------------

Results ----------------------------------------------------------------------------------------

5.1 Aus Rice Analysis--------------------------------------------------------------

5.2 Aman Rice Analysis----------------------------------------------------------

5.3 Boro Rice Analysis------------------------------------------------------------

Conclusion and Future work -----------------------------------------------------------------

Reference --------------------------------------------------------------------------------------

**Abstract**

**Acknowledgement**

I cordially thank to Md. Rafsan Jani, Lecturer, Department of CSE, Jahangirnagar University, Savar for his constant support, careful supervision and guidance, keen interest, valuable criticism, active co-operation, suggestion and un-interrupted encouragement to carry out this research work.

Finally, I remember and express my gratitude to Almighty Allah for enabling me going fine with this research work continuously alongside my family and a lot of other critical unusual issues.

**Introduction**

* 1. **Background**

Bangladesh is a large deltaic plain formed under the influence of three mighty rivers –Ganges, Brahmaputra and Meghna with its flat topography, abundant water and sub tropical climate, constitutes an excellent habitat for the rice plant. Rice, as such, evolved as the staple food of the people of the country and historically associated with their culture, rites and rituals. With time, as the population increased at a rapid pace, the gap between rice production and food requirement for the millions widened. To feed the increasing population through radical change in rice production, replacement of the low-yielding traditional varieties and age old production practices of rice by high-yielding varieties and improved production technologies became essential. Rice research in this part of the sub-continent started in 1910 that got momentum in mid 60s.

Realizing the importance of rice in food security and political stability of the country, an autonomous organization in the name of East Pakistan Rice research Institute (EPRRI) was established on 1 October 1970 on 76.82 hectare of land at Gazipur, 36 km north of the capital city Dhaka, which was renamed as the Bangladesh Rice Research Institute (BRRI) after the independence in 1971.

Since its establishment in 1970, BRRI has been serving the nation through development of high-yielding rice varieties and improved production practices, which have been instrumental almost in tripling the annual rice production within the last 45 years. For this, BRRI has earned a very high reputation in Bangladesh as well as in the world rice community.

The high-yielding modern varieties (MVs) developed by BRRI [**presently covers**](http://brri.portal.gov.bd/sites/default/files/files/brri.portal.gov.bd/page/4c36b7d3_92e1_473d_8cd0_3d0e00a8115b/Rice%20Area%20in%20Bangladesh_21.01.2018.pdf) 82% of the Boro (winter rice), 36% of the Aus (summer rice), and 47% of the transplant Aman (wet season rice) areas of Bangladesh. The overall adoption MVs in Bangladesh is 79.55%. These varieties account for about 85% of the total [annual rice production](http://brri.portal.gov.bd/sites/default/files/files/brri.portal.gov.bd/page/4c36b7d3_92e1_473d_8cd0_3d0e00a8115b/Rice%20Production%20in%20Bangladesh_21.01.2018.pdf)in the country. The BRRI MVs and technology packages played the key role in boosting annual rice production in Bangladesh from  9.77 million metric tons in 1970-71 to 34201.50 thousand metric tons in 2016-17. Without BRRI MVs, rice production would have increased at the rate of 1% annually, almost half the rate at which the population grew during this period.

BRRI MVs and production technologies benefited the nation in the form of low cost of living particularly, of huge population living in rural areas and urban slums. But, there were indirect additional benefits to the society too. The government saved huge amount of foreign exchange which, if there were no BRRI MVs, could be spent for food grain imports to feed the country’s growing population. It has been estimated that the rate of return per taka investment in rice research and development is Tk 46. Additional production from BRRI MVs also kept the domestic rice price relatively stable and within the reach of the common people. In fact, since early 1980s the import of food grains declined steadily and the country approached self-sufficiency by 1990s.

BRRI technologies also contributed to income generation and employment in rural Bangladesh over the last 44 years. In areas where the MV technology has been introduced, the proportion of population living below the poverty line is 51% compared with 78% for areas without such technological progress. The net return per agricultural holding using MV technology is about 50% higher than a similar holding using traditional varieties. The expansion of modern irrigation facilities, with the expansion of MV rice acreage, has also led to increased employment opportunities in both agricultural and non-agricultural sectors, with a rise in the income of the rural population. MV adoption also created indirect employment opportunity such as in fertilizer trade and in the maintenance of pumps and other equipments, for example.

As recognition of its outstanding contribution in the field of rice research and development, BRRI has been honoured with the following prestigious national and international awards:

* Bangabandhu Award in 1974
* President's Gold Medal in 1977
* Independence Day Gold Medal in 1978
* President's Gold Medal in 1980
* FAO Bronze Plaque in 1980
* President's Gold Medal in 1984
* National Environmental Award in 2009.
* Mercantile Bank Limited Award  2013.
* Metropolitan Chamber of Commerce and Industry (MCCI) Award 2014.
* KIB Agriculture Award-2015
* Awarded National ICT Award-2016 at Digital World 2016 for ICT Excellence through Innovative Service Delivery.
  1. **Problem Statement**

The proposed research project work aims to develop Bangladesh’s rice production prediction model. The change of climate has great impact on rice production in Bangladesh. In this research work, we will consider area, rainfall as independent factor on rice production. We have used rice production, rainfall data since 1971-2017 for building prediction model.

* 1. **Introduction to Data Used for this Research**

To understand the problem, we researched the data in some statistical way and prepared the data for the machine learning. In the preparation stages we went through the following stages.

**Raw Data Stage:** BMD (Bangladesh Meteorological Department). What we got form the BMD is the row data as below.

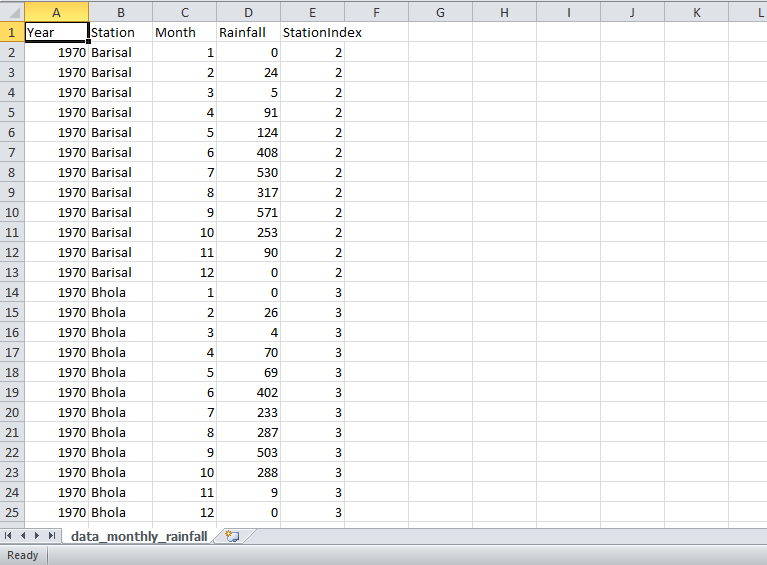


Fig-Rainfall Raw data from BMD

Rainfall data has column:

* Station
* Year
* Months
* 1(Day 1)
* 2(Day 2)
* 3(Day 3)

-

-

- 30(Day 30)

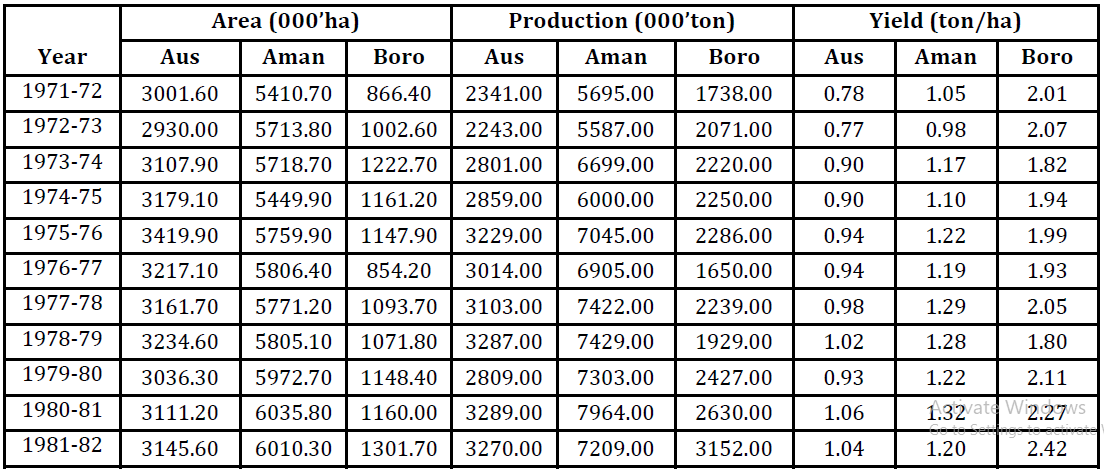
- 31(Day 31)

Rainfall Data structure like:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Day | Day |  | Day | Day |
| Station | Year | Month | 1 | 2 | ------- | 30 | 31 |
| Dhaka | 1971 | 1 | 0 | 0 | ---- | 0 | 0 |
| Dhaka | 1971 | 2 | 0 | 0 | ---- | 0 | 0 |
| Dhaka | 1971 | 3 | 0 | 0 | ------ | 1 | 4 |
| Dhaka | 1971 | 4 | 0 | 22 | ---- | 4 | 0 |

Again, we have collected rice production data from Bangladesh Rice Research Institute.

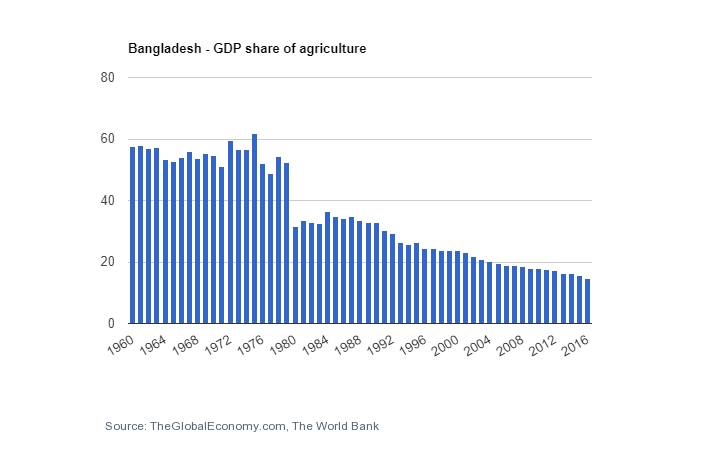
Raw Rice Production data looks like:



**Literature Review:**

Bangladesh’s economy is largely based on the agricultural sector which contributes 15% of national GDP and in which 39.07% of the labour force are engaged. Rice production is central to this sector, not only do the majority of Bangladesh’s farmers depend directly and indirectly on the success of the rice crop each year, as it is the main food staple, rice production is a big factor in the national effort to promote food security. Despite the importance of rice farming in the Bangladeshian landscape, it has traditionally been dependant on rainfall.

Agricultural contribution on our national GDP:



It is noticeable that the contribution of agriculture on GDP is decreasing day by day. This is warning this to us also.

There are some research on rice productivity in Bangladesh. Crop productivity is a function of meteorological(rainfall, temperature, relative humidity and incoming solar radiation quality) and non-meteorological factors such as the type of seeds, site characteristics and management practices (including irrigation, fertilizer and pesticide application). It is often a challenging task to differentiate the impact of non-meteorological factors, more specifically the technological inputs. In order to study the pattern of trends and to quantify the growth rate based on technological inputs, the actual productivity was fitted with timer set linear curve over time. The technological productivity (TP*i*) was calculated by referring to the research of Subash

and Mohan [37]:

Where are constants determined empirically and  1, 2, 3,… n represented from 1996–1997 to 2008–2009 for rice. To normalize the productivity data, the rice productivity index (RPI) was used that tends to extract the percentage of the technological driven productivity to the actual productivity. The normalized RPI for the year is

Where is the rice productivity index of rice for the year, *Pi* is the actual productivity for the year and is the technological trend productivity for the year.

There some papers about impacts of climate changes on rice productivity. Most of the papers have used Linear Regression with single independent factor. In this research paper we have taken two independent factor for rice productivity.

Our hypothesis :

Where is constants and are independent factors ( )

**System Model**

**3.1 Supervised Learning**

In supervised learning, we are given a data set and already know what our correct output should look like, having the idea that there is a relationship between the input and the output.

In order to solve a given problem of supervised learning, one has to perform the following steps:

1. Determine the type of training examples. Before doing anything else, the user should decide what kind of data is to be used as a training set. In case of handwriting analysis, for example, this might be a single handwritten character, an entire handwritten word, or an entire line of handwriting.
2. Gather a training set. The training set needs to be representative of the real-world use of the function. Thus, a set of input objects is gathered and corresponding outputs are also gathered, either from human experts or from measurements.
3. Determine the input feature representation of the learned function. The accuracy of the learned function depends strongly on how the input object is represented. Typically, the input object is transformed into a [feature vector](https://en.wikipedia.org/wiki/Feature_vector), which contains a number of features that are descriptive of the object. The number of features should not be too large, because of the [curse of dimensionality](https://en.wikipedia.org/wiki/Curse_of_dimensionality); but should contain enough information to accurately predict the output.
4. Determine the structure of the learned function and corresponding learning algorithm. For example, the engineer may choose to use [support vector machines](https://en.wikipedia.org/wiki/Support_vector_machine) or [decision trees](https://en.wikipedia.org/wiki/Decision_tree_learning).
5. Complete the design. Run the learning algorithm on the gathered training set. Some supervised learning algorithms require the user to determine certain control parameters. These parameters may be adjusted by optimizing performance on a subset (called a *validation* set) of the training set, or via [cross-validation](https://en.wikipedia.org/wiki/Cross-validation_(statistics)).
6. Evaluate the accuracy of the learned function. After parameter adjustment and learning, the performance of the resulting function should be measured on a test set that is separate from the training set.

Supervised learning problems are categorized into "regression" and "classification" problems. In a regression problem, we are trying to predict results within a continuous output, meaning that we are trying to map input variables to some continuous function. In a classification problem, we are instead trying to predict results in a discrete output. In other words, we are trying to map input variables into discrete categories.

**Example 1:**

Given data about the size of houses on the real estate market, try to predict their price. Price as a function of size is a continuous output, so this is a regression problem.

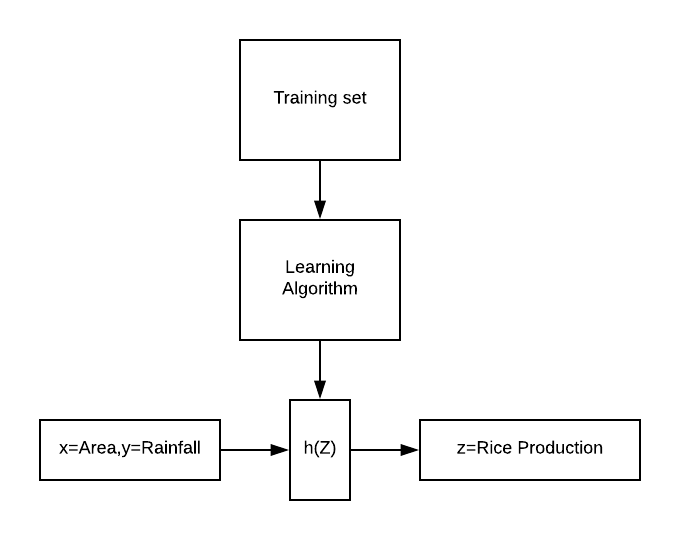
We could turn this example into a classification problem by instead making our output about whether the house "sells for more or less than the asking price." Here we are classifying the houses based on price into two discrete categories.

**Example 2**:

(a) Regression - Given a picture of a person, we have to predict their age on the basis of the given picture

(b) Classification - Given a patient with a tumor, we have to predict whether the tumor is malignant or benign.

**3.2 Model Representation**

****

To establish notation for future use, we’ll use   to denote the “input” variables(Area,Rainfall), also called input features, and  to denote the “output” or target variable that we are trying to predict(rice production). A pair () is called a training example, and the dataset that we’ll be using to learn.

**3.3 Cost Function Intuition**

We can measure the accuracy of our hypothesis function by using a **cost function**. This takes an average difference (actually a fancier version of an average) of all the results of the hypothesis with inputs from x's and the actual output z's.

This function is otherwise called the "Squared error function", or "Mean squared error". The mean is halved () as a convenience for the computation of the gradient descent, as the derivative term of the square function will cancel out the ​term.

Here,

H(x) is hypothesis

is predicted output

is error between exact value and predicted value.

3.4 Gradient Descent Intuition

So we have our hypothesis function and we have a way of measuring how well it fits into the data. Now we need to estimate the parameters in the hypothesis function. That's where gradient descent comes in.

Imagine that we graph our hypothesis function based on its fields *θ*0​ and *θ*1​ (actually we are graphing the cost function as a function of the parameter estimates). We are not graphing x and y itself, but the parameter range of our hypothesis function and the cost resulting from selecting a particular set of parameters.

We put *θ*0​ on the x axis and *θ*1​ on the y axis, with the cost function on the vertical z axis. The points on our graph will be the result of the cost function using our hypothesis with those specific theta parameters. The graph below depicts such a setup.



We will know that we have succeeded when our cost function is at the very bottom of the pits in our graph, i.e. when its value is the minimum. The red arrows show the minimum points in the graph.

The way we do this is by taking the derivative (the tangential line to a function) of our cost function. The slope of the tangent is the derivative at that point and it will give us a direction to move towards. We make steps down the cost function in the direction with the steepest descent. The size of each step is determined by the parameter α, which is called the learning rate.

For example, the distance between each 'star' in the graph above represents a step determined by our parameter α. A smaller α would result in a smaller step and a larger α results in a larger step. The direction in which the step is taken is determined by the partial derivative of *J*(*θ*0​,*θ*1​). Depending on where one starts on the graph, one could end up at different points. The image above shows us two different starting points that end up in two different places.

The gradient descent algorithm is:

repeat until convergence:

Where j=0,1 represents the feature index number.

At each iteration j, one should simultaneously update the parameters *θ*1 ​,*θ*2​,...,*θn*​. Updating a specific parameter prior to calculating another one on the  iteration would yield to a wrong implementation.

When specifically applied to the case of linear regression, a new form of the gradient descent equation can be derived. We can substitute our actual cost function and our actual hypothesis function and modify the equation to :

|  |
| --- |
| repeat until convergence: {  } |

where m is the size of the training set, *θ*0​ a constant that will be changing simultaneously with *θ*1​ and values of the given training set (data).

Note that we have separated out the two cases for *θj*​  into separate equations for *θ*0​ and *θ*1​; and that for *θ*1​ we are multiplying  at the end due to the derivative. The following is a derivation of  for a single example:



The point of all this is that if we start with a guess for our hypothesis and then repeatedly apply these gradient descent equations, our hypothesis will become more and more accurate.

So, this is simply gradient descent on the original cost function J. This method looks at every example in the entire training set on every step, and is called **batch gradient descent**. Note that, while gradient descent can be susceptible to local minima in general, the optimization problem we have posed here for linear regression has only one global, and no other local, optima; thus gradient descent always converges (assuming the learning rate α is not too large) to the global minimum. Indeed, J is a convex quadratic function. Here is an example of gradient descent as it is run to minimize a quadratic function.



The ellipses shown above are the contours of a quadratic function. Also shown is the trajectory taken by gradient descent, which was initialized at (48,30). The x’s in the figure (joined by straight lines) mark the successive values of θ that gradient descent went through as it converged to its minimum.

**3.5 Linear Regression with multiple variables**

Linear regression with multiple variables is also known as "multivariate linear regression".

We now introduce notation for equations where we can have any number of input variables.

|  |
| --- |
| = value of feature *j* in the   training example  *M = Number of training examples*  *N = Number of features* |

The multivariable form of the hypothesis function accommodating these multiple features is as follows:

Using the definition of matrix multiplication, our multivariable hypothesis function can be concisely represented as:

This is a vectorization of our hypothesis function for one training example; see the lessons on vectorization to learn more.

The gradient descent equation itself is generally the same form; we just have to repeat it for our 'n' features:

Repeat until convergence : {

}

In other words:

Repeat until convergence: {

for j:=0……n

}

**3.6 Feature Scaling**

We can speed up gradient descent by having each of our input values in roughly the same range. This is because θ will descend quickly on small ranges and slowly on large ranges, and so will oscillate inefficiently down to the optimum when the variables are very uneven.

The way to prevent this is to modify the ranges of our input variables so that they are all roughly the same. Ideally:

Or

These aren't exact requirements; we are only trying to speed things up. The goal is to get all input variables into roughly one of these ranges, give or take a few.

Two techniques to help with this are **feature scaling** and **mean normalization**. Feature scaling involves dividing the input values by the range (i.e. the maximum value minus the minimum value) of the input variable, resulting in a new range of just 1. Mean normalization involves subtracting the average value for an input variable from the values for that input variable resulting in a new average value for the input variable of just zero. To implement both of these techniques, adjust your input values as shown in this formula:

Where  is the **average** of all the values for feature (i) and  is the range of values (max - min), or  is the standard deviation.

**3.7 Learning Rate**

Debugging gradient descent**.** Make a plot with *number of iterations* on the x-axis. Now plot the cost function, J(θ) over the number of iterations of gradient descent. If J(θ) ever increases, then you probably need to decrease α.

Automatic convergence test**.** Declare convergence if J(θ) decreases by less than E in one iteration, where E is some small value such as . However in practice it's difficult to choose this threshold value.

To summarize:

If  is too small: slow convergence.

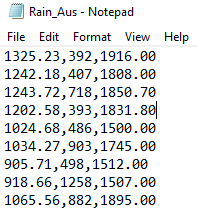
If  is too large: may not decrease on every iteration and thus may not converge.

**Implementation**

**4.1 Input data format**

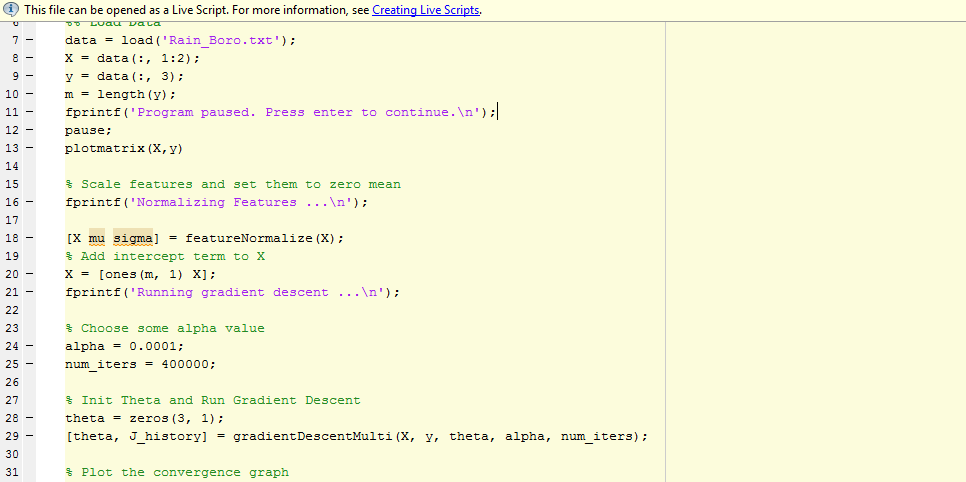
|  |  |  |
| --- | --- | --- |
| Cultivation Area | Rainfall | Rice Production |
| 1325.23 | 392 | 1916.00 |
| 1242.18 | 407 | 1808.00 |

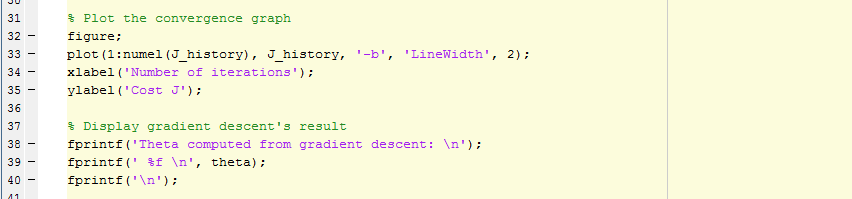
Data snap shot:



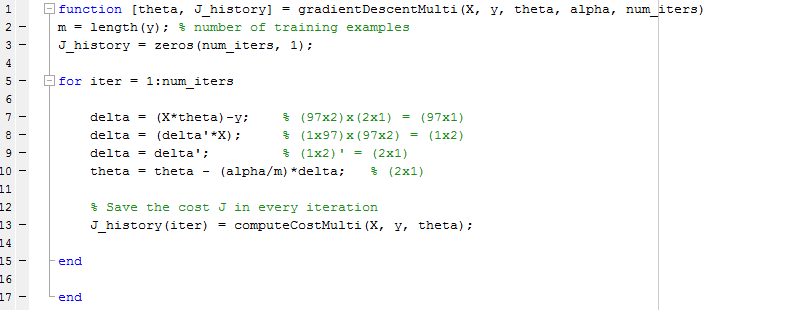
**4.2 Coding snap shot**

Main program:

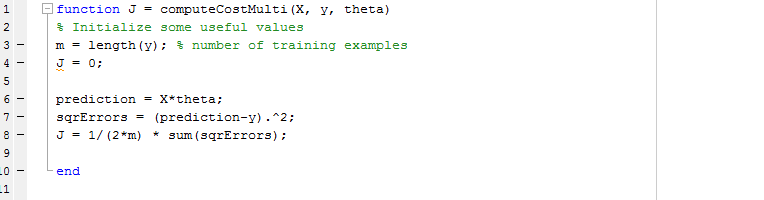




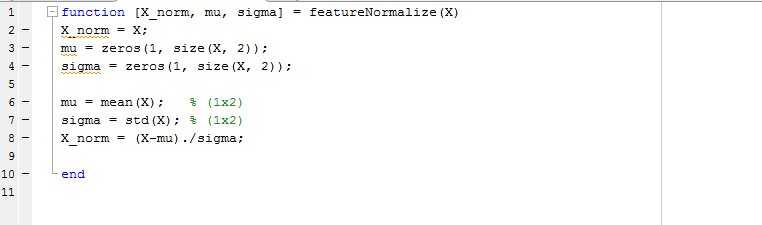
Gradient Descent Algorithm:



Cost Function:



Feature Scaling:



**Results**

**5.1 Aus Rice Analysis**

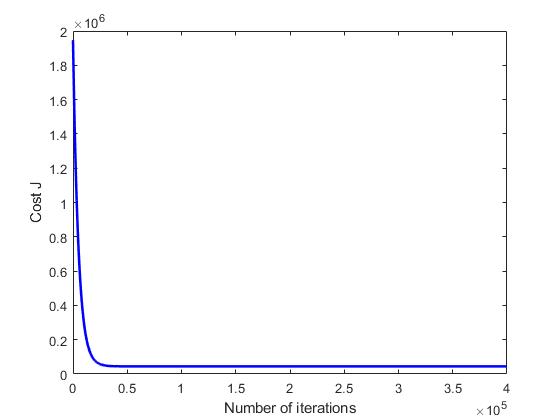


Fig-Convergence of Gradient Descent for Aus Rice

**Optimal constant :**

**5.2 Aman Rice Analysis**

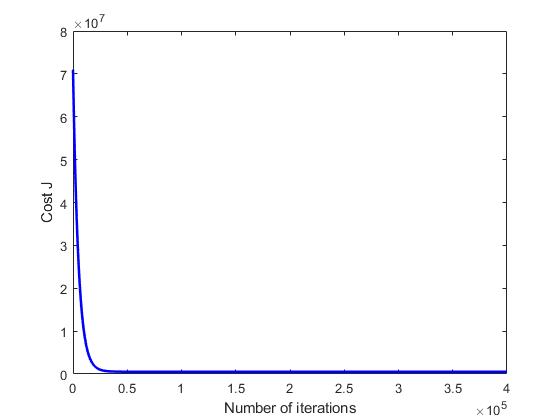
****

Fig-Convergence of Gradient Descent for Aman Rice

Optimal constant:

**5.3 Boro Rice Analysis**

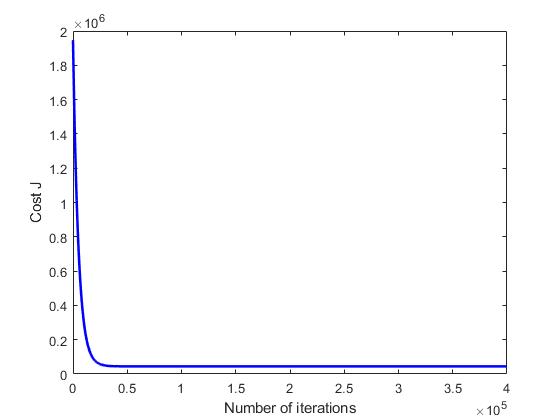
****

Fig-Convergence of Gradient Descent for Boro Rice

Optimal constant:

**Conclusion and Future work**

The observed annual rainfall over Bangladesh varies from location to location. It is the lowest of 1607.8 mm (at Satkhira) and the highest of 2608.7 mm (at Patuakhali). The standard deviation of annual rainfall over this region varies from 334.0 to 586.3 mm.

Boro rice is an irrigated crop in Bangladesh, mainly depended on irrigation application water but the variability of rainfall in winter season is an important factor on Boro production in Bangladesh. Variability of rainfall affects the rice crop at different times. If the variability is associated with the onset of the rain, stand establishment and the growth duration of rice are affected. If variability is associated with an untimely cessation at the reproductive or ripening stage of the rice crop, yield reduction is severe.

From the analysis of historical data of rainfall over the last 46 years (1971-2016), shows decreasing trend of rainfall pattern in winter season among the most weather stations in Bangladesh. This trend is also consistent with the general climate change predictions (IPCC, Third assessment report). On the basis of rainfall trend in winter season and various predictions results, this simulation study has been conducted to reduce rainfall amount 5 mm and 10 mm to see the effect on rice production in Bangladesh.

If the population grows to 233.2 million by the year 2050, then the problem will be more acute. Currently, maximum area in Bangladesh is under HYV Boro rice cultivation. It is expected that in future the total rice area would be under HYV cultivation. However, even then the rice production is not projected to be adequate to meet the demand. Therefore, development of cultivars that would be able to withstand high temperature (20C to 40C temperature) and water stress and have greater photosynthesis efficiency may mitigate the production problems projected for the next 50 to 100 years. Besides proper management practices (irrigation and fertilizer applications) would also help to meet the food demand under changing climatic condition in future. Public awareness of the impact of climate change on the agricultural production systems deserves priority consideration, and mitigating technologies must be developed, which will require increased public and private investment.

Increasing productivity requires new knowledge-both to maintain yields and to improve the quality of production. The needed knowledge is primarily biological in nature, but also includes the social science and technical knowledge.

Maximum agricultural researches in public sectors, very few number private organizations have directly relationship with agricultural research. But the private organizations may contribute a significant role to improve these sectors.

There will need to be more reliance on scientific knowledge and assessment of viable options and bridging the gap among policy makers, research organizations, agricultural extension workers and farmers. Besides very few number of our farmers have knowledge about suitable growing period of crop, irrigation and fertilizer application schedule for adapting new climatic conditions.

In this research project work, we have considered Rainfall and Area as independent factor for producing rice. Since we also consider temperature, humidity impact on rice production then prediction model would work more accurately. Since we can build a application using this prediction model then government could measure whether there is shortage of rice or not.

**Reference**

[1] Sindhu P. Menson,Ramith Bharadhwaj,Pooja shetty. (2017).Prediction of Temperature using Linear Regression. ICEECCOT

[2] Haruhisa Asada, Jun Matsumoto. Effect of rainfall variation on rice production in the Ganges-Brahmaputra Basin,Vol. 38: 249-260,2009

[3] Md Abiar Rahman, suchul kang, Nidhi Nagabhatla, Robert Macnee. (2017). Impacts of temperature and rainfall variation on rice productivity in major ecosystems of Bangladesh.Agric & Food Secur, PP: 6-10

[4] Kabir H, Golder. Rainfall Variability and Its Impact on Crop Agriculture in Southest Region of Bangladesh. J Climatol Weather Forecasting, PP:1-5, 2017

[5] Yahya Eru Cakra, Bayu Distiawan Trisedya. Stock Price Prediction using Linear Regression based on sentiment Analysis. ICACSIS,2005