MACHINE LEARNING

My home ► My courses ► Managed Courses ► Semester 1601 ► 1601-COL341 ► Assignments ► Assignemnt1: Basic linear algebra using R

Description Submission Edit Submission view

Assignemnt1: Basic linear algebra using R

Due date: Friday, 19 August 2016, 10:00 PM
Maximum number of files: 1
Type of work: Individual work

Objective:

In this assignment we will be testing your understanding of Linear Algebra. You should have a good idea of column space of a matrix, null space of a matrix and rank of a matrix. Before starting this assignment. The aim of this assignment is to make you understand the concepts of rank, null space and column space of a matrix. Solving homogeneous and non-homogeneous matrix equation and calculating the rank of a matrix. Possible cases and no of solutions to the problem.

Basic Learning:

1) Column space of a matrix: It is the set of all possible linear combinations of columns of A.

$$\text{For } A = \begin{bmatrix} 3 & 1 & 4 \\ 2 & 0 & 1 \\ 1 & 3 & 5 \end{bmatrix}, column\ space = c_1 \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} + c_3 \begin{bmatrix} 4 \\ 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 3c_1 + c_2 + 4c_3 \\ 2c_1 + c_3 \\ c_1 + 3c_2 + 5c_3 \end{bmatrix}$$

All the rows are independent as there are three constants that can have any value hence the rank = 3 = dimention of column space (remember, it is set of all possible vectors, not just one vector).

The null matrix always have a rank 0. A matrix spanning the column space is a matrix whose linear combination of columns can give any vector in the column space. One such Matrix for above A will be

$$\begin{bmatrix} 7 & 5 & 4 \\ 3 & 1 & 2 \\ 6 & 8 & 4 \end{bmatrix}$$

Or it can be A itself. U will always be of the dimension mxk where m is the no of rows in A and k is the rank of column space of A.

2) The null space is the solution set of the homogeneous equation Ax=0. In case of exactly one solution x is a null matrix and its rank is zero. In case of infinite no of solutions its rank will be greater than 0 as already seen in the above case.

Problem statement:

The programming must be done in R and you are **NOT** allowed to use the rank or any other high-level function in R. Your input will be in file input.txt containing a matrix A of size m x n and a vector y of size m x 1. The input file will be csv with a Matrix of size m x n+1 where the first n columns will correspond to matrix A and the last column will be y. There will be three sub-problems of the assignment. For format of output, refer to the 'input/output format' section.

For all the questions below, assume the dimensions of A to be m x n.

1) Find rank (rc) of column space of A, and a matrix U whose columns span the column space of A i.e. every vector in column space of A could be represented as a linear combination of columns of U. There are many possible solutions for U you need to find one such solution. Output: The matrix U, size m x rc in csv format

2) Find rank of null space (rn) of A and a matrix U whose columns span the null space of A. There are many possible solutions for U you need to find one such solution. Output: The matrix U,size n x rn in csv format.

3) Find all possible solutions of the equation Ax = y. Output the integer 0 if there exist no solution to the problem. If the solution is unique, then output it as a vector (i^{th} component of the vector in i^{th} line). However, in the case of multiple solutions we need a way to generate all possible solutions of Ax = y. For this, note that if x1, x2, … xk are the solutions of Ax = y then the vector

$$X = X_1 + \sum_{i=1}^{k} W_i(X_i - X_1)$$

is also a solution for all values of the vector w. Output the solutions x1, x2, .. xk in the form of a nxk matrix such that they generate all the solutions. You need to think the minimum no of solutions (k) required to generate all the possible solution and then output those k solutions. Output: A Matrix U size mxk whose k columns are the k solutions.

Input / Output Format:

Input Ranges: 100 <= m <= 10,000; 100 <= n <=10,000

Input must be taken from a file named input.txt and output written to three files out_1.txt , out_2.txt and out_3.txt for parts 1, 2 and 3 respectively.

There will be one input per file. The input will be A in csv format and the last column will be the vector y. the output should be written in three files out_1.txt, out_2.txt, out_3.txt. Each of them should contain the matrix U in csv format. In the last case if there is no solution output 0 (1 line in output file), if there is one solution output the vector x (size nx1, so output file will have n lines ith component of vector in ith line).

We will be giving you example files for input and corresponding output formats for further clarification.
6 sample inputs are available at:
http://moodle.iitd.ac.in/pluginfile.php/44429/mod_vpl/intro/input.zip

References:

Introduction of Linear Algebra (South Asian Edition, 4th Edition, 2009) Gilbert Strang, Chapter 3 – Vector Spaces and Subspaces.

R tutorials

http://www.tutorialspoint.com/r

In case of any query mail (with subject 'MLASS1') to:
Harshit Patel: psharshit17@gmail.com
Mohd Saood: the.saood@gmail.com

VPL 3.1.3

NAVIGATION