Unit I

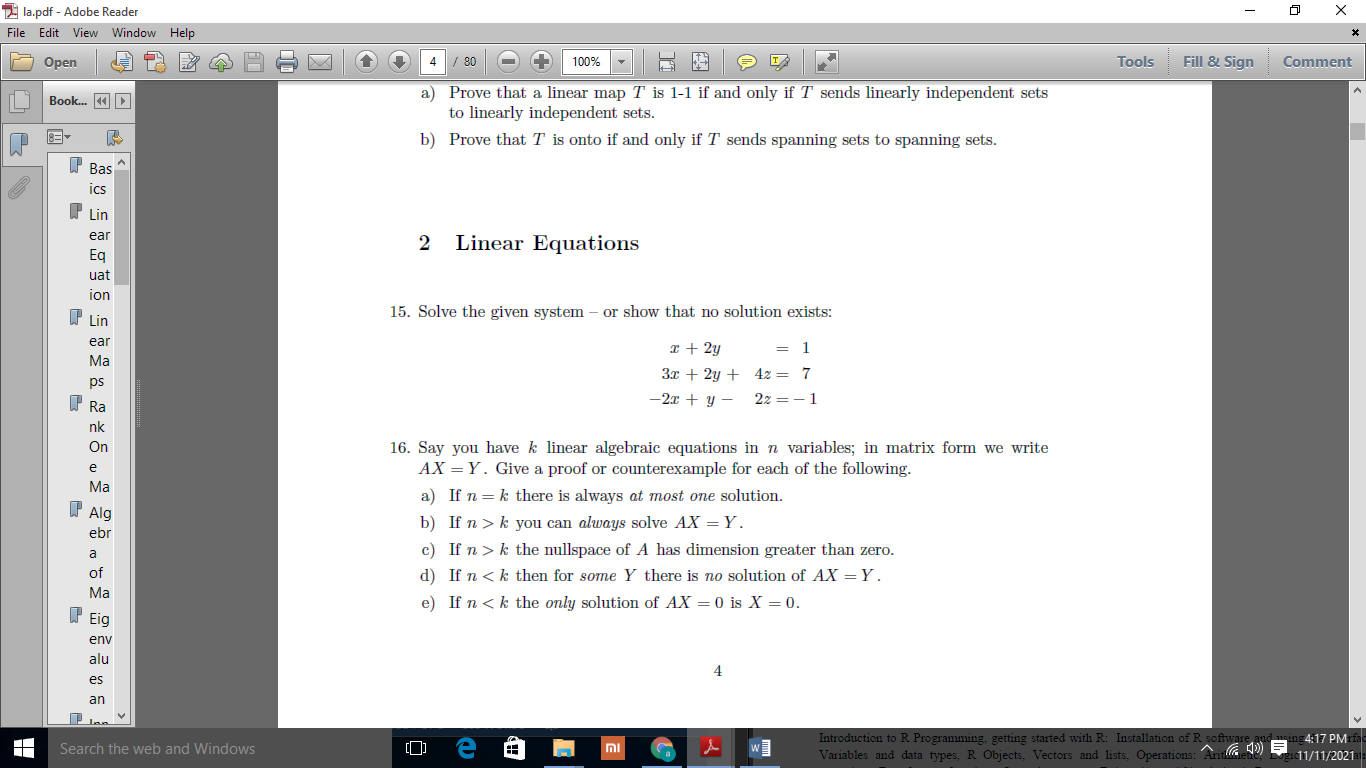
How do you find half spaces?

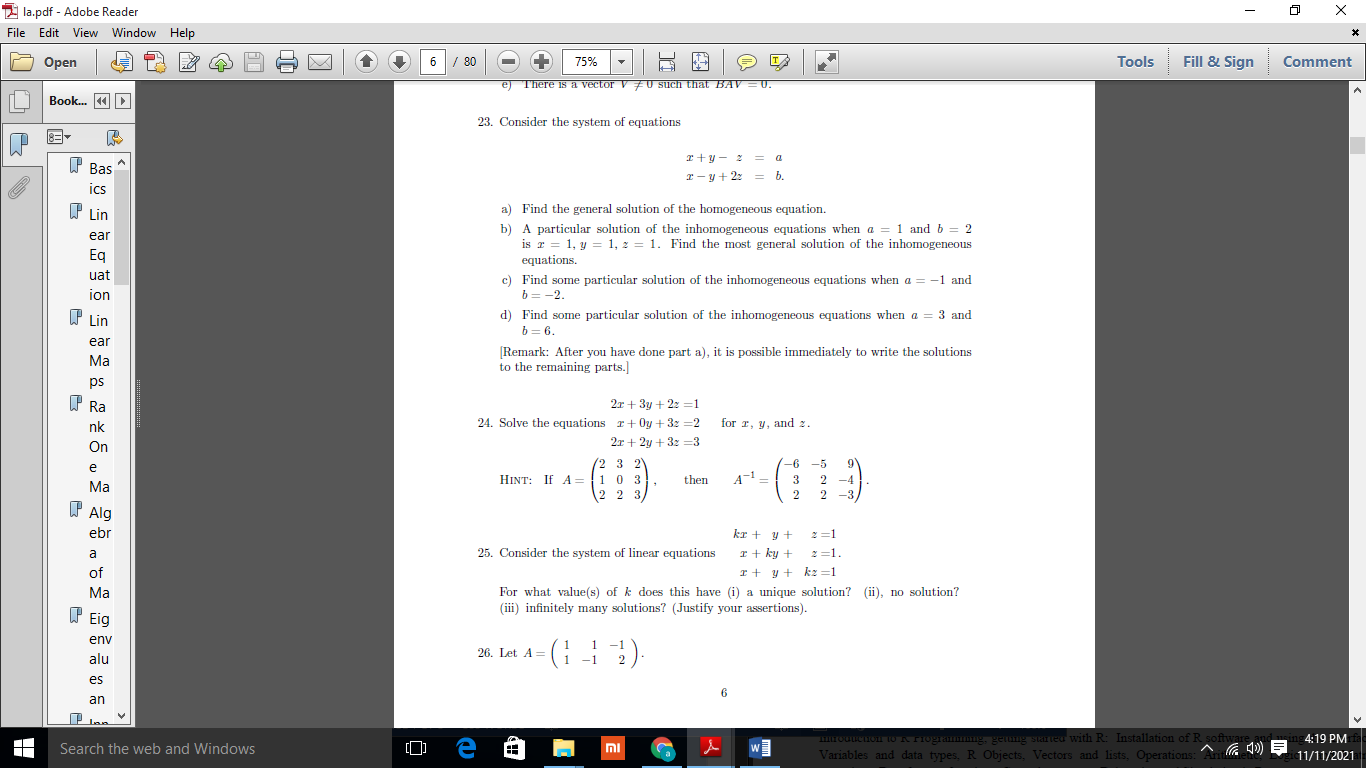
What are Halfspaces?

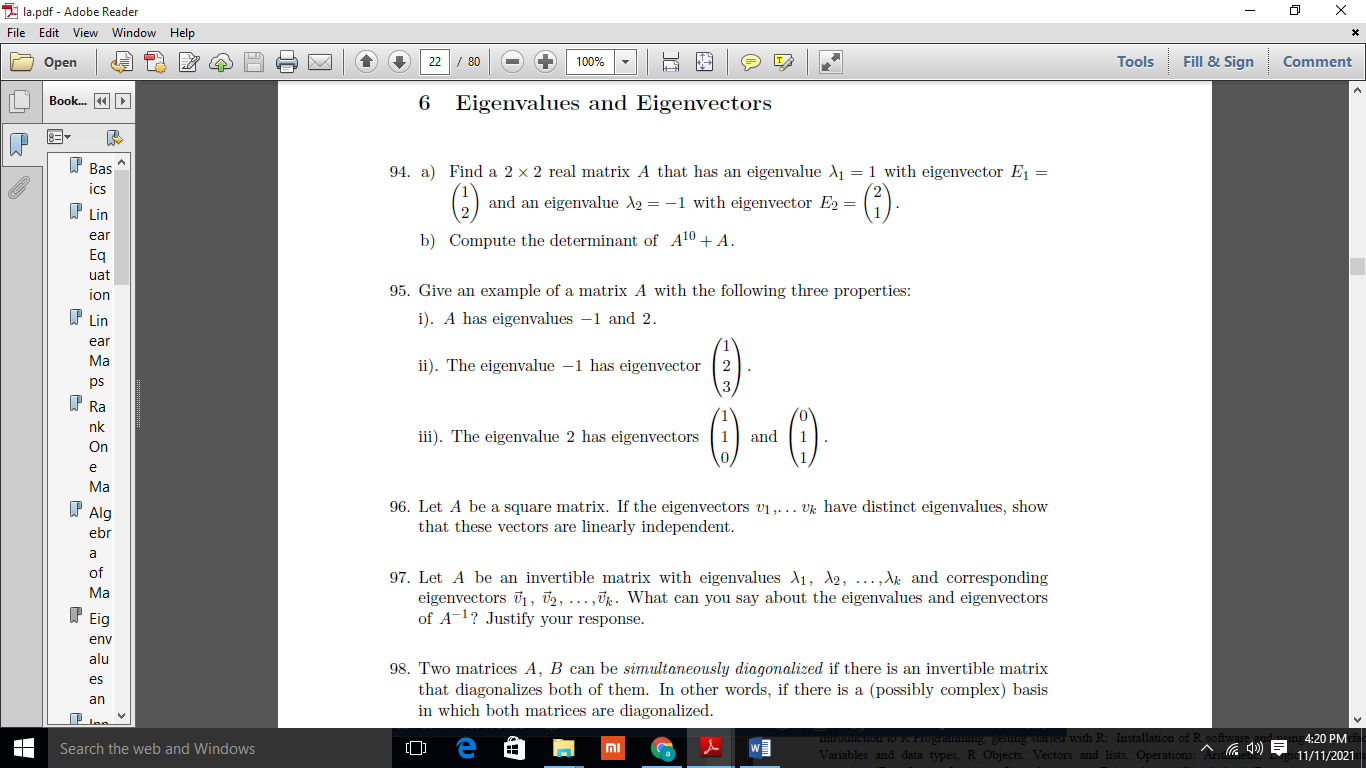
What is half-space in optimization?

What is open half plane?

Why is half space not affine?







Unit II

1. What is statistical modeling with example?
2. What comes under statistical Modelling?
3. What is the common goal of statistical Modelling?
4. What are the 3 types of random variable?
5. What is a random variable explain with an example?
6. What is an example of testing a hypothesis?
7. What are some examples of how hypothesis testing can be applied in everyday life?
8. How do you calculate hypothesis testing?
9. Are hypothesis questions?
10. Why hypothesis testing is used?

Unit III

1. What are the applications of predictive Modelling?
2. Which algorithm's would you use to build a predictive model?
3. What fields can predictive modeling be used in?
4. What is a real life example of linear regression?
5. What is an example of a question that can be put to a regression analysis?
6. What are the major problems of linear regression?
7. Is linear regression supervised learning?
8. How do you write a linear regression question?
9. What is an example of a question that can be put to a regression analysis?
10. What is the model used for simple linear regression?
11. How do you write a linear regression research question?
12. Simple Linear Regression model building
13. What is the purpose of building a regression model?
14. What kind of questions can be answered by multiple linear regression?
15. How do you write a multiple regression research question?
16. What is multiple linear regression example?
17. What is the problem in multiple linear regression?
18. What type of research uses multiple regression?
19. What are the assumptions of multiple regression?
20. What do you mean by the Logistic Regression?

## What are the different types of Logistic Regression?

## Explain the intuition behind Logistic Regression in detail.

## What factors can attribute to the popularity of Logistic Regression?

UNIT IV

### 1. What is R?

### 2. Can you write and explain some of the most common syntax in R?

### 3. How do you list the preloaded datasets in R?

### 4. What are some advantages of R?

### 5. What are the disadvantages of R?

### 6. What are the similarities and differences between R and Python?

### 7. Write code to accomplish a task

### 8. When is it appropriate to use the “next” statement in R?

### 9. How do you assign a variable in R?

### 10. What are the different data types/objects in R?

### 11. What are the objects you use most frequently?

### 12. Why use R?

### 13. What are some of your favorite functions in R?

### 14. Write a custom function in R

### 15. How do you import data in R?

### 16. How do you install a package in R?

### 17. What is the use of with() in R?

### 18. What is the use of by() in R?

### 19. When is it appropriate to use mode()?

### 20. What is a factor variable, and why would you use one?

### 21. When is it appropriate to use the which() function?

### 22. How do you concatenate strings in R?

### 23. How do you read a CSV file in R?

### 24. What are 3 sorting algorithms available in R?

### 25. Can you create an R decision tree?

### 26. Why is R useful for data science?

### 27. Describe how R can be used for predictive analysis

1) Explain what is R?

2) List out some of the function that R provides?

3) Explain how you can start the R commander GUI?

Questions and Answers ????

4) In R how you can import Data?

5) Mention what does not ‘R’ language do?

6) Explain how R commands are written?

7) How can you save your data in R?

8) Mention how you can produce co-relations and covariances?

9) Explain what is t-tests in R?

10) Explain what is With () and By () function in R is used for?

11) What are the data structures in R that is used to perform statistical analyses and create graphs?

12) Explain general format of Matrices in R?

13) In R how missing values are represented ?

14) Explain what is transpose?

15) Explain how data is aggregated in R?

16) What is the function used for adding datasets in R?

17) What is the use of subset() function and sample() function in R ?

18) Explain how you can create a table in R without external file?

### 1) What is R?

### 2) Differentiate between vector, List, Matrix, and Data frame.

### 3) Give names of those packages which are used for data imputation.

### 4) Explain initialize() function in R?

### 5) How can we find the mean of one column with respect to another?

### 6) What is a Random Walk model?

### 7) What is a White Noise model?

### 8) Give any five features of R.

### 9) Differentiate between R and Python in terms of functionality?

### 10) What are the applications of R?

### 11) Explain RStudio.

### 12) What are the advantages and disadvantages of R?

### 13) What is the purpose behind R and Hadoop integration?

### 14) Give the name of the Hadoop integration methods.

### 15) What will be the output of the expression all(NA==NA)?

### 16) What is the difference b/w sample() and subset() in R?

### 17) Why do we use the command - install.packages(file.choose(), repos=NULL)?

### 18) Give the command to create a histogram and to remove a vector from the R workspace?

### 19) Differentiate b/w "%%" and "%/%".

### 20) Why do we use apply() function in R?

### 21) Differentiate between library() and require() functions.

### 22) What is the t-test() in R?

### 23) What is the use of with() and by() functions in R?

### 24) Differentiate b/w lapply and sapply.

### 25) Explain aggregate() function.

### 26) Explain the doBy package?

### 27) Explain the use of the table() function.

### 28) Explain fitdistr() function?

### 29) What are GGobi and iPlots?

### 30) Explain the lattice package.

### 31) Explain anova() function.

### 32) Explain cv.lm() and stepAIC() function.

### 33) Explain leaps() function.

### 34) Explain relaimpo and robust package.

### 35) Give full form of MANOVA and what is the use of it.

### 36) Explain mashapiro.test() and barlett.test().

### 37) Explain the use of the forecast package.

### 38) Differentiate between qda() and lda() function.

### 39) Explain the auto.arima() and principal() function.

### 40) Explain FactoMineR.

### 41) What is the full form of SEM and CFA?

### 42) Define cluster.stats() and pvclust() function().

### 43) Define MATLAB and party packages.

### 44) Explain S3 and S4 systems.

### 45) Give names of visualization packages.

### 46) Explain Chi-Square Test

### 47) Explain Random Forest.

### 48) Explain Time Series Analysis.

### 49) Explain Pie chart in R.

### 50) Explain Histogram.

How do you handle errors in R?

UNIT V

## 1. What is K means Clustering Algorithm?

## 2. What is Lloyd’s algorithm for Clustering?

## 3. Is Feature Scaling required for the K means Algorithm

## 4. Why do you prefer Euclidean distance over Manhattan distance in the K means Algorithm?

## 5. Why is the plot of the within-cluster sum of squares error (inertia) vs K in K means clustering

## 6. Which metrics can you use to find the accuracy of the K means Algorithm?

## 7. What is a centroid point in K means Clustering?

## 8. Does centroid initialization affect K means Algorithm?

## 9. Discuss the optimization function for the K means Algorithm.

## 10. What are the advantages and disadvantages of the K means Algorithm?

## 11. What are the challenges associated with K means Clustering?

## 12. What are the ways to avoid the problem of initialization sensitivity in the K means Algorithm?

## 13. What is the difference between K means and K means++ Clustering?

## 14. How K means++ clustering Algorithm works?

## 15. How to decide the optimal number of K in the K means Algorithm?

## 16. What is the training and testing complexity of the K means Algorithm?

## 17. Is it possible that the assignment of data points to clusters does not change between

## 18. Explain some cases where K means clustering fails to give good results.

## 19. How to perform K means on larger datasets to make it faster?

## 20. What are the possible stopping conditions in the K means Algorithm?

## 21. What is the effect of the number of variables on the K means Algorithm?

## What is kNN Algorithm?

## How to select appropriate k value?

## 1. What is the KNN Algorithm?

## 2. Why is KNN a non-parametric Algorithm?

## 3. What is “K” in the KNN Algorithm?

## 4. Why is the odd value of “K” preferred over even values in the KNN Algorithm?

## 5. How does the KNN algorithm make the predictions on the unseen dataset?

## 6. Is Feature Scaling required for the KNN Algorithm? Explain with proper justification.

## 7. What is space and time complexity of the KNN Algorithm?

## 8. Can the KNN algorithm be used for regression problem statements?

## 9. Why is the KNN Algorithm known as Lazy Learner?

## 10. Why is it recommended not to use the KNN Algorithm for large datasets?

## 11. How to handle categorical variables in the KNN Algorithm?

## 12. How to choose the optimal value of K in the KNN Algorithm?

## 13. How can you relate KNN Algorithm to the Bias-Variance tradeoff?

## 14. Which algorithm can be used for value imputation in both categorical and continuous

## 15. Explain the statement- “The KNN algorithm does more computation on test time rather than train time”.

## 16. What are the things which should be kept in our mind while choosing the value of k in the

## 17. What are the advantages of the KNN Algorithm?

## 18. What are the disadvantages of the KNN Algorithm?

## 19. Is it possible to use the KNN algorithm for Image processing?

## 20. What are the real-life applications of KNN Algorithms?

## What is Logistic Regression ?

## What are the types of Logistic Regression techniques ?

## How does Logistic Regression work?