Linear Algebra

Bill Chung

January 12, 2023

```
library(far)
library(MASS)
library(pracma)
library(expm)
```

Welcome

- Goal:
  - Be able to explain the difference between ϵ and $\hat{\epsilon}$
  - Why $\hat{\epsilon}$ will orthogonal to all the features in your MLR

Snap shot from wk9

$$
\begin{aligned}
V[\hat{\vec{\beta}}] &= E[(\hat{\vec{\beta}} - \vec{\beta})^2] \\
&= E[(\hat{\vec{\beta}} - \vec{\beta})(\hat{\vec{\beta}} - \vec{\beta})^T]
\end{aligned}
$$

- now look at the expression for $\hat{\vec{\beta}} = (\mathbb{X}^T \cdot \mathbb{X})^{-1}\mathbb{X}^T(\vec{y})$, sub this expression back to $\hat{\vec{\beta}}$, then

$$
\begin{aligned}
V[\hat{\vec{\beta}}] &= E[(\hat{\vec{\beta}} - \vec{\beta})(\hat{\vec{\beta}} - \vec{\beta})^T] \\
&= E[\big((\mathbb{X}^T \cdot \mathbb{X})^{-1}\mathbb{X}^T\vec{\epsilon}\big)\big((\mathbb{X}^T \cdot \mathbb{X})^{-1}\mathbb{X}^T\vec{\epsilon}\big)^T] \\
&\text{gram matrix is symmetric matrix} \\
&= E[\big((\mathbb{X}^T \cdot \mathbb{X})^{-1}\mathbb{X}^T\vec{\epsilon}\big)\big(\vec{\epsilon}^T\mathbb{X} \cdot (\mathbb{X}^T\mathbb{X})^{-1}\big)] \\
&= E[\big((\mathbb{X}^T \cdot \mathbb{X})^{-1}] \cdot E[\mathbb{X}^T\vec{\epsilon}\vec{\epsilon}^T\mathbb{X}] \cdot E[(\mathbb{X}^T\mathbb{X}^{-1})] \\
&\text{the gram matrix becomes variance and covariance matrix when centered} \\
&= \frac{E[\mathbb{X}^T\vec{\epsilon}\vec{\epsilon}^T\mathbb{X}]}{(E[\mathbb{X}^T \cdot \mathbb{X}])^2}
\end{aligned}
$$


STOP SHOWING
UP EVERYWHERE I GO!!
LINEAR ALGEBRA

Recommended books

- `Linear Algebra and its application` by David C.Lay 4th edition
- Linear and Nonlinear Programming by Stephen G. Nash and Ariela Sofer
- The fundamental theorem of linear algebra, Strang, Gilbert
- `I wonder by Sam: Linear Algebra for Data Scientist` (soon to be published!)

Dancing with Wu Li Masters

- `Young man, in mathmatics, you don't understand things. You just get used to them` by John Von Neumann from *Dancing with Wu Li Masters*

Who is John Von Neumann?

- Leonoid Kantorovich (1912 - 1986): `A new method of solving some classes of extrmal problems (1937)`
- George B. Dantzig (1914 - 2005) : `SIMPLEX (1947)`
- Jerzy Neyman (1894 - 1982) : `Confidence Interval, P-value`
- John Von Neumann : `The duality theorem (1947)`

Schedule

- I will try to cover up to wk 7 material

Week	Topic	Key concepts
1	Attributes and method of `vector` and `matrix`	see notes below
2	Slight detour to probabilities: Joint, conditional, marginal and Bayes formula. Markov chain, eigenvalue, eigenvectors	Linear combinations
3	What is rref(A) and what does it tell you about your matrix?	Basis, subspace, space, span, projection, inverse
4	Fundamental four subspaces of matrix. Given a vector, can you find out where it `lives`?	Shall we span?
5	Projection, projection, projection	linear combination, change of basis
6	Findings vector multiplication that looks like projection	projection, orthogonal matrix, spanning Space
7	Change of basis and solving systems of equations	matrix decomposition
8	It does not matter how slowly you move as long as you are making progress	eignevalue, eigenvector, Markov chain
9	Eignedecomposition	eigenvalue, eigenvector, eigenspace, nullspace

Week	Topic	Key concepts
10	Markov chain	irreducible, reduccible, ergodic, regular, absorbing MC. What type of matrix do you have?
11	Meeting matrix again	PSD, PD, ID, NSD, ND, Condition number, symmetric matrix, gram matrix, diagonailzable matrix
12	Singular value decomposition	SVD and PCA
13	SIMPLEX method and	*The Martians*
14	The duality theorem	and .. Basis can a function! (What? really? FFT, EEMD)

Background

Notation

$$\mathbb{A} \cdot \vec{x} = \vec{b}$$

$$\vec{v}$$

$$\mathbb{A}$$

vectors

Attribute

- Size of a vector
- Direction that it can `move`
- Direction that it can `see`
- Norm
- Subspace where it `lives`
- Space where it `lives`

Method

- Span
- linear combination
- transpose
- dot product
- projection

Space

- Contains ∞ number of subspaces

Subspace

- Created by spanning a vector or set of vectors
- Always contains $\vec{0}$ and closed under `addition` and `multiplication`
- basis
- Has orthogonal complement subspace (`they are like best friends`)

matrix

Attribute

- Dimension of matrix
- Column Space, $C(\mathbb{A})$, Left Nullspace, $N(\mathbb{A}^T)$
- Row Space, $R(\mathbb{A})$, Nullspace, $N(\mathbb{A})$
- Input space (related to domain)
- Output space (related to codomain and Range)
- basis
- eigenvalue, eignevector
- singular value, singular vector
- condition number
- Rank
- PD, PSD, ID, ND, NSD
- Rank-nullity theorem
- inverse (not every square matrix has it..)
- Gram matrix

method

- transpose
- inverse
- decomposition
  - singular value decomposition
  - eigen decomposition
- projection
- rref($\mathbb{A}$)

Solving systems of equations

- Homogeneous equations
- Homogeneous equations
- Augmented matrix

How to create matrix and vector in R

```r
a1 <- matrix(c(3,0,-1,-5,2,4),nrow=1,byrow=T)
print(a1)
```

```
##      [,1] [,2] [,3] [,4] [,5] [,6]
## [1,]    3    0   -1   -5    2    4
```

```r
a2 <- matrix(c(3,0,-1,5,2,4),nrow=1,byrow=T)

A <- rbind(a1,a2)
print(A)
```

```
##      [,1] [,2] [,3] [,4] [,5] [,6]
## [1,]    3    0   -1   -5    2    4
## [2,]    3    0   -1    5    2    4
```

```r
Rank(A)
```

```
## [1] 2
```

```r
dim(A)
```

```
## [1] 2 6
```

```r
A
```

```
##      [,1] [,2] [,3] [,4] [,5] [,6]
## [1,]    3    0   -1   -5    2    4
## [2,]    3    0   -1    5    2    4
```

```r
x <- c(1,2,3,4,5,6)
x
```

```
## [1] 1 2 3 4 5 6
```

```r
b<- x/Norm(x)

Norm(b)
```

```
## [1] 1
```

```r
A%*%x
```

```
##      [,1]
## [1,]   14
## [2,]   54
```

select columns 1, 3 and 6 and put them into $\mathbb{B}$
select columns 2, 4 and 5 and put them into $\mathbb{N}$

```r
B <- A[,c(1,3,6)]
B
```

```
##      [,1] [,2] [,3]
## [1,]    3   -1    4
## [2,]    3   -1    4
```

$$\mathbb{B} \cdot \vec{x}_B + \mathbb{N} \cdot \vec{x}_N = \mathbb{A} \cdot \vec{x}$$

Creating sample vector

```r
#randomly selects number
a <- sample(-5:5, replace=TRUE, 12)
#find out number of elements in the vector
length(a)
```

```
## [1] 12
```

```r
A <- matrix(a, ncol = 4, byrow= TRUE)
A
```

```
##      [,1] [,2] [,3] [,4]
## [1,]    3    2   -5    4
## [2,]    3   -2   -2   -4
## [3,]   -2   -1   -1   -1
```

```r
A <- matrix(sample(-5:5, replace=TRUE, 12), ncol = 4, byrow= TRUE)
A
```

```
##      [,1] [,2] [,3] [,4]
## [1,]    5   -3   -2    5
## [2,]    3   -2   -5    2
## [3,]    5   -1    0    5
```

```r
b <- matrix(sample(-5:5, replace=TRUE, 3), ncol = 1, byrow= TRUE)

H <- cbind(A,b)
rref(H)
```

```
##      [,1] [,2] [,3]       [,4]  [,5]
## [1,]    1    0    0  0.9444444 -1.75
## [2,]    0    1    0 -0.2777778 -4.75
## [3,]    0    0    1  0.2777778  1.25
```

C and P

$$
{}_nP_r = \frac{n!}{(n-r)!}
$$

$$
{}_nC_r = \frac{n!}{r!(n-r)!}
$$

Another way to express combination

$$
\binom{N}{r} = \frac{n!}{(n-r)!r!}
$$

Recommended Chapters and reading from David Lay

CH1

- CH1.1 example 1, 2,3
- 1.1 Exercise 11,12,13,14
- page 27, linear combinations
- page 30, definition
- CH1.3(page32) 13-6
- page 35, definition, page 36, theorem
- page 39, Theorem 5
- CH1.4 Exercise 5,6,7,8,11,12
- CH1.5 example 1, 2 (this is related to nullspace) example 3 (this is an example of hyperplane)
- CH1.5 Exercise 1,2,3,4
- Ch1.7 page 56 (definition). example 1, page 57 (the yellow box) example 2, 3,5
- page 64, example 1, page 65 (definition)
- page 93, theorem 1, page 95, definition, page 95 example 3, example 4
- page 98, theorem 2 page 99 theorem 3
- page 103, inverse of matrix definition, page 105, theorem 6
- page 112, theorem 8 (when A is invertible..then we know the following)
- page 114, numerical notes

CH2.8

- Subspaces of Rn
- page 146, 147, 148 definitions and theorem 12
- example 6 (important), example 7,
- CH2.9, page 154 and 155, definition,example 3 (page 155)
- page 156 Theorem 14 and 15 (and more about rank and invertible matrix)
- CH2.9 Exercise 9 to 12

CH3, determinant (skip)

CH4

- p190 Definition of vector space, p193, subspaces
- CH4.2, page 198 definition, example 1, theorem 2,
- page 200, example 3, page 201 column space
- page 203 and 204, Kernel and Range of a Linear transformation (table and definition)
- CH4.2 exercise 37,38,39

- CH 4.3 page 208, theorem 4, page 209 definition
- CH 4.3 Exercise 13,14
- CH 4.4. Coordinate systems (theorem 7 and definition), page 218 example 4
- CH 4.5 theorem 9 , exercise 13,14,15,16,17 and 18
- CH 4.6 Rank page 231, theorem 13, example 2, page 233, def and theorem 14 (super important)
- page 235 more about information that a nonsigular matrix provides
- CH 4.6 Exercise 1 to 6. 19 and 20
- page 240, change of basis theorem 15 (will talk about if we get to talk about eigendecomposition)
- page 253. Section 4.9 example 1-5 (again will talk about this if we talk about eigenvalue analysis)
- page 265, okay this is chapter about eigenvalues and eigenvectors.

CH 5.1 to 5.6

CH6 6.1 to 6.6 (this is OLS!)