Linear algebra

Readings

Readings

- Appendix D

Introduction

Goal

Use matrices to derive and express estimators and their properties

(in linear-ish models)

Matrices

An $m \times n$ **matrix** is an array of reals with m rows and n columns,

$$\mathbf{A} = [a_{ij}] = \begin{bmatrix} a_{11} & a_{12} & & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & & a_{mn} \end{bmatrix}$$

A matrix is **square** if $m = n$.

Matrices (R)

```r
A <- matrix(c(1,2,3,4),ncol=2,nrow=2)
A
```

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

Column vectors

A **column vector** of length n is an $n \times 1$ matrix, written

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

Column vectors (R)

```r
a <- A[,1]
a
```

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

R prints all objects with one dimension as rows.

Row vectors

A **row vector** of length n is an $1 \times n$ matrix, written

$$\mathbf{x} = (x_1, \cdots, x_n)$$

Row vectors (R)

```r
b <- A[1,]
b
```

```
## [1] 1 3
```

Diagonal matrix

A square matrix is **diagonal** if $a_{ij} = 0$ for all $i \neq j$.

Diagonal matrix (R)

```r
diag(10*rnorm(4))
```

```
##          [,1]      [,2]      [,3]      [,4]
## [1,] 6.733394  0.000000 0.0000000   0.00000
## [2,] 0.000000 -1.701653 0.0000000   0.00000
## [3,] 0.000000  0.000000 0.3999658   0.00000
## [4,] 0.000000  0.000000 0.0000000 -24.19603
```

Identity matrix

The $n-$ dimensional diagonal matrix with $a_{ii} = 1$ for all i is the **identity matrix**, called I_n:

$$I_n = \begin{bmatrix} 1 & 0 & & 0 \\ 0 & 1 & & 0 \\ \vdots & & & \\ 0 & 0 & & 1 \end{bmatrix}$$

Identity matrix (R)

```r
n <- 5
iota_n <- numeric(5)+1
I_n <- diag(iota_n)
iota_n
```

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

```r
I_n
```

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

Zero matrix

The matrix $O_{m \times n}$ is an $m \times n$ matrix with zeros everywhere

Transposition

- **Transpose**: $\mathbf{A}'$ turns the rows into columns, columns into rows.
- Let $A = [a_{ij}]$, then $B = A' = [b_{ij}]$ where $b_{ij} = a_{ji}$ for all i and j
- $(AB)' = B'A'$

Special vectors

Important vectors:

$$\iota' = (1, 1, \cdots, 1)$$
$$e_2' = (0, 1, 0, \cdots, 0)$$
$$e_3' = (0, 0, 1, 0, \cdots, 0)$$

Symmetric matrix

A matrix A is **symmetric** if $A' = A$

Equation systems

Vectors and matrices are useful for solving linear systems of equations. The system

$$a_{11}x_1 + a_{12}x_2 = b_1$$
$$a_{21}x_1 + a_{22}x_2 = b_2$$

can be written as

$$\mathbf{Ax} = \mathbf{b}$$

where $\mathbf{A}$ is a 2×2 matrix, $\mathbf{x}' = (x_1, x_2)$ and $\mathbf{b}' = (b_1, b_2)$

Beer and coffee

You find two receipts from two visits to a bar.

- A coffee and three beers: CAD 10
- Three coffees and a beer: CAD 14

Solution?

Beer and coffee (R)

```
A <-  matrix(c(1,3,3,1),nrow=2)
b <- matrix(c(10,14),nrow=2)
cbind(A,b)
```

```
##      [,1] [,2] [,3]
## [1,]    1    3   10
## [2,]    3    1   14
```

```
x <- solve(A) %*% b
x
```

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

Systems of equations

For a given system of equations $Ax = b$:

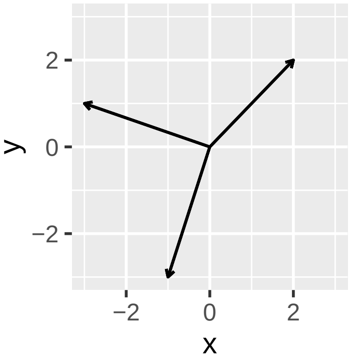
- How many solutions x exist?
  - 0
  - 1
  - many
- If there is no exact solution, can we approximate x so that $Ax \approx b$?
- How do we obtain exact or approximate solutions?

Vector operations

Visualizing

Vectors as the path starting from 0:

```
library(ggplot2)
df <- data.frame(x=c(0,2,4),y=c(1,2,-2))
ggplot(df,aes(x,y)) + geom_segment(aes(x = 0, y = 0, xend =
```



Vector addition

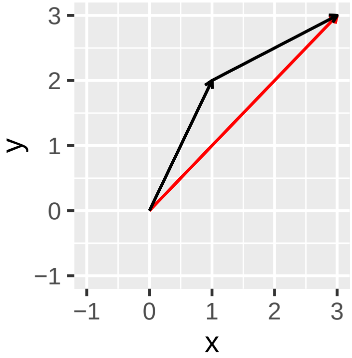
If x and y are two vectors of length n, then $x + y$ is defined as:

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}$$

Vectors must be **conformable**

Visualizing

```
library(ggplot2)
df <- data.frame(x=c(0,2,4),y=c(1,2,-2))
ggplot(df,aes(x,y)) + geom_segment(aes(x = 0, y = 0, xend =
```



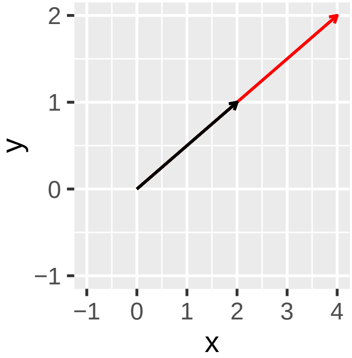
Scalar multiplication

Let $c \in \mathbb{R}$ and let x be an n vector, i.e. $x \in \mathbb{R}^n$. Then

$$cx = \begin{bmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{bmatrix}$$

Visualizing

```
library(ggplot2)
df <- data.frame(x=c(0,2,4),y=c(1,2,-2))
ggplot(df,aes(x,y)) + geom_segment(aes(x = 0, y = 0, xend =
```



Inner product

For two vectors x, y of length n, the **inner product** is defined as

$$x'y = \sum_{i=1}^{n} x_i y_i$$

Inner product

What is the inner product of $x = (1, 2, 3, 4, 5)$ with itself?

Inner product (R)

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

```
## [1] 55
```

Matrix operations

Matrix addition

For two $m \times n$ matrices A, B, $C = A + B$ is defined as the element-wise sum, $[c_{ij}] = [a_{ij} + b_{ij}]$

```
A <- matrix(c(1,2,3,4),ncol=2)
B <- matrix(c(5,6,7,8),ncol=2)
A
```

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

```
A+B
```

```
##      [,1] [,2]
## [1,]    6   10
## [2,]    8   12
```

Scalar multiplication

Scalar multiplies each element.

```
c <- 2
c*A
```

```
##      [,1] [,2]
## [1,]    2    6
## [2,]    4    8
```

```
c <- -3
c*A
```

```
##      [,1] [,2]
## [1,]   -3   -9
## [2,]   -6  -12
```

Matrix multiplication

Matrix Multiplication

To multiply matrix **A** by matrix **B** to form the product **AB**, the *column* dimension of **A** must equal the *row* dimension of **B**. Therefore, let **A** be an $m \times n$ matrix and let **B** be an $n \times p$ matrix. Then, **matrix multiplication** is defined as

$$\mathbf{AB} = \left[\sum_{k=1}^{n} a_{ik} b_{kj} \right].$$

In other words, the $(i,j)^{\text{th}}$ element of the new matrix **AB** is obtained by multiplying each element in the i^{th} row of **A** by the corresponding element in the j^{th} column of **B** and adding these n products together. A schematic may help make this process more transparent:

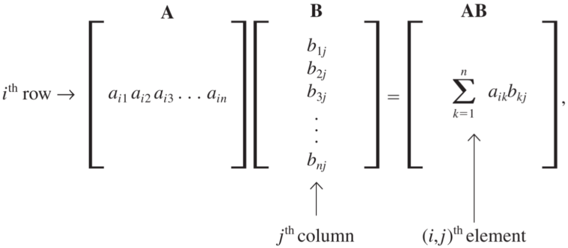


Figure 1: matrix multiplication

Matrix multiplication (R)

```
A
```

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

```
B
```

```
##      [,1] [,2]
## [1,]    5    7
## [2,]    6    8
```

```
C <- A%*%B
```

What is C?

Matrix multiplication (R)

```
C
```

```
##      [,1] [,2]
## [1,]   23   31
## [2,]   34   46
```

Matrix multiplication 2

```
A <- matrix(c(1,2,3,4,5,6),ncol=3)
B <- matrix(c(1,2,3,4),nrow=2)
A
```

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

```
B
```

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

What is AB?

Properties (14)

$AB \neq BA$

1. $A = \iota_5', \; B = A'$
2. $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$
   - What does A do in AB?
   - What does A do in BA?

Multiplication: I_n, $O_{m \times n}$

1. What happens when you multiply by I_n?
   - Proof: *whiteboard:sketch4.2*
2. ... when $O_{m \times n}$?

Exact solutions

Inversion

For a square matrix of dimension n, A, and conformable vectors x and b, consider the linear system:

$$Ax = b$$

If there exists a matrix A^{-1} such that

$$A^{-1}A = I_n$$

then

$$x = A^{-1}b$$

Inverse

A square matrix A of dimension n is **invertible** if there exists a matrix A^{-1} such that $AA^{-1} = A^{-1}A = I_n$

A^{-1} is the **inverse** of A.

- What is the inverse of 5?
- Does 0 have an inverse?

Inverse (properties)

1. If it exists, the inverse is unique
2. If $\alpha \neq 0$ and A is invertible, then $(\alpha A)^{-1} = \frac{1}{alpha} A^{-1}$
3. Assuming invertibility and conformability, $(AB)^{-1} = B^{-1}A^{-1}$
4. $(A')^{-1} = (A^{-1})'$

Inverse (practice)

```
A <- matrix(c(2,0,0,2),nrow=2)
A
```

```
##      [,1] [,2]
## [1,]    2    0
## [2,]    0    2
```

Inverse (practice)

```
A <- matrix(c(1,2,3,0),nrow=2)
A
```

```
##      [,1] [,2]
## [1,]    1    3
## [2,]    2    0
```

Inverse: existence

When do we know an inverse exist?

- Required: must be square
- Sufficient: full **rank**
- Sufficient: positive **definite**

Linear independence (1)

Consider the matrix

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

- Is it invertible?
- Why: *whiteboard:twodependentvecs*

Linear independence (2)

Take a set $(x_1, \cdots, x_r)$ of $n \times 1$ vectors.

They are **linearly independent** if and only if

$$a_1 x_1 + a_2 x_2 + \cdots + a_r x_r = 0$$

implies $a_1 = \cdots = a_r = 0$

- Can you construct one of the vectors as a linear combination of the other ones?

Linear independence (3)

Linear independence is closely related to finding a solution to $Ax = b$

Picture

- Can you use a linear combination of the column vectors of A to get to point b?
- That linear combination is x
- Given two linearly independent vectors in $\mathbb{R}^2$, you can go anywhere in $\mathbb{R}^2$
- If they are linearly dependent, you can only travel over the line $x_1 a_1$

Rank

The **rank** of a matrix is the maximum number of linearly independent columns.

Rank (practice)

```r
matrix(c(1,2,2,4,3,6),nrow=2)
```

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

Rank - inverse

If a square $n \times n$ matrix A has rank n, then it is invertible

Positive definiteness

For a symmetric, $n \times n$ matrix A, the quadratic form is

$$x'Ax = \sum_i a_{ii} x_i^2 + 2 \sum_{i=1}^{n} \sum_{j>i} a_{ij} x_i x_j$$

Then A is called

- **positive definite** (pd) if $x'Ax > 0$ whenever $x \neq 0$
- **positive semidefinite** (psd) if $x'Ax \geq 0$ for all x

PD and invertibility

1. If A is pd, A^{-1} exists and is pd. **Invertible**
2. For any $n \times k$ matrix X, $X'X$ and XX' are psd. **Maybe invertible.**
3. For any full rank matrix X, $X'X$ is pd. **Invertible.**

Projections

One solution

If A has full rank, then we can solve

$$Ax = b$$

for $x = A^{-1}b$

Many solutions

If A is rank-deficient, then

$$Ax = b$$

has many solutions, e.g.

$$x_1 + x_2 = 27$$

Approximate solutions

If the rank of A exceeds the dimension of x, then there are more equations than unknowns

$$\begin{aligned} x_1 &= 2 \\ x_1 + x_2 &= 4 \\ x_2 &= 3 \end{aligned}$$

Minimization

If $Ax = b$ has no solution, we can still

pick x to minimize the distance from Ax to b

Let *the distance* be defined as

$$SS(x) = (Ax - b)'(Ax - b)$$

the inner product. Then the approximate solution is:

$$\hat{x} = argmin SS(x)$$

Minimization: picture

whiteboard:sketch4-4

Minimization of quadratic form

The inner product is a quadratic form, and can be written out as

$$SS(b) = x'A'Ax - b'Ax - x'A'b + b'b$$

Calculus for matrices tells us that the differential is

$$dSS(b) = (dx)'A'Ax + x'A'Adx - b'Adx - (dx)'A'b$$

Because all quantities are scalars, that is equal to

$$dSS(b) = 2x'A'Adx - 2b'Adx$$

FOC

The derivative is

$$2x'A'A - 2b'A$$

so transposition yields the FOC

$$2A'A\hat{x} = 2A'b$$

The 2's and the A's drop out, and . . . ?

Projection matrix

Since A has full rank (no solution), $A'A$ is pd, and therefore invertible. Therefore,

$$\hat{x} = (A'A)^{-1}A'b$$

1. $P_A = (A'A)^{-1}A'$ is the **projection matrix** for matrix A
2. $x^* = P_A b$ is the **orthogonal projection** of b onto A
3. $b - P_A b = (I - P_A)b = M_A b$ is called the residual

Idempotent matrices

A square, $n-$dimensional matrix A is idempotent if $AA = A$.

- Can you think of two idempotent numbers?
- Can you think of two idempotent matrices?

Projection matrix

1. P_A is idempotent
2. M_A is idempotent