CSMMA16

Mathematics and Statistics

Vectors and Matrices

26th September 2016

Matrices

Matrix = Rectangular array of numbers, complex numbers or functions
For example,

$$\mathbf{A} = [a_{ij}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$

is a matrix of n rows and m columns with element a_{ij} in the *ith* row and *jth* column.
A is said to be of order or dimension $n \times m$
- rows are listed first and columns second, by convention.
A is square if $m = n$. Thus, for a square matrix the elements $\{a_{11}, a_{22}, \cdots, a_{nn}\}$ form the main diagonal.

Vectors

A vector has one row or one column

$$\mathbf{R} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix}$$

R is a row vector
C is a column vector
Note that in R software:

- Arrays are numeric objects with dimension attributes
- Matrices are two dimensional arrays
- Vectors are not matrices - have no (NULL) dimension attributes

Addition

Matrix addition (subtraction) defined only for matrices of the same order.
The sum (difference) also has the same order
Consider two matrices **A** and **B** of the same order and let

$$\mathbf{S} = [s_{ij}] = \mathbf{A} + \mathbf{B}$$

Then

$$s_{ij} = a_{ij} + b_{ij}$$

Result **S** also has same order as **A** and **B**
In R:

```
A<-matrix(c(1,2,3,4,5,6), nrow=2,byrow=T) # defines A.
B<-matrix(1,nrow=2,ncol=3)                # same for B
S<-A+B                            #evaluate sum. Assign to S
S                       #output S
```

Note: function c( ) forms a vector from its arguments

Equal Matrices

$\mathbf{A} = \mathbf{B}$ if $a_{ij} = b_{ij}$ for all i, j
That is, they are of the same size and corresponding elements are identical.

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}; \qquad \mathbf{B} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} 1 & 1 \\ 3 & 4 \end{bmatrix}; \qquad \mathbf{D} = \begin{bmatrix} 1 & 2 & 0 \\ 3 & 4 & 0 \end{bmatrix}$$

$$\mathbf{A} = \mathbf{B}, \qquad \mathbf{A} \neq \mathbf{C}, \qquad \mathbf{A} \neq \mathbf{D}$$

In R:

```
identical(A,B)
[1] TRUE
identical(A,C)
[1] FALSE
```

Note: other options available

Multiplication

By scalar: To form the product $\mathbf{R} = k\mathbf{A}$, multiply each element in $\mathbf{A}$ by the scalar k. Thus,

$$r_{ij} = k \times a_{ij} \ .$$

Note that $k \times (\mathbf{A} + \mathbf{B}) = k * \mathbf{A} + k * \mathbf{B}$ (distributive)

By matrix: The product $\mathbf{R} = \mathbf{AB}$ defined only if number of columns in $\mathbf{A}$ equals number of rows in $\mathbf{B}$

In resultant matrix:

$$r_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \ .$$

Example:

$$\begin{bmatrix} 2 & 1 & 3 \\ 1 & 4 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 3 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 \times 2 + 1 \times 0 + 3 \times 1 & 2 \times 1 + 1 \times 3 + 3 \times 2 \\ 1 \times 2 + 4 \times 0 + 1 \times 1 & 1 \times 1 + 4 \times 3 + 1 \times 2 \end{bmatrix}$$

$$= \begin{bmatrix} 7 & 11 \\ 3 & 15 \end{bmatrix}$$

Note that $\mathbf{A}$ is 2×3 and $\mathbf{B}$ is 3×2 so the product is 2×2

Some rules

- The product $\mathbf{R}$ has number of rows as $\mathbf{A}$ and number of columns as $\mathbf{B}$
  -If $\mathbf{A}$ is $m \times n$, $\mathbf{B}$ is $n \times p$ then $\mathbf{A} \times \mathbf{B}$ is $m \times p$ matrix.
- If $\mathbf{A} \times \mathbf{B}$ is possible then $\mathbf{B} \times \mathbf{A}$ may not be.
- If both are possible, $\mathbf{A} \times \mathbf{B} \neq \mathbf{B} \times \mathbf{A}$ in general.
- $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \times \mathbf{B}) \times \mathbf{C} = \mathbf{A} \times \mathbf{B} \times \mathbf{C}$
- $\mathbf{A} \times (\mathbf{B} + \mathbf{C}) = (\mathbf{A} \times \mathbf{B}) + (\mathbf{A} \times \mathbf{C})$
- For scalar k,
  $(k \times \mathbf{A}) \times \mathbf{B} = k \times (\mathbf{A} \times \mathbf{B}) = \mathbf{A} \times (k \times \mathbf{B}) = (\mathbf{A} \times \mathbf{B}) \times k$

Transpose

Here the matrix is flipped: rows become columns and vive-versa
If $\mathbf{R} = \mathbf{A}^T$, then $r_{ij} = a_{ji}$.

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} \quad \text{then} \quad \mathbf{A}^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}$$

If $\mathbf{A}$ is size $m \times n$, then $\mathbf{A}^T$ is size $n \times m$.

In R:

```
A.T<-t(A)
```


Some rules

- $(\mathbf{A}^T)^T = \mathbf{A}$
- $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$
- $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$
- $(k\mathbf{A})^T = k\mathbf{A}^T$
- If $\mathbf{A}^T = \mathbf{A}$ then $\mathbf{A}$ is symmetric
- If $\mathbf{B}^T = -\mathbf{B}$ then $\mathbf{B}$ is skew-symmetric

Special Matrices

$$\begin{bmatrix}2&0&0\\0&3&0\\0&0&4\end{bmatrix} \quad \begin{bmatrix}1&3&2\\0&3&4\\0&0&4\end{bmatrix} \quad \begin{bmatrix}2&0&0\\3&1&0\\4&6&0\end{bmatrix}$$

Diagonal *UpperTriangular* *LowerTriangular*

$$\begin{bmatrix}2&0&0\\0&2&0\\0&0&2\end{bmatrix} \quad \begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix} \quad \begin{bmatrix}0&0&0\\0&0&0\\0&0&0\end{bmatrix}$$

Scalar *Identity* **I** *Zero*

- If **S** is scalar, $\mathbf{SA} = \mathbf{AS}$. In particular $\mathbf{IA} = \mathbf{AI} = \mathbf{A}$
- To convert a scalar k to a matrix, multiply scalar by **I**
- $\mathbf{I}_n$ denotes $n \times n$ identity matrix

In R:

```
diag(x = c(2,3,4))
diag(x=1,nrow=3,ncol=3) #for Identity
lower.tri(A,diag=F) #matrix of logicals with TRUE in lower triangle.
A[lower.tri(A,diag=F)]<-0 #upper triangular matrix from A
```


Special Matrices and Equations

Consider

$$\begin{bmatrix}2&0&0\\0&3&0\\0&0&4\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix} = \begin{bmatrix}8\\27\\20\end{bmatrix}$$

This is easy to evaluate. Clearly $x = 4; \quad y = 9; \quad z = 5$. Another example:

$$\begin{bmatrix}1&3&2\\0&2&4\\0&0&6\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix} = \begin{bmatrix}22\\18\\12\end{bmatrix}$$

This is equivalent to

$$\begin{aligned} x + 3y + 2z &= 22 \\ 2y + 4z &= 18 \\ 6z &= 12 \end{aligned}$$

This is also quite easy to evaluate: $x = 3; \quad y = 5; \quad z = 2$.

Determinant and Cofactor

The determinant of a square matrix $\mathbf{A}$ of numbers is a number associated with that matrix. It is written $\det(\mathbf{A})$ or $|\mathbf{A}|$.

The cofactor $\mathbf{C}_{ij}$ of the $(i,j)^{th}$ element a_{ij} of the matrix $\mathbf{A}$ is $-1^{i+j}\mathbf{M}_{ij}$ where $\mathbf{M}_{ij}$ is determinant of matrix $\mathbf{A}$ without row i and column j

Example:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}; \quad \mathbf{M}_{1,3} = \begin{vmatrix} 4 & 5 & \\ & & \\ 7 & 8 & \end{vmatrix}; \quad \mathbf{M}_{2,1} = \begin{vmatrix} & 2 & 3 \\ & & \\ & 8 & 9 \end{vmatrix}; \quad \mathbf{C}_{2,1} = -\begin{vmatrix} 2 & 3 \\ 8 & 9 \end{vmatrix}$$

The determinant of a single element matrix is that element.

To find determinant of an $n \times n$ matrix pick any row or column in the matrix, multiply every element in the chosen row or column by its cofactor and sum. For example, choosing the first row,

$$|\mathbf{A}| = \sum_{j=1}^{n} a_{1,j}\mathbf{C}_{1,j}$$

For $\mathbf{A}$ above, $|\mathbf{A}| = 1 \times \mathbf{C}_{1,1} + 2 \times \mathbf{C}_{1,2} + 3 \times \mathbf{C}_{1,3}$

Determinants

For 2×2 matrices:

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}; \quad \mathbf{C}_{1,1} = |a_{2,2}|; \quad \mathbf{C}_{1,2} = -|a_{2,1}|$$

So $|\mathbf{A}| = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}$

For 3×3 matrices:

$$\begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix} = a_{1,1}\begin{vmatrix} a_{2,2} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{vmatrix} - a_{1,2}\begin{vmatrix} a_{2,1} & a_{2,3} \\ a_{3,1} & a_{3,3} \end{vmatrix} + a_{1,3}\begin{vmatrix} a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} \end{vmatrix}$$

$$= a_{1,1}(a_{2,2}a_{3,3} - a_{2,3}a_{3,2}) - a_{1,2}(a_{2,1}a_{3,3} - a_{2,3}a_{3,1}) + a_{1,3}(a_{2,1}a_{3,2} - a_{2,2}a_{3,1})$$

Determinant Examples

A 2×2 matrix:

$$\mathbf{A} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}; \quad \mathbf{C}_{1,1} = |4|; \quad \mathbf{C}_{1,2} = -|3|$$

So $|\mathbf{A}| = 1 \times 4 - 2 \times 3 = -2$

A 3×3 matrix:

$$\begin{vmatrix} 3 & 1 & 5 \\ 4 & 1 & 2 \\ 3 & 1 & 4 \end{vmatrix} = 3 \begin{vmatrix} 1 & 2 \\ 1 & 4 \end{vmatrix} - \begin{vmatrix} 4 & 2 \\ 3 & 4 \end{vmatrix} + 5 \begin{vmatrix} 4 & 1 \\ 3 & 1 \end{vmatrix} = 6 - 10 + 5 = 1$$

In R:

```
A<-matrix(c(1,2,3,4), nrow=2)
det(A)
[1] -2
unlist(determinant(A, logarithm=F))
modulus    sign
      2      -1
```


Uses of Determinant

- Find inverses
- The Jacobian of the transformation in change of variable in an integral is a determinant - The Jacobian keeps track of stretching and warping of the function
- Eigenvalues (later) of a matrix are roots of a determinant
- Wronskian of solutions of a linear ODE (next lecture) is a determinant.

Inverse

The (multiplicative) inverse of $\mathbf{A}$, denoted $\mathbf{A}^{-1}$, satisfies

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$$

One way of finding inverse of square matrix is

$$\mathbf{A}^{-1} = \frac{adj(\mathbf{A})}{det(\mathbf{A})}$$

where $adj(\mathbf{A})$ is the adjoint of $\mathbf{A}$ and is the transpose of the matrix of cofactors of $\mathbf{A}$. That is

$$adj(\mathbf{A}) = [C_{ji}] = [C_{ij}]^T$$

Note, a matrix has no inverse if its determinant is 0.

Inverse Example

For $A = \begin{pmatrix} 3 & 1 & 5 \\ 4 & 1 & 2 \\ 3 & 1 & 4 \end{pmatrix}, \quad [C_{ij}] = \begin{pmatrix} 2 & -10 & 1 \\ 1 & -3 & 0 \\ -3 & 14 & -1 \end{pmatrix}$

So $adj(\mathbf{A}) = \begin{pmatrix} 2 & 1 & -3 \\ -10 & -3 & 14 \\ 1 & 0 & -1 \end{pmatrix}$ and $\mathbf{A}^{-1} = \frac{adj(\mathbf{A})}{det(\mathbf{A})} = \begin{pmatrix} 2 & 1 & -3 \\ -10 & -3 & 14 \\ 1 & 0 & -1 \end{pmatrix}$

In R:

```
A<-matrix(c(3,1,5,4,1,2,3,1,4),nrow=3,byrow=T)
solve(A)
     [,1] [,2] [,3]
[1,]    2    1   -3
[2,]  -10   -3   14
[3,]    1    0   -1
```

Some properties

- $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$
- $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-\mathbf{1}})^T$
- $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}{-}1$ note change of order
- If $\mathbf{A}^T = \mathbf{A}^{-1}$ then matrix $\mathbf{A}$ is an orthogonal matrix.

Diagonal matrices (with non-zero elements) are orthogonal

Rank of Matrix

Rank of matrix is largest square sub-matrix whose determinant $\neq 0$.
A sub-matrix of $\mathbf{A}$ is $\mathbf{A}$ less some rows or columns
For example:

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 3 & 4 \\ 1 & 1 & 2 & 0 \\ 2 & 4 & 6 & 8 \end{bmatrix}$$

has four 3×3 sub-matrices all whose det $= 0$
But, for instance, sub-matrix

$$\begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} \neq 0$$

so rank($\mathbf{A}$)=2
Properties of Rank:

- Rank($\mathbf{A}$)=0 only if $\mathbf{A}$ is the zero matrix.
- Rank($\mathbf{A}$) = Rank($\mathbf{A}$)^T
- Elementary row operations (later) don't affect matrix rank.

Rank is a quite useful concept in linear models and control theory.

Other Functions

Let $\mathbf{A}$ be a square $(n \times n)$ matrix

- $\mathbf{A}^p = \mathbf{AAA} \cdots \mathbf{A} \quad (p\ \mathbf{A}'s)$
- $\exp(\mathbf{A}) = \mathbf{I}_n + \mathbf{A} + \frac{1}{2!}\mathbf{A}^2 + \frac{1}{3!}\mathbf{A}^3 + \cdots$
- Trace of $\mathbf{A}$, $tr(\mathbf{A}) = \sum_{i=1}^{n} a_{i,i}$ - sum of diagonal elements

Gaussian Elimination

Technique used to solve systems of equations.
Consider linear equations with the general form $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{x}$ and $\mathbf{b}$ are vectors.
Example:

$$\begin{bmatrix} 3 & 4 & 0 \\ 2 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 10 \\ 7 \\ 7 \end{bmatrix}$$

First form augmented matrix $\tilde{\mathbf{A}} = \begin{bmatrix} \mathbf{A} & \mathbf{b} \end{bmatrix}$:

$$\tilde{\mathbf{A}} = \begin{bmatrix} 3 & 4 & 0 & 10 \\ 2 & 0 & 1 & 7 \\ 0 & 1 & 2 & 7 \end{bmatrix}$$

Important: Each row is equivalent to one equation.
Do "row operations" on each row, row X $:= a \times$ row X $+ b \times$ row Y
Aim: make $\mathbf{A}$ upper triangular matrix (so called echelon form)

Example

$$\tilde{\mathbf{A}} = \begin{bmatrix} 3 & 4 & 0 & 10 \\ 2 & 0 & 1 & 7 \\ 0 & 1 & 2 & 7 \end{bmatrix}$$

Aim to turn 2 in row 2 to 0. Row 2 := 2 × row 1 − 3 × row 2

$$\tilde{\mathbf{A}} = \begin{bmatrix} 3 & 4 & 0 & 10 \\ 0 & 8 & -3 & -1 \\ 0 & 1 & 2 & 7 \end{bmatrix}$$

Now turn 1 in row 3 to 0. Row 3 := row 2 − 8 × row 3

$$\tilde{\mathbf{A}} = \begin{bmatrix} 3 & 4 & 0 & 10 \\ 0 & 8 & -3 & -1 \\ 0 & 0 & -19 & -57 \end{bmatrix} \Longrightarrow \begin{bmatrix} 3 & 4 & 0 \\ 0 & 8 & -3 \\ 0 & 0 & -19 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 10 \\ -1 \\ -57 \end{bmatrix}$$

Hence $x = 2; \quad y = 1; \quad z = 3$

Gauss Jordan - for Matrix Inverse

To invert $\mathbf{A}$ form matrix $[\mathbf{A} \quad \mathbf{I}]$
Do row operations until $\mathbf{A}$ part is unit matrix $\mathbf{I}$
Inverse matrix is in second part
Example:

$$\mathbf{A} = \begin{bmatrix} 3 & -1 & 1 \\ -1 & 3 & 4 \\ -1 & 1 & 2 \end{bmatrix} \quad \text{then} \quad \begin{bmatrix} \mathbf{A} & \mathbf{I} \end{bmatrix} = \begin{bmatrix} 3 & -1 & 1 & 1 & 0 & 0 \\ -1 & 3 & 4 & 0 & 1 & 0 \\ -1 & 1 & 2 & 0 & 0 & 1 \end{bmatrix}$$

$$\text{Make } a_{2,1} \text{ and } a_{3,1} \text{ zero} : \begin{bmatrix} \mathbf{A} & \mathbf{I} \end{bmatrix} = \begin{bmatrix} 3 & -1 & 1 & 1 & 0 & 0 \\ 0 & 8 & 13 & 1 & 3 & 0 \\ 0 & 2 & 7 & 1 & 0 & 3 \end{bmatrix}$$

$$\text{Make } a_{3,2} \text{ zero} : \begin{bmatrix} \mathbf{A} & \mathbf{I} \end{bmatrix} = \begin{bmatrix} 3 & -1 & 1 & 1 & 0 & 0 \\ 0 & 8 & 13 & 1 & 3 & 0 \\ 0 & 0 & -15 & -3 & 3 & -12 \end{bmatrix}$$

Continued

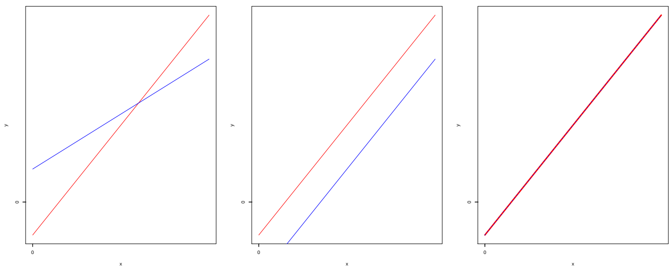
$$\text{Make } a_{1,3} \text{ and } a_{2,3} \text{ zero} : \begin{bmatrix} \mathbf{A} & \mathbf{I} \end{bmatrix} = \begin{bmatrix} 45 & -15 & 0 & 12 & 3 & -12 \\ 0 & 120 & 0 & -24 & 84 & -156 \\ 0 & 0 & -15 & -3 & 3 & -12 \end{bmatrix}$$

$$\text{Make } a_{1,2} \text{ zero} : \begin{bmatrix} \mathbf{A} & \mathbf{I} \end{bmatrix} = \begin{bmatrix} 360 & 0 & 0 & 72 & 108 & -252 \\ 0 & 120 & 0 & -24 & 84 & -156 \\ 0 & 0 & -15 & -3 & 3 & -12 \end{bmatrix}$$

Now set diagonals of the left half of the matrix to 1:
Divide Row 1 by 360, Row 2 by 120 and Row 3 by -15

$$\begin{bmatrix} \mathbf{A} & \mathbf{I} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0.2 & 0.3 & -0.7 \\ 0 & 1 & 0 & -0.2 & 0.7 & -1.3 \\ 0 & 0 & 1 & 0.2 & -0.2 & 0.8 \end{bmatrix} \implies \mathbf{A}^{-1} = \begin{bmatrix} 0.2 & 0.3 & -0.7 \\ -0.2 & 0.7 & -1.3 \\ 0.2 & -0.2 & 0.8 \end{bmatrix}$$

Systems of linear equations



Consider these graphs with 2 lines of form $ymx = c$.
First graph: one value of x and y satisfies both equations, at the intersection of the lines - one solution.
Second graph: two lines are parallel - no solution.
Third graph: two lines overlap - infinite solutions.

Any set of linear equations has 0, 1 or ∞ solutions

Finding 0 solutions by G.E.

Example:

$$\begin{aligned} 2x + y + 3z &= 4 \\ x + y + 2z &= 0 \\ 2x + 4y + 6z &= 8 \end{aligned} \implies \tilde{\mathbf{A}} = \begin{bmatrix} 2 & 1 & 3 & 4 \\ 1 & 1 & 2 & 0 \\ 2 & 4 & 6 & 8 \end{bmatrix}$$

$$\text{Make } a_{2,1} \text{ and } a_{3,1} \text{ zero : } \tilde{\mathbf{A}} = \begin{bmatrix} 2 & 1 & 3 & 4 \\ 0 & -1 & -1 & 4 \\ 0 & -3 & -3 & -4 \end{bmatrix}$$

Row2 := Row1-2*Row2; Row3 := Row1-Row3

$$\text{Make } a_{3,2} \text{ zero : } \tilde{\mathbf{A}} = \begin{bmatrix} 2 & 1 & 3 & 4 \\ 0 & -1 & -1 & 4 \\ 0 & 0 & 0 & 16 \end{bmatrix}$$

Row3 := 3*Row2 - Row3

Last row means $0 = 16$! So there is no solution.

Finding Infinite Solutions

$$\begin{aligned} 2x + y + 3z &= 4 \\ x + y + 2z &= 0 \\ 2x + 4y + 6z &= 8 \end{aligned} \implies \tilde{\mathbf{A}} = \begin{bmatrix} 2 & 1 & 3 & 4 \\ 1 & 1 & 2 & 0 \\ 2 & 4 & 6 & -8 \end{bmatrix}$$

$$\text{Make } a_{2,1} \text{ and } a_{3,1} \text{ zero : } \tilde{\mathbf{A}} = \begin{bmatrix} 2 & 1 & 3 & 4 \\ 0 & -1 & -1 & 4 \\ 0 & -3 & -3 & 12 \end{bmatrix}$$

Row2 := Row1-2*Row2; Row3 := Row1-Row3

$$\text{Make } a_{3,2} \text{ zero : } \tilde{\mathbf{A}} = \begin{bmatrix} 2 & 1 & 3 & 4 \\ 0 & -1 & -1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Row3 := 3*Row2 - Row3

Last row means $0 = 0$; Equations true for all values of x, y and z so ∞ solutions to equations.
The equations linearly dependent.
If there are solutions, equations are linearly independent.

Fundamental Theorem of Linear Systems

Can determine how many solutions using Rank of matrix

If system defined by m row matrix equation $\mathbf{A}x = \mathbf{b}$

The system has solutions only if $\text{Rank}(\mathbf{A}) = \text{Rank}(\tilde{\mathbf{A}})$

- If $\text{Rank}(\mathbf{A}) = m$ - one solution
- If $\text{Rank}(\mathbf{A}) < m$ - infinite number of solutions

Homogeneous Systems

If system is $\mathbf{A}x = 0$, i.e. $\mathbf{b} = 0$, then system is homogenous.
Such a system has a trivial solution $x_1 = x_2 = \cdots = x_n = 0$.
A non-trivial solution exists if $\text{Rank}(\mathbf{A}) < m$.

Cramer's Rule:
For homogeneous systems

- If $D = |\mathbf{A}| \neq 0$, the only solution is $\mathbf{x} = \mathbf{0}$
- If $D = 0$, the system has non-trivial solutions

This is useful for eigenvalues and eigenvectors.

Cramer's Theorem

Alternative to Gaussian Elimination
Consider a linear system $\mathbf{A}x = \mathbf{b}$ where $\mathbf{A}$ is square $(n \times n)$
The solution is

$$\begin{aligned} x_1 &= D_1/D \\ x_2 &= D_2/D \\ \cdots \\ x_n &= D_n/D \end{aligned}$$

where $D = |\mathbf{A}| \neq 0$ and D_k is determinant of matrix formed by taking $\mathbf{A}$ and replacing its k^{th} column with $\mathbf{b}$.

Impracticable in large matrices as hard to find determinant.

Example

Solve the system

$$\begin{bmatrix} 0.96 & -0.8 \\ 0.28 & 0.6 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 100 \end{bmatrix}$$

Here, $D = 0.96 \times 0.6 - -0.8 \times 0.28 = 0.8$
Replacing first column of $\mathbf{A}$ with $\mathbf{b}$:

$$D_1 = \begin{vmatrix} 0 & -0.8 \\ 300 & 0.6 \end{vmatrix} = 240 \text{ thus } x = \frac{240}{0.8} = 300$$

Replacing 2nd column of $\mathbf{A}$ with $\mathbf{b}$:

$$D_2 = \begin{vmatrix} 0.96 & 0 \\ 0.28 & 300 \end{vmatrix} = 288 \text{ thus } x = \frac{288}{0.8} = 360$$

Eigenvalues and Eigenvectors

Let $\mathbf{A}$ be $n \times n$ and consider the equation

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

where $\mathbf{x}$ is an $n \times 1$ vector and λ a scalar.

- The n scalars λ satisfying $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ are the eigenvalues (or characteristic values or latent roots) of $\mathbf{A}$
- For each λ, an $\mathbf{x}$ satisfying $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ is an eigenvector of $\mathbf{A}$

To find λ and $\mathbf{x}$:
Equation can be reorganised as

$$\mathbf{A}\mathbf{x} - \lambda\mathbf{x} = \mathbf{0} \text{ or } (\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = \mathbf{0}$$

This is a homogeneous equation with trivial solution $\mathbf{x} = \mathbf{0}$.
By Cramers Rule, non trivial solution exists if

$$|\mathbf{A} - \lambda\mathbf{I}| = \mathbf{0}$$

The n eigenvalues found by solving this equation (characteristic equation)

Example

$$\mathbf{A} = \begin{bmatrix} -4 & 1 \\ 2 & -3 \end{bmatrix}; \quad \mathbf{A} - \lambda\mathbf{I} = \begin{bmatrix} -4 & 1 \\ 2 & -3 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} -4-\lambda & 1 \\ 2 & -3-\lambda \end{bmatrix}$$

Thus

$$\begin{aligned} |\mathbf{A} - \lambda\mathbf{I}| &= (-4-\lambda)(-3-\lambda) - 1 \times 2 \\ &= 12 + \lambda^2 + 4\lambda + 3\lambda - 2 \end{aligned}$$

and, therefore the characteristic equation is

$$\lambda^2 + 7\lambda + 10 = 0 \ .$$

This gives the characteristic polynomial

$$(\lambda + 5)(\lambda + 2)$$

with roots $\lambda = -5$ and $\lambda = -2$
These are the two eigenvalues for the matrix A.

As eigenvectors of $\mathbf{A}$ are $\mathbf{x}$ satisfying $(\mathbf{A} - \lambda \mathbf{I})\mathbf{x} = 0$ for each λ:

$$\text{For } \lambda = 5: \quad \begin{bmatrix} -4 - -5 & 1 \\ 2 & -3 - -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

This represents equations

$$\begin{aligned} x_1 + x_2 &= 0 \\ 2x_1 + 2x_2 &= 0 \end{aligned}$$

Not independent - infinitely many solutions - always happens

So choose $x_1 = 1$, say, then $x_2 = -1$. Thus an eigenvector is $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ as are all multiples of this vector

For the other eigenvector we get,

$$\begin{aligned} -2x_1 + x_2 &= 0 \\ 2x_1 - x_2 &= 0 \end{aligned}$$

Again not independent: let $x_1 = 1$, so $x_2 = 2$ to get eigenvector $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$

Application - Diagonalisation

To diagonalise an $n \times n$ matrix $\mathbf{A}$ (eigen decomposition).
Let its eigenvalues and eigenvectors be $\lambda_1, \lambda_2, \ldots, \lambda_n$, and $\mathbf{\Lambda}_1, \mathbf{\Lambda}_2, \ldots, \mathbf{\Lambda}_n$. Write

$$\mathbf{U} = \begin{bmatrix} \mathbf{\Lambda}_1 & \mathbf{\Lambda}_2 & \cdots & \mathbf{\Lambda}_n \end{bmatrix} .$$

Then

$$\mathbf{U}^{-1}\mathbf{A}\mathbf{U} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \lambda_n \end{bmatrix}$$

Example:

For $\mathbf{A} = \begin{bmatrix} -4 & 1 \\ 2 & -3 \end{bmatrix}$; We have $\mathbf{U} = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}$ and $\mathbf{U}^{-1} = \frac{1}{3}\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}$

$$\text{So } \mathbf{U}^{-1}\mathbf{A}\mathbf{U} = \frac{1}{3}\begin{bmatrix} -10 & 5 \\ -2 & -2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} -15 & 0 \\ 0 & -6 \end{bmatrix}$$

In R:

```
A<-cbind(c(-4,2),c(1,-3))
eigen(A,only.values = FALSE)
eigen(A,only.values = TRUE)   #For eigenvalues only
```

Note: normalised (later) eigenvectors are produced

If $\mathbf{U}$ not invertible, then eigen decomposition not possible

Singular Values

For an $m \times n$ matrix $\mathbf{A}$.

- Form the symmetric matrix $\mathbf{A}^T\mathbf{A}$.
- Find the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of this symmetric matrix

The singular values are $\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n}$

These are used in Principal Component Analysis (later), Singular Value Decomposition, etc

Note: Eigenvectors of $\mathbf{A}^T\mathbf{A}$ (corresponding to distinct eigenvalues) are orthogonal (later).

That is matrix of eigenvectors $\mathbf{U}$ satisfy $\mathbf{U}^T\mathbf{U} = \mathbf{I}$

Singular Value Decomposition

For any matrix $\mathbf{A}$, singular value decomposition (SVD) constructs matrices $\mathbf{U}$ and $\mathbf{V}$ and diagonal matrix $\mathbf{D}$ such that $\mathbf{A} = \mathbf{UDV}$

- $\mathbf{D} = diag(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n})$, singular values of $\mathbf{AA}^T$
- columns of $\mathbf{U}$ are orthonormal vectors of $\mathbf{AA}^T$
- columns of $\mathbf{V}$ are orthonormal vectors of $\mathbf{A}^T\mathbf{A}$

Example: $\mathbf{A} = \begin{bmatrix} 2 & -3 \\ -5 & 1 \\ 1 & 2 \\ 2 & 0 \end{bmatrix}$; $\mathbf{AA}^T = \begin{bmatrix} 13 & -13 & -4 & 4 \\ -13 & 26 & -3 & -10 \\ -4 & -3 & 5 & 2 \\ 4 & -10 & 2 & 4 \end{bmatrix}$ $\mathbf{A}^T\mathbf{A} = \begin{bmatrix} 34 & -9 \\ -9 & 14 \end{bmatrix}$

```
eig<-eigen(A%*%t(A))$values
eig
[1]  3.745362e+01  1.054638e+01  2.842171e-14 -1.165734e-15
U<-eigen(A%*%t(A))$vector
V<-eigen(t(A)%*%A)$vector

D<-diag(sqrt(eig[1:2]));   U<--U[,1:2];   V[,2]<--V[,2]
```

```
U%*%D%*%V
     [,1]          [,2]
[1,]    2  3.000000e+00
[2,]   -5 -1.000000e+00
[3,]    1 -2.000000e+00
[4,]    2 -1.887379e-15
```

More simply:

```
sv<-svd(A)
> sv
$d
[1] 6.119937 3.247518

$v
          [,1]      [,2]
[1,] -0.9336204 0.3582638
[2,]  0.3582638 0.9336204
```

```
$u
            [,1]       [,2]
[1,] -0.48072918 -0.6418235
[2,]  0.82131011 -0.2641089
[3,] -0.03547306  0.6852939
[4,] -0.30510787  0.2206385
```


Positive Definite Matrices

Suppose **A** is a symmetric $n \times n$ matrix (i.e. $\mathbf{A} = \mathbf{A}^T$.)
For any non-zero column vector **x** containing n real numbers

Definition

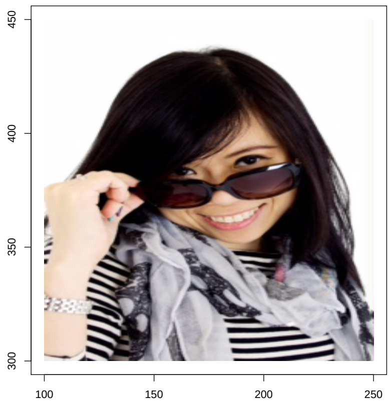
A is positive definite if $\mathbf{x}^T\mathbf{A}\mathbf{x}$ is positive (> 0)
A is positive semi-definite if $\mathbf{x}^T\mathbf{A}\mathbf{x} \geq 0$
A is negative definite if $\mathbf{x}^T\mathbf{A}\mathbf{x} < 0$

- A symmetric matrix **A** is positive definite if all its eigenvalues are positive
- Matrix **A** is also positive definite if
  1. all the diagonal entries are positive, and
  2. each diagonal entry is greater than the sum of the absolute values of all other entries in the corresponding row/column.

Images in R

```
library(bmp)
photo_greyscale.bmp <- read.bmp(".../image1_gs.bmp")
dim(photo_greyscale.bmp)
[1] 784 785   ## range 0 - 255
par(mfrow=c(1,2))
image(photo_greyscale.bmp,col = gray((0:32)/32),axes=F)
image(rotate.m(photo_greyscale.bmp),col = gray((0:32)/32),axes=F)
```


```
library(jpeg)
photo_colour.jpg <- readJPEG(".../image1.jpg")
dim(photo_colour.jpg)
[1] 320 214   3  ## range 0 - 1
plot(c(100, 250), c(300, 450), type = "n", xlab = "", ylab = "")
rasterImage(photo_colour.jpg,xleft=100, ybottom=300, xright=250, ytop=450
```



More on Vectors

A scalar has magnitude, e.g. mass, speed, temperature, position
A vector has magnitude and direction e.g. force, velocity
Can define position in (x, y, z) terms, e.g.
$\mathbf{P} = (1, 2, 5), \quad \mathbf{Q} = (4, -2, 1)$
A vector defines movement from one position to another
Vector $\mathbf{P}$ to $\mathbf{Q}$ has components
4-1 in x, -2-2 in y and 1-5 in z
In general, if $\mathbf{P}$ is point (x_1, y_1, z_1) and $\mathbf{Q}$ is (x_2, y_2, z_2)
then the vector from $\mathbf{P}$ to $\mathbf{Q}$ is $(x_2 - x_1, y_2 - y_1, z_2 - z_1)$

Vector Norm

Write $\mathbf{x} = (x_1, x_2, \ldots, x_n)$
Length of the vector $\mathbf{x}$, sometimes called the Euclidean distance, is

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

Vectors of length 1 are unit vectors

This is another norm, sometimes called Taxicab or Manhattan norm,

$$\|\mathbf{x}\|_1 = |x_1| + |x_2| + \cdots + |x_n|$$

In general, the L_p norm ($p \geq 1$ and real) is

$$\|\mathbf{x}\|_p = (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{\frac{1}{p}}$$

- If $p = 1$ this is Manhattan norm
- If $p = 2$ this is Euclidean norm

Matrix Norm

Extension of vector norm to matrices.

For $m \times n$ matrix $\mathbf{A}$ and $n \times 1$ vector $\mathbf{x}$, the p-norm of $\mathbf{A}$ is

$$\|\mathbf{A}\|_p = \sup_{\mathbf{x} \neq 0} \frac{\|\mathbf{Ax}\|_p}{\|\mathbf{x}\|_p} = \max_{\|\mathbf{x}\|_p = 1} \|\mathbf{Ax}\|_p$$

Here sup is the supremum or least upper bound

- For $p = 1$ the norm is the maximum of the sum of each column in $\mathbf{A}$
- For $p = 2$ the norm is the largest singular value of $\mathbf{A}$

$$\|\mathbf{A}\|_2 = \max\left(\sqrt{eig(\mathbf{A}^T\mathbf{A})}\right)$$

where $eig()$ is the set of eigenvalues.

Inner product

A convenient representation of vectors in 3D is to use unit vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$ where

$$\mathbf{i} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}; \quad \mathbf{j} = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}; \quad \mathbf{k} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$$

For example $\begin{bmatrix} 2 & 3 & 1 \end{bmatrix}$ is the same as $2\mathbf{i} + 3\mathbf{j} + \mathbf{k}$

The inner (dot) product of two vectors $\mathbf{a}$ and $\mathbf{b}$ is a scalar

$$\mathbf{a} \bullet \mathbf{b} = \|a\|\|b\| \cos(\theta) \text{ where } \theta \text{ is angle between } \mathbf{a} \text{ and } \mathbf{b}$$

In particular $\mathbf{a} \bullet \mathbf{a} = \|a\|^2 \cos(0) = \|a\|^2$. Thus $\|a\| = \sqrt{(}\mathbf{a} \bullet \mathbf{a})$

Suppose $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ and $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$, then

$$\mathbf{a} \bullet \mathbf{b} = a_1b_1 + a_2b_2 + a_3b_3$$

In general, if $\mathbf{a} = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$ and $b = \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix}$ then

$$\mathbf{a} \bullet \mathbf{b} = a_1b_1 + a_2b_2 + \ldots + a_nb_n$$

For 2 orthogonal vectors ($\theta = 90°$) the dot product is 0.

Cross product

Cross product $\mathbf{v} = \mathbf{a} \times \mathbf{a}$ of 2 vectors $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ and $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$ is also a vector at right angles to both $\mathbf{a}$ and $\mathbf{b}$

Length of $\mathbf{v}$ is defined as

$$\|a\|\|b\|\sin(\theta) \text{ where } \theta \text{ is angle between } \mathbf{a} \text{ and } \mathbf{b}$$

The components of $\mathbf{v}$ are

$$\begin{bmatrix} a_2b_3 - a_3b_2 & a_3b_1 - a_1b_3 & a_1b_2 - a_2b_1 \end{bmatrix}$$

Not easy to remember - but there is a trick from Matrices:

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

Note: $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$

Examples

Suppose $\mathbf{a} = 4\mathbf{i} + \mathbf{j} + 2\mathbf{k}$ and $\mathbf{b} = 4\mathbf{i} + 3\mathbf{j} + \mathbf{k}$ then

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 4 & 1 & 2 \\ 4 & 3 & 1 \end{vmatrix} = -5\mathbf{i} + (-4)(-\mathbf{j}) + 8\mathbf{k} = -5\mathbf{i} + 4\mathbf{j} + 8\mathbf{k}$$

The Scalar Triple Product

$$\mathbf{a} \bullet (\mathbf{b} \times \mathbf{c}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

Applications exist for these products (computer graphics, fluid flow, etc)

Vector Space

A linear vector space is a set $\mathbf{V}$ whose elements are vectors on which addition and multiplication can act.

In other words, if vectors $\mathbf{x}$ and $\mathbf{y}$ are in $\mathbf{V}$; $\mathbf{x}, \mathbf{y} \in \mathbf{V}$ and a and b are real numbers,

- $\mathbf{x} + \mathbf{y} \in \mathbf{V}$ and $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$
- a zero vector $\mathbf{0} \in \mathbf{V}$ exists such that $\mathbf{x} + \mathbf{0} = \mathbf{x}$ for all $\mathbf{x} \in \mathbf{V}$
- $a\mathbf{x} \in \mathbf{V}$ for all $\mathbf{x} \in \mathbf{V}$
- $a(\mathbf{x} + \mathbf{y}) = a\mathbf{x} + b\mathbf{y}; (a + b)\mathbf{x} = a\mathbf{x} + b\mathbf{x}$

These properties apply to $\mathbb{R}^n$, n dimensions of real numbers $\mathbb{R}$

Sets - briefly

- A subset $\mathbf{U}$ of $\mathbb{R}^n$, denoted $\mathbf{U} \subset \mathbb{R}^n$, is said to be open if for all vectors $\mathbf{x} \in \mathbf{U}$ a neighbourhood of radius δ,
$$N(\mathbf{x}, \delta) = \{\mathbf{y} \in \mathbb{R}^n; \|\mathbf{y} - \mathbf{x}\| < \delta\}$$
of $\mathbf{x}$ can be found with $N(\mathbf{x}, \delta) \subset \mathbf{U}$
- A set is closed if its complement in $\mathbb{R}^n$ is open.
- A point $\mathbf{x}$ is a boundary point of $\mathbf{U} \subset \mathbb{R}^n$ if every neighbourhood of $\mathbf{x}$ contains at least one point belonging to $\mathbf{U}$ and one not.
- The boundary of $\mathbf{U}$, denoted $\partial\mathbf{U}$, is the set of all boundary points of $\mathbf{U}$
- The interior of $\mathbf{U}$ is the set of points in $\mathbf{U}$ that are not in $\partial\mathbf{U}$
- The closure $\mathbf{U}$ is the union of $\mathbf{U}$ and its boundary $\partial\mathbf{U}$
- An open set is equal to its interior, a closed set is equal to its closure.
- A subset $\mathbf{U}$ of $\mathbb{R}^n$ is bounded if there is an $r > 0$ such that $\|\mathbf{x}\| < r$ for all $\mathbf{x} \in \mathbf{U}$
- A subset $\mathbf{U}$ of $\mathbb{R}^n$ is compact if it is closed and bounded