More Linear Algebra[1]

STA 302: Fall 2017

[1] See Chapter 2 of *Linear models in statistics* for more detail. This slide show is an open-source document. See last slide for copyright information.

Overview

You already know about

- Matrices $A = (a_{ij})$
- Column vectors $\mathbf{v} = (v_j)$
- Matrix addition and subtraction $A + B = (a_{ij} + b_{ij})$
- Scalar multiplication $aB = (a\, b_{ij})$
- Matrix multiplication $AB = \left(\sum_k a_{ik} b_{kj} \right)$

  In words: The i,j element of AB is the inner product of row i of A with column j of B.
- Inverse: $A^{-1}A = AA^{-1} = I$
- Transpose $A' = (a_{ji})$
- Symmetric matrices: $A = A'$
- Determinants
- Linear independence

Inverses: Proving $B = A^{-1}$

- $B = A^{-1}$ means $AB = BA = I$.
- It looks like you have two things to show.
- But if A and B are square matrices of the same size, you only need to do it in one direction.

Theorem

If A and B are square matrices and $AB = I$, then A and B are inverses.

Proof: Suppose $AB = I$

- A and B must both have inverses, for otherwise $|AB| = |A|\,|B| = 0 \neq |I| = 1$. Now,
- $AB = I \Rightarrow ABB^{-1} = IB^{-1} \Rightarrow A = B^{-1}$.
- $AB = I \Rightarrow A^{-1}AB = A^{-1}I \Rightarrow B = A^{-1}$.

How to show $A^{-1\prime} = A^{\prime -1}$

- Let $B = A^{-1}$.
- Want to prove that B' is the inverse of A'.
- It is enough to show that $B'A' = I$.
- $AB = I \Rightarrow B'A' = I' = I \quad \blacksquare$

Three mistakes that will get you a zero

Numbers are 1×1 matrices, but larger matrices are not just numbers.

You will get a zero if you

- Write $AB = BA$. It's not true in general.
- Write A^{-1} when A is not a square matrix. The inverse is not even defined.
- Represent the inverse of a matrix (even if it exists) by writing it in the denominator, like $\mathbf{a}'B^{-1}\mathbf{a} = \frac{\mathbf{a}'\mathbf{a}}{B}$.

Matrices are not just numbers.

If you commit one of these crimes, the mark for the question (or part of a question, like 3c) is zero. The rest of your answer will be ignored, and will also get a zero.

Half marks off, at least

You will lose *at least* half marks for writing a product like AB when the number of colmns in A does not equal the number of rows in B.

Linear combination of vectors

Let $\mathbf{x}_1, \ldots, \mathbf{x}_p$ be $n \times 1$ vectors and $a_1, \ldots, a_p$ be scalars. A *linear combination* of the vectors is

$$\begin{aligned}
\mathbf{c} &= a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_p\mathbf{x}_p \\
&= a_1 \begin{pmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{pmatrix} + a_2 \begin{pmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{n2} \end{pmatrix} + \cdots + a_p \begin{pmatrix} x_{1p} \\ x_{2p} \\ \vdots \\ x_{np} \end{pmatrix}
\end{aligned}$$

Linear independence

A set of vectors $\mathbf{x}_1, \ldots, \mathbf{x}_p$ is said to be *linearly dependent* if there is a set of scalars $a_1, \ldots, a_p$, not all zero, with

$$a_1 \begin{pmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{pmatrix} + a_2 \begin{pmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{n2} \end{pmatrix} + \cdots + a_p \begin{pmatrix} x_{1p} \\ x_{2p} \\ \vdots \\ x_{np} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

If no such constants $a_1, \ldots, a_p$ exist, the vectors are linearly independent. That is,

If $a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_p\mathbf{x}_p = \mathbf{0}$ implies $a_1 = a_2 \cdots = a_p = 0$, then the vectors are said to be *linearly independent*.

Bind the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_p$ into a matrix

$$
\begin{aligned}
& a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_p\mathbf{x}_p \\
= & \begin{pmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{pmatrix} a_1 + \begin{pmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{n2} \end{pmatrix} a_2 + \cdots + \begin{pmatrix} x_{1p} \\ x_{2p} \\ \vdots \\ x_{np} \end{pmatrix} a_p \\
= & \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & n_{np} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix} \\
= & X\mathbf{a}
\end{aligned}
$$

A more convenient definition of linear independence

$a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_p\mathbf{x}_p = X\mathbf{a}$

Let X be an $n \times p$ matrix of constants. The columns of X are said to be *linearly dependent* if there exists $\mathbf{a} \neq \mathbf{0}$ with $X\mathbf{a} = \mathbf{0}$. We will say that the columns of X are linearly *independent* if $X\mathbf{a} = \mathbf{0}$ implies $\mathbf{a} = \mathbf{0}$.

For example, show that the existence of B^{-1} implies that the columns of B are linearly independent.

$$B\mathbf{a} = \mathbf{0} \Rightarrow B^{-1}B\mathbf{a} = B^{-1}\mathbf{0} \Rightarrow \mathbf{a} = \mathbf{0} \quad \blacksquare$$

Trace of a square matrix

- Sum of diagonal elements
- Obvious: $tr(A+B) = tr(A) + tr(B)$
- Not obvious: $tr(AB) = tr(BA)$
- Even though $AB \neq BA$.

$tr(AB) = tr(BA)$

Let A be $p \times q$ and B be $q \times p$, so that AB is $p \times p$ and BA is $q \times q$.

First, agree that $\sum_{i=1}^n x_i = \sum_{j=1}^n x_j$.

$$
\begin{aligned}
tr(AB) &= tr\left(\left[\sum_{k=1}^{q} a_{ik} b_{kj}\right]\right) \\
&= \sum_{i=1}^{p} \sum_{k=1}^{q} a_{ik} b_{ki} \\
&= \sum_{k=1}^{q} \sum_{i=1}^{p} b_{ki} a_{ik} \\
&= \sum_{i=1}^{q} \sum_{k=1}^{p} b_{ik} a_{ki} \\
&= tr\left(\left[\sum_{k=1}^{p} b_{ik} a_{kj}\right]\right) \\
&= tr(BA)
\end{aligned}
$$

Eigenvalues and eigenvectors

Let $A = [a_{i,j}]$ be an $n \times n$ matrix, so that the following applies to square matrices. A is said to have an *eigenvalue* λ and (non-zero) *eigenvector* $\mathbf{x} \neq \mathbf{0}$ corresponding to λ if

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

Eigenvectors can be scaled to have length one, so that $\mathbf{x}'\mathbf{x} = 1$.

- Eigenvalues are the λ values that solve the determinantal equation $|A - \lambda I| = 0$.
- The determinant is the product of the eigenvalues: $|A| = \prod_{i=1}^{n} \lambda_i$

Spectral decomposition of symmetric matrices

The *Spectral decomposition theorem* says that every square and symmetric matrix $A = [a_{i,j}]$ may be written

$$A = CDC',$$

where the columns of C (which may also be denoted $\mathbf{x}_1, \ldots, \mathbf{x}_n$) are the eigenvectors of A, and the diagonal matrix D contains the corresponding eigenvalues.

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

The eigenvectors may be chosen to be orthonormal, so that C is an orthogonal matrix. That is, $CC' = C'C = I$.

Positive definite matrices

The $n \times n$ matrix A is said to be *positive definite* if

$$\mathbf{y}'A\mathbf{y} > 0$$

for *all* $n \times 1$ vectors $\mathbf{y} \neq \mathbf{0}$. It is called *non-negative definite* (or sometimes positive semi-definite) if $\mathbf{y}'A\mathbf{y} \geq 0$.

Example: Show $X'X$ non-negative definite

Let X be an $n \times p$ matrix of real constants and let $\mathbf{y}$ be $p \times 1$. Then $\mathbf{z} = X\mathbf{y}$ is $n \times 1$, and

$$
\begin{aligned}
& \mathbf{y}'\,(X'X)\,\mathbf{y} \\
= \; & (X\mathbf{y})'(X\mathbf{y}) \\
= \; & \mathbf{z}'\mathbf{z} \\
= \; & \sum_{i=1}^{n} z_i^2 \geq 0 \;\; \blacksquare
\end{aligned}
$$

Some properties of symmetric positive definite matrices

Variance-covariance matrices are often assumed positive definite.

For a symmetric matrix,

Positive definite

$\Downarrow$

All eigenvalues positive

$\Downarrow$

Inverse exists $\Leftrightarrow$ Columns (rows) linearly independent.

If a real symmetric matrix is also non-negative definite, as a variance-covariance matrix *must* be, Inverse exists $\Rightarrow$ Positive definite

Showing Positive definite $\Rightarrow$ Eigenvalues positive

Let the $p \times p$ matrix A be positive definite, so that $\mathbf{y}' A \mathbf{y} > 0$ for all $\mathbf{y} \neq \mathbf{0}$.

λ an eigenvalue means $A\mathbf{x} = \lambda\mathbf{x}$ with $\mathbf{x}'\mathbf{x} = 1$.

$\Rightarrow \mathbf{x}' A \mathbf{x} = \mathbf{x}' \lambda \mathbf{x} = \lambda \mathbf{x}'\mathbf{x} = \lambda > 0. \quad \blacksquare$

Inverse of a diagonal matrix

To set things up

Suppose $D = [d_{i,j}]$ is a diagonal matrix with non-zero diagonal elements. It is easy to verify that

$$\left(\begin{array}{cccc} 1/d_{1,1} & 0 & \cdots & 0 \\ 0 & 1/d_{2,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/d_{n,n} \end{array} \right) \left(\begin{array}{cccc} d_{1,1} & 0 & \cdots & 0 \\ 0 & d_{2,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{n,n} \end{array} \right) = I$$

So D^{-1} exists.

Showing Eigenvalues positive $\Rightarrow$ Inverse exists

For a symmetric, positive definite matrix

Let A be symmetric and positive definite. Then $A = CDC'$, and its eigenvalues are positive.

Let $B = CD^{-1}C'$. Show $B = A^{-1}$.

$$AB = CDC' \, CD^{-1}C' = I$$

So

$$A^{-1} = CD^{-1}C'$$

Square root matrices

For symmetric, non-negative definite matrices

To set things up, define

$$D^{1/2} = \begin{pmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_n} \end{pmatrix}$$

So that

$$\begin{aligned} D^{1/2}D^{1/2} &= \begin{pmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_n} \end{pmatrix} \begin{pmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_n} \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} = D \end{aligned}$$

For a non-negative definite, symmetric matrix A

Define

$$A^{1/2} = CD^{1/2}C'$$

So that

$$\begin{aligned} A^{1/2}A^{1/2} &= CD^{1/2}C'CD^{1/2}C' \\ &= CD^{1/2}\,I\,D^{1/2}C' \\ &= CD^{1/2}D^{1/2}C' \\ &= CDC' \\ &= A \end{aligned}$$

The square root of the inverse is the inverse of the square root

Let A be symmetric and positive definite, with $A = CDC'$.

Let $B = CD^{-1/2}C'$. What is $D^{-1/2}$?

Show $B = \left(A^{-1}\right)^{1/2}$.

$$\begin{aligned} BB &= CD^{-1/2}C'CD^{-1/2}C' \\ &= CD^{-1}C' = A^{-1} \end{aligned}$$

Show $B = \left(A^{1/2}\right)^{-1}$

$$A^{1/2}B = CD^{1/2}C'CD^{-1/2}C' = I$$

Just write $$A^{-1/2} = CD^{-1/2}C'$$

Extras

You may not know about these, but we may use them occasionally

- Rank
- Partitioned matrices

Rank

- Row rank is the number of linearly independent rows.
- Column rank is the number of linearly independent columns.
- Rank of a matrix is the minimum of row rank and column rank.
- $rank(AB) = \min\left(rank(A), rank(B)\right)$.

Partitioned matrix

- A matrix of matrices

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$$

- Row by column (matrix) multiplication works, provided the matrices are the right sizes.

Matrix calculation with R

```
> is.matrix(3)  # Is the number 3 a 1x1 matrix?

[1] FALSE


> treecorr = cor(trees); treecorr

           Girth    Height    Volume
Girth  1.0000000 0.5192801 0.9671194
Height 0.5192801 1.0000000 0.5982497
Volume 0.9671194 0.5982497 1.0000000


> is.matrix(treecorr)

[1] TRUE
```

Creating matrices
Bind rows into a matrix

```
> # Bind rows of a matrix together
> A = rbind(  c(3, 2, 6,8),
+             c(2,10,-7,4),
+             c(6, 6, 9,1)   ); A

     [,1] [,2] [,3] [,4]
[1,]    3    2    6    8
[2,]    2   10   -7    4
[3,]    6    6    9    1


> # Transpose
> t(A)

     [,1] [,2] [,3]
[1,]    3    2    6
[2,]    2   10    6
[3,]    6   -7    9
[4,]    8    4    1
```

Matrix multiplication

Remember, A is 3×4

```
> # U = AA' (3x3), V = A'A (4x4)
> U = A %*% t(A)
> V = t(A) %*% A; V

     [,1] [,2] [,3] [,4]
[1,]   49   62   58   38
[2,]   62  140   -4   62
[3,]   58   -4  166   29
[4,]   38   62   29   81
```

Determinants

```
> # U = AA' (3x3), V = A'A (4x4)
> # So rank(V) cannot exceed 3 and det(V)=0
> det(U); det(V)

[1] 1490273
[1] -3.622862e-09
```

Inverse of U exists, but inverse of V does not.

Inverses

- The `solve` function is for solving systems of linear equations like $M\mathbf{x} = \mathbf{b}$.
- Just typing `solve(M)` gives M^{-1}.

```
> # Recall U = AA' (3x3), V = A'A (4x4)
> solve(U)

              [,1]          [,2]          [,3]
[1,]  0.0173505123 -8.508508e-04 -1.029342e-02
[2,] -0.0008508508  5.997559e-03  2.013054e-06
[3,] -0.0102934160  2.013054e-06  1.264265e-02


> solve(V)


Error in solve.default(V) :
  system is computationally singular: reciprocal condition
  number = 6.64193e-18
```

Eigenvalues and eigenvectors

```
> # Recall U = AA' (3x3), V = A'A (4x4)
> eigen(U)

$values
[1] 234.01162 162.89294  39.09544

$vectors
           [,1]       [,2]        [,3]
[1,] -0.6025375  0.1592598  0.78203893
[2,] -0.2964610 -0.9544379 -0.03404605
[3,] -0.7409854  0.2523581 -0.62229894
```

V should have at least one zero eigenvalue

Because A is 3×4, $V = A'A$, and the rank of a product is the minimum rank of the matrices.

```
> eigen(V)

$values
[1]  2.340116e+02  1.628929e+02  3.909544e+01 -1.012719e-14

$vectors
           [,1]         [,2]       [,3]         [,4]
[1,] -0.4475551  0.006507269 -0.2328249  0.863391352
[2,] -0.5632053 -0.604226296 -0.4014589 -0.395652773
[3,] -0.5366171  0.776297432 -0.1071763 -0.312917928
[4,] -0.4410627 -0.179528649  0.8792818  0.009829883
```

Spectral decomposition $V = CDC'$

```
> eigenV = eigen(V)
> C = eigenV$vectors; D = diag(eigenV$values); D

         [,1]     [,2]     [,3]          [,4]
[1,] 234.0116   0.0000  0.00000  0.000000e+00
[2,]   0.0000 162.8929  0.00000  0.000000e+00
[3,]   0.0000   0.0000 39.09544  0.000000e+00
[4,]   0.0000   0.0000  0.00000 -1.012719e-14
```

```
> # C is an orthoganal matrix
> C %*% t(C)

              [,1]         [,2]         [,3]          [,4]
[1,]  1.000000e+00 5.551115e-17 0.000000e+00 -3.989864e-17
[2,]  5.551115e-17 1.000000e+00 2.636780e-16  3.556183e-17
[3,]  0.000000e+00 2.636780e-16 1.000000e+00  2.558717e-16
[4,] -3.989864e-17 3.556183e-17 2.558717e-16  1.000000e+00
```

Verify $V = CDC'$

```
> V;    C %*% D %*% t(C)

     [,1] [,2] [,3] [,4]
[1,]   49   62   58   38
[2,]   62  140   -4   62
[3,]   58   -4  166   29
[4,]   38   62   29   81

     [,1] [,2] [,3] [,4]
[1,]   49   62   58   38
[2,]   62  140   -4   62
[3,]   58   -4  166   29
[4,]   38   62   29   81
```

Square root matrix $V^{1/2} = CD^{1/2}C'$

```
> sqrtV = C %*% sqrt(D) %*% t(C)

Warning message:
In sqrt(D) : NaNs produced

> # Multiply to get V
> sqrtV %*% sqrtV;  V

     [,1] [,2] [,3] [,4]
[1,]  NaN  NaN  NaN  NaN
[2,]  NaN  NaN  NaN  NaN
[3,]  NaN  NaN  NaN  NaN
[4,]  NaN  NaN  NaN  NaN
     [,1] [,2] [,3] [,4]
[1,]   49   62   58   38
[2,]   62  140   -4   62
[3,]   58   -4  166   29
[4,]   38   62   29   81
```

What happened?

```
> D; sqrt(D)

         [,1]     [,2]     [,3]          [,4]
[1,] 234.0116   0.0000  0.00000  0.000000e+00
[2,]   0.0000 162.8929  0.00000  0.000000e+00
[3,]   0.0000   0.0000 39.09544  0.000000e+00
[4,]   0.0000   0.0000  0.00000 -1.012719e-14

         [,1]     [,2]     [,3] [,4]
[1,] 15.29744  0.00000 0.000000    0
[2,]  0.00000 12.76295 0.000000    0
[3,]  0.00000  0.00000 6.252635    0
[4,]  0.00000  0.00000 0.000000  NaN

Warning message:
In sqrt(D) : NaNs produced
```

Copyright Information