## Matrix Structure

## Computational Mathematics and Statistics Camp

University of Chicago September 2018

1. Recall the general rule for calculating the determinant of an  $n \times n$  matrix:

$$|\mathbf{X}| = \sum_{j=1}^{n} (-1)^{i+j} x_{ij} |\mathbf{X}_{[ij]}|$$

where the *ij*th **minor** of **X** for  $x_{ij}$ ,  $|\mathbf{X}_{[ij]}|$ , is the determinant of the  $(n-1) \times (n-1)$  submatrix that results from taking the *i*th row and *j*th column out. The **cofactor** of **X** is the minor signed as  $(-1)^{i+j}x_{ij}|\mathbf{X}_{[ij]}|$ . To calculate the determinant we cycle recursively through the columns and take sums with a formula that multiplies the cofactor by the determining value.

Given this rule, obtain the trace and determinant of the following matrix. You can do this the hard way, or the easy(ier) way. I encourage you to think a bit before starting the calculations for the determinant - how can you make the problem easier?

$$\left[\begin{array}{cccc}
6 & 6 & 1 & 0 \\
0 & 4 & 0 & 1 \\
4 & 2 & 1 & 1 \\
1 & 1 & 0 & 0
\end{array}\right]$$

a. Calculate the trace. This is straightforward - calculate the sum of the diagonal elements.

$$\operatorname{tr}\left(\begin{array}{ccc} 6 & 6 & 1 & 0\\ 0 & 4 & 0 & 1\\ 4 & 2 & 1 & 1\\ 1 & 1 & 0 & 0 \end{array}\right) = 6 + 4 + 1 + 0$$

$$= 11$$

b. Calculate the determinant. The best way to tackle this problem is to start the recursive operation on a row with the most zeros - this is because  $x_{ij}$  will be 0, so the resulting submatricies will drop out of the equation. Here I choose to start on the second row:

$$\begin{vmatrix} 6 & 6 & 1 & 0 \\ 0 & 4 & 0 & 1 \\ 4 & 2 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{vmatrix} = (-1)^{2+1}(0) \begin{vmatrix} 6 & 1 & 0 \\ 2 & 1 & 1 \\ 1 & 0 & 0 \end{vmatrix} + (-1)^{2+2}(4) \begin{vmatrix} 6 & 1 & 0 \\ 4 & 1 & 1 \\ 1 & 0 & 0 \end{vmatrix}$$
$$+ (-1)^{2+3}(0) \begin{vmatrix} 6 & 6 & 0 \\ 4 & 2 & 1 \\ 1 & 1 & 0 \end{vmatrix} + (-1)^{2+4}(1) \begin{vmatrix} 6 & 6 & 1 \\ 4 & 2 & 1 \\ 1 & 1 & 0 \end{vmatrix}$$
$$= 4 \begin{vmatrix} 6 & 1 & 0 \\ 4 & 1 & 1 \\ 1 & 0 & 0 \end{vmatrix} + 1 \begin{vmatrix} 6 & 6 & 1 \\ 4 & 2 & 1 \\ 1 & 1 & 0 \end{vmatrix}$$

Now that two of the terms have dropped out, we repeat the process on the remaining two submatricies. Again, we want to start with rows that have the most zeros, so for both submatricies I start on the third row.

1

$$= 4 \begin{vmatrix} 6 & 1 & 0 \\ 4 & 1 & 1 \\ 1 & 0 & 0 \end{vmatrix} + 1 \begin{vmatrix} 6 & 6 & 1 \\ 4 & 2 & 1 \\ 1 & 1 & 0 \end{vmatrix}$$

$$= 4 \left[ (-1)^{3+1} (1) \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} + (-1)^{3+2} (0) \begin{vmatrix} 6 & 0 \\ 4 & 1 \end{vmatrix} + (-1)^{3+3} (0) \begin{vmatrix} 6 & 1 \\ 4 & 1 \end{vmatrix} \right]$$

$$+ 1 \left[ (-1)^{3+1} (1) \begin{vmatrix} 6 & 1 \\ 2 & 1 \end{vmatrix} + (-1)^{3+2} (1) \begin{vmatrix} 6 & 1 \\ 4 & 1 \end{vmatrix} + (-1)^{3+3} (0) \begin{vmatrix} 6 & 6 \\ 4 & 2 \end{vmatrix} \right]$$

$$= 4 \left[ 1 \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} \right] + 1 \left[ 1 \begin{vmatrix} 6 & 1 \\ 2 & 1 \end{vmatrix} - 1 \begin{vmatrix} 6 & 1 \\ 4 & 1 \end{vmatrix} \right]$$

$$= 4 \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} + \begin{vmatrix} 6 & 1 \\ 2 & 1 \end{vmatrix} - \begin{vmatrix} 6 & 1 \\ 4 & 1 \end{vmatrix}$$

$$= 4 \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} + \begin{vmatrix} 6 & 1 \\ 2 & 1 \end{vmatrix} - \begin{vmatrix} 6 & 1 \\ 4 & 1 \end{vmatrix}$$

Now that the submatricies are all  $2 \times 2$ , we can use the standard formula to calculate the determinants:

$$= 4(1 \times 1 - 0 \times 1) + (6 \times 1 - 1 \times 2) - (6 \times 1 - 1 \times 4)$$

$$= 4(1 - 0) + (6 - 2) - (6 - 4)$$

$$= 4 + 4 - 2$$

$$= 6$$

2. Invert each of the following matricies. Verify you have the correct inverse by calculating  $\mathbf{X}\mathbf{X}^{-1} = \mathbf{I}$ . Not all of the matricies may be invertible - if not, show why.

a. 
$$\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

Recall the rule for inverting  $2 \times 2$  matricies:

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}$$
$$= |\mathbf{X}|^{-1} \begin{bmatrix} x_{22} & -x_{12} \end{bmatrix}$$

$$\mathbf{X}^{-1} = |\mathbf{X}|^{-1} \begin{bmatrix} x_{22} & -x_{12} \\ -x_{21} & x_{11} \end{bmatrix}$$
$$= \frac{1}{|\mathbf{X}|} \begin{bmatrix} x_{22} & -x_{12} \\ -x_{21} & x_{11} \end{bmatrix}$$

Given this rule, first calculate the determinant of the matrix.

$$|\mathbf{X}| = (2 \times 1) - (1 \times 1)$$
  
= 2 - 1

Now we can easily solve for the inverse:

$$\mathbf{X}^{-1} = \frac{1}{1} \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$$
$$= \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$$

b. 
$$\begin{bmatrix} 4 & 5 \\ 2 & 4 \end{bmatrix}$$

1. Solve for the determinant

$$|\mathbf{X}| = (4 \times 4) - (5 \times 2)$$
  
= 16 - 10  
= 6

2. Solve for the inverse

$$\mathbf{X}^{-1} = \frac{1}{6} \begin{bmatrix} 4 & -5 \\ -2 & 4 \end{bmatrix}$$
$$= \begin{bmatrix} \frac{4}{6} & -\frac{5}{6} \\ -\frac{2}{6} & \frac{4}{6} \end{bmatrix}$$
$$= \begin{bmatrix} \frac{2}{3} & -\frac{5}{6} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

c. 
$$\begin{bmatrix} 2 & 1 \\ -4 & -2 \end{bmatrix}$$

1. Solve for the determinant

$$|\mathbf{X}| = (2 \times -2) - (1 \times -4)$$
  
= -4 - (-4)  
= 0

At this point we are done. The matrix has a determinant of zero, making it singular. Singular matricies cannot be inverted.

d. 
$$\begin{bmatrix} 2 & 4 & 0 \\ 4 & 6 & 3 \\ -6 & -10 & 0 \end{bmatrix}$$

With a  $3 \times 3$  matrix, we need to apply Gauss-Jordan elimination to obtain the inverse.

1. Setup the augmented matrix with the identity matrix

$$\left[\begin{array}{ccc|cccc}
2 & 4 & 0 & 1 & 0 & 0 \\
4 & 6 & 3 & 0 & 1 & 0 \\
-6 & -10 & 0 & 0 & 0 & 1
\end{array}\right]$$

2. Swap row 1 with row 3

$$\left[ \begin{array}{ccc|ccc|c}
-6 & -10 & 0 & 0 & 0 & 1 \\
4 & 6 & 3 & 0 & 1 & 0 \\
2 & 4 & 0 & 1 & 0 & 0
\end{array} \right]$$

3. Add  $\frac{2}{3} \times \text{ row } 1 \text{ to row } 2$ 

$$\begin{bmatrix}
-6 & -10 & 0 & 0 & 0 & 1 \\
0 & -2/3 & 3 & 0 & 1 & 2/3 \\
2 & 4 & 0 & 1 & 0 & 0
\end{bmatrix}$$

3

4. Add  $\frac{1}{3} \times \text{ row } 1 \text{ to row } 3$ 

$$\left[ 
\begin{array}{ccc|cccc}
-6 & -10 & 0 & 0 & 0 & 1 \\
0 & -2/3 & 3 & 0 & 1 & 2/3 \\
0 & 2/3 & 0 & 1 & 0 & 1/3
\end{array} \right]$$

5. Add row 2 to row 3

$$\begin{bmatrix}
-6 & -10 & 0 & 0 & 0 & 1 \\
0 & -2/3 & 3 & 0 & 1 & 2/3 \\
0 & 0 & 3 & 1 & 1 & 1
\end{bmatrix}$$

6. Divide row 3 by 3

$$\left[ 
\begin{array}{ccc|ccc|c}
-6 & -10 & 0 & 0 & 0 & 1 \\
0 & -2/3 & 3 & 0 & 1 & 2/3 \\
0 & 0 & 1 & 1/3 & 1/3 & 1/3
\end{array} \right]$$

7. Subtract  $3 \times \text{ row } 3 \text{ from row } 2$ 

$$\begin{bmatrix}
-6 & -10 & 0 & 0 & 0 & 1 \\
0 & -2/3 & 0 & -1 & 0 & -1/3 \\
0 & 0 & 1 & 1/3 & 1/3 & 1/3
\end{bmatrix}$$

8. Multiply row 2 by  $-\frac{3}{2}$ 

$$\left[ \begin{array}{ccc|c}
-6 & -10 & 0 & 0 & 0 & 1 \\
0 & 1 & 0 & 3/2 & 0 & 1/2 \\
0 & 0 & 1 & 1/3 & 1/3 & 1/3
\end{array} \right]$$

9. Add  $10 \times \text{ row } 2 \text{ to row } 1$ 

$$\begin{bmatrix}
-6 & 0 & 0 & 15 & 0 & 6 \\
0 & 1 & 0 & 3/2 & 0 & 1/2 \\
0 & 0 & 1 & 1/3 & 1/3 & 1/3
\end{bmatrix}$$

10. Divide row 1 by -6

$$\left[\begin{array}{ccc|ccc|c}
1 & 0 & 0 & -5/2 & 0 & -1 \\
0 & 1 & 0 & 3/2 & 0 & 1/2 \\
0 & 0 & 1 & 1/3 & 1/3 & 1/3
\end{array}\right]$$

11. The inverse of the original matrix is the right part of the augmented matrix.

$$\begin{bmatrix} 2 & 4 & 0 \\ 4 & 6 & 3 \\ -6 & -10 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} -5/2 & 0 & -1 \\ 3/2 & 0 & 1/2 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}$$

12. Factor out common terms

$$\begin{bmatrix} 2 & 4 & 0 \\ 4 & 6 & 3 \\ -6 & -10 & 0 \end{bmatrix}^{-1} = \frac{1}{6} \begin{bmatrix} -15 & 0 & -6 \\ 9 & 0 & 3 \\ 2 & 2 & 2 \end{bmatrix}$$

4

3. When it comes to real numbers, we know that if xy = 0, then either x = 0 or y = 0 or both. One might believe that a similar idea applies to matricies, but one would be wrong. Prove that if the matrix product AB = 0 (by which we mean a matrix of appropriate dimensionality made up entirely of zeroes), then it is not necessarily true that either A = 0 or B = 0. Hint: in order to prove that something is not always true, simply identify one example where AB = 0,  $A, B \neq 0$ .

**Solution:** Generally speaking, it is easy to show that something is *not* necessarily true. All that is needed is a single counterexample! And in this case, there are infinitely many counterexamples. Here's one:

$$\mathbf{A} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$\mathbf{AB} = \begin{bmatrix} 1(1) + 1(-1) & 1(1) + 1(-1) \\ 1(-1) + 1(1) & 1(-1) + 1(1) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

4. Ordinary least squares regression is a common method for obtaining regression parameters relating a set of explanatory variables with a continuous outcome of interest. The vector  $\hat{\mathbf{b}}$  that contains the intercept and the regression slope is calculated by the equation:

$$\hat{\mathbf{b}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

If an explanatory variable is nominal (i.e. ordering does not matter) with more than two classes (e.g. {White, Black, Asian, Mixed, Other}), the variable must be modified to include in the regression model. A common technique known as **one-hot encoding** converts the column into a series of n-1 binary (0/1) columns where each column represents a single class and n is the total number of unique classes in the original column. Explain why this method converts the column into n-1 columns, rather than n columns, in terms of linear algebra.

**Solution:** In order to calculate  $\hat{\mathbf{b}}$ , we must be able to calculate  $(\mathbf{X}'\mathbf{X})^{-1}$ . And we can only invert  $\mathbf{X}'\mathbf{X}$  if the matrix is **nonsingular**. What could make a matrix singular? If at least one column is **linearly dependent** (i.e. its value can be produced by linear combinations of other columns in the matrix), then the matrix will not be **full rank**. A square matrix that is not full rank will produce a determinant of 0, which as you'll recall in the case of a  $2 \times 2$  matrix would require division by zero.

$$\mathbf{X}^{-1} = \frac{1}{0} \left[ \begin{array}{cc} x_{22} & -x_{12} \\ -x_{21} & x_{11} \end{array} \right]$$

So  $\mathbf{X}'\mathbf{X}$  must be full rank in order to invert it. How does this effect our one-hot encoding scheme? If we were to convert the expanatory variable into n binary variables, each of the n columns is a linear combination of the other n-1 binary variables. We could not invert this matrix and therefore OLS would not be possible.

- 5. Solve the following systems of equations for x, y, z, either via matrix inversion or substitution:
  - a. System #1

$$x + y + 2z = 2$$
$$3x - 2y + z = 1$$
$$y - z = 3$$

• Using matrix inversion:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 2 \\ 3 & -2 & 2 \\ 0 & 1 & -1 \end{bmatrix} \quad \mathbf{y} = [2, 1, 3]' \quad \mathbf{x} = [x, y, z]$$
$$\mathbf{A}\mathbf{x} = \mathbf{y}$$
$$\mathbf{A}^{-1}\mathbf{y} = \mathbf{x}$$

You can use (a lot) of Gauss-Jordan elimination to invert the matrix. Or I can just use R.

```
A \leftarrow matrix(c(1, 1, 2, 3, -2, 1, 0, 1, -1),
            nrow = 3,
            ncol = 3, byrow = TRUE)
        [,1] [,2] [,3]
## [1,]
          1
## [2,]
## [3,]
y \leftarrow c(2, 1, 3)
## [1] 2 1 3
solve(A)
              # inverse of A
        [,1] [,2] [,3]
## [1,] 0.1 0.3 0.5
## [2,] 0.3 -0.1 0.5
## [3,] 0.3 -0.1 -0.5
solve(A, y) # inverse of A times y = x
```

- ## [1] 2 2 -1
- Using systems of equations
  - 1.  $1 \times 1$  x third row added to second row and  $2 \times 1$  third row added to first row.

$$x + 3y = 8$$
$$3x - y = 4$$
$$y - z = 3$$

2. -3 x first row added to second row

$$x + 3y = 8$$
$$-10y = -20$$
$$y - z = 3$$

3. Solve for y and z

$$-10y = -20 \rightarrow y = 2$$
$$y - z = 3 \rightarrow z = -1$$

4. Substitute y into the first equation

$$x + 3(2) = 8 \rightarrow x = 2$$

$$x = 2, y = -1, z = 2$$

b. System #2

$$2x + 3y - z = -8$$
$$x + 2y - z = 2$$
$$-x - 4y + z = -6$$

• Using matrix inversion

```
A \leftarrow matrix(c(2, 3, -1, 1, 2, -1, -1, -4, 1),
           nrow = 3,
           ncol = 3, byrow = TRUE)
        [,1] [,2] [,3]
##
## [1,]
          2 3 -1
        1
## [2,]
             2 -1
## [3,] -1
y \leftarrow c(-8, 2, -6)
## [1] -8 2 -6
solve(A) # inverse of A
        [,1] [,2] [,3]
## [1,]
        1 -0.5 0.5
## [2,]
          0 -0.5 -0.5
## [3,]
          1 -2.5 -0.5
solve(A, y) # inverse of A times y = x
```

• Using systems of equations

## [1] -12 2 -10

1. Add third row to first and second rows

$$x - y = -14$$
$$-2y = -4$$
$$-x - 4y + z = -6$$

2. Solve for y

$$-2y = -4 \rightarrow y = 2$$

3. Substitution

$$x-2 = -14 \to x = -12$$
 
$$-(-12) - 4(2) + z = -6 \to z = -10$$

$$x = -12, y = 2, z = -10$$

c. System #3

$$x - y + 2z = 2$$
$$4x + y - 2z = 10$$
$$x + 3y + z = 0$$

• Using matrix inversion

```
A \leftarrow matrix(c(1, -1, 2, 4, 1, -2, 1, 3, 1),
            nrow = 3,
            ncol = 3, byrow = TRUE)
Α
        [,1] [,2] [,3]
## [1,]
          1 -1
## [2,]
## [3,]
y \leftarrow c(2, 10, 0)
## [1] 2 10 0
solve(A)
              # inverse of A
               [,1]
                           [,2]
## [1,] 0.2000000 0.20000000 1.387779e-17
## [2,] -0.1714286 -0.02857143 2.857143e-01
## [3,] 0.3142857 -0.11428571 1.428571e-01
```

solve(A, y) # inverse of A times y = x

- Using systems of equations
  - 1. Add row 1 to row 2

$$x - y + 2z = 2$$
$$5x = 12$$
$$x + 3y + z = 0$$

2. Solve for x

$$5x = 12 \rightarrow x = \frac{12}{5}$$

3. Plug in x = 2 and add row 1 x 3 to row 3

$$\frac{12}{5} - y + 2z = 2$$
$$4\left(\frac{12}{5}\right) + 7z = 6$$

4. Solve for z

$$4\left(\frac{12}{5}\right) + 7z = 6 \to z = -\frac{18}{35}$$

5. Solve for y

$$\frac{12}{5} - y + 2\left(-\frac{18}{35}\right) = 2 \to y = -\frac{22}{35}$$
$$x = \frac{12}{5}, y = -\frac{22}{35}, z = -\frac{18}{35}$$

6. Recall from Gill 4.8 that every  $p \times p$  matrix **X** has p scalar values,  $\lambda_i, i = 1, \ldots, p$  such that:

$$\mathbf{X}\mathbf{h}_i = \lambda_i \mathbf{h}_i$$

for some corresponding vector  $\mathbf{h}_i$ .  $\lambda_i$  is called an **eigenvalue** of  $\mathbf{X}$  and  $h_i$  is called an **eigenvector** of  $\mathbf{X}$ .

Now consider the matrix:

$$\mathbf{M} = \left[ \begin{array}{rrr} 3 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & -1 & -3 \end{array} \right]$$

along with the eigenvalues  $\lambda_1=-3$  and  $\lambda_2=2,$  and eigenvectors:

$$\mathbf{x}_1 = \begin{bmatrix} -1 \\ -1 \\ 5 \end{bmatrix} \qquad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

a. For each eigenvalue/vector pair i, show that  $\mathbf{M}\mathbf{x}_i = \lambda_i \mathbf{x}_i$ .

We can straightforwardly verify for  $\mathbf{x}_1$ :

$$\mathbf{M}\mathbf{x}_{1} = \begin{bmatrix} 3 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & -1 & -3 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \\ 5 \end{bmatrix}$$
$$= \begin{bmatrix} -3+1+5 \\ -1-1+5 \\ -1+1-5 \end{bmatrix}$$
$$= \begin{bmatrix} 3 \\ 3 \\ -15 \end{bmatrix} = -3\mathbf{x}_{1} = \lambda_{1}\mathbf{x}_{1}$$

and for  $\mathbf{x}_2$ :

$$\mathbf{M}\mathbf{x}_{1} = \begin{bmatrix} 3 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & -1 & -3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$
$$= \begin{bmatrix} 3 - 1 + 0 \\ 1 + 1 + 0 \\ 1 - 1 + 0 \end{bmatrix}$$
$$= \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix} = 2\mathbf{x}_{2} = \lambda_{2}\mathbf{x}_{2}$$

b. One way to calculate the eigenvalues of **A** is to find the values of  $\lambda$  that solve the equation:

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

where I is the identity matrix. Show that this fact holds for M given above.

Let's first write out  $\mathbf{M} - \lambda \mathbf{I}$ :

$$\mathbf{M} - \lambda \mathbf{I} = \begin{bmatrix} 3 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & -1 & -3 \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}$$
$$= \begin{bmatrix} 3 - \lambda & -1 & 1 \\ 1 & 1 - \lambda & 1 \\ 1 & -1 & -3 - \lambda \end{bmatrix}$$

Now, we can find the determinant of this new matrix and set it equal to 0. This is called the characteristic polynomial of M:

$$0 = (3 - \lambda)(1 - \lambda)(-3 - \lambda) + (-1)(1)(1) + (1)(1)(-1) - (1)(1 - \lambda)(1) - (-1)(1)(-3 - \lambda) - (3 - \lambda)(1)(-1)$$
  
$$0 = (3 - \lambda)(1 - \lambda)(-3 - \lambda) - 2 - (1 - \lambda) + (-3 - \lambda) + (3 - \lambda)$$

If we didn't already know the eigenvalues, we could solve for the roots of this polynomial to find them. In this case, we can simply plug in the given values  $\lambda = -3$  and  $\lambda = 2$  and verify that they solve the equation.