MATRICES AND LINEAR REGRESSION

How can we combine them?

RECAP OF LAST TIME:

$$\underset{(k\times n)}{\mathbf{X}}\underset{(n\times p)}{\mathbf{Y}} = \begin{bmatrix} \sum_{i=1}^{n} x_{1i}y_{i1} & \sum_{i=1}^{n} x_{1i}y_{i2} & \cdots & \sum_{i=1}^{n} x_{1i}y_{ip} \\ \sum_{i=1}^{n} x_{2i}y_{i1} & \sum_{i=1}^{n} x_{2i}y_{i2} & \cdots & \sum_{i=1}^{n} x_{2i}y_{ip} \\ \vdots & & \ddots & \vdots \\ \sum_{i=1}^{n} x_{ki}y_{i1} & \cdots & \cdots & \sum_{i=1}^{n} x_{ki}y_{ip} \end{bmatrix}$$

DISCLAIMER

a matrix. For example, a matrix that has three columns but only two columns of unique information is given by $\begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}$. This is also true for the matrix $\begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 2 \end{bmatrix}$, because the third column is just two times the second column and therefore has no new relational information to offer.

The book is sometimes thoroughly confused.

- You saw it with outer-product and cross-product.
- Now with rows and columns.

GOAL FOR TODAY

1. Priority.

- Learn how to get from Y = XB + $\vec{e} \rightarrow \hat{b} = (X'X)^{-1}X'y$
- This will take us through Inversions and Determinants
  - How to calculate them and what they mean.

- How to do it by hand & computer.

If you have time.

- A guide on eigenvalues and eigenvectors used to find PCA.

LINEAR MODEL GONE MATRIX.

- Definition of linear regression:
  - $Y = XB + \vec{e}$
  - What should X contain? (Design matrix)
  - What should Beta Contain? (How many rows should be in the beta matrix)
  - What is the errors? How many rows?
- Let's take an example.
  - Attitude Military ~ Current Budget + War
  - Attitude Military 7-point Likert scale. (Really like – Hate)
  - $X = \begin{matrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1.5 & 0 \end{matrix} \quad B = \begin{matrix} 2 \\ 0.8 \\ 3 \end{matrix} \quad XB = \begin{matrix} 2 + 0.8 + 3 * 1 = 5.8 \\ 2 + 2 * 0.8 + 3 * 1 = 6.6 \\ 2 + 1.5 * 0.8 + 3 * 0 = 3.2 \end{matrix} \quad XB + \vec{\epsilon} = \begin{matrix} 5.8 + e_1 \\ 6.6 + e_2 \\ 3.2 + e_3 \end{matrix}$

IN REAL LIFE!

- We don't have the beta values but want to find the beta values based on having X and Y while minimizing the errors.
- Y = XB + $\vec{e}$ (Isolate Beta)
- $a = b \ast c \qquad or \qquad \frac{a}{b} = c \qquad or \quad a \ast b^{-1} = c \qquad a \ast \frac{1}{b} = c$
- Unfortunately, matrix inversion is not the same as scalar inversion.
  - What is the inverse of a matrix?
  - First, we need to understand the determinant of a matrix.

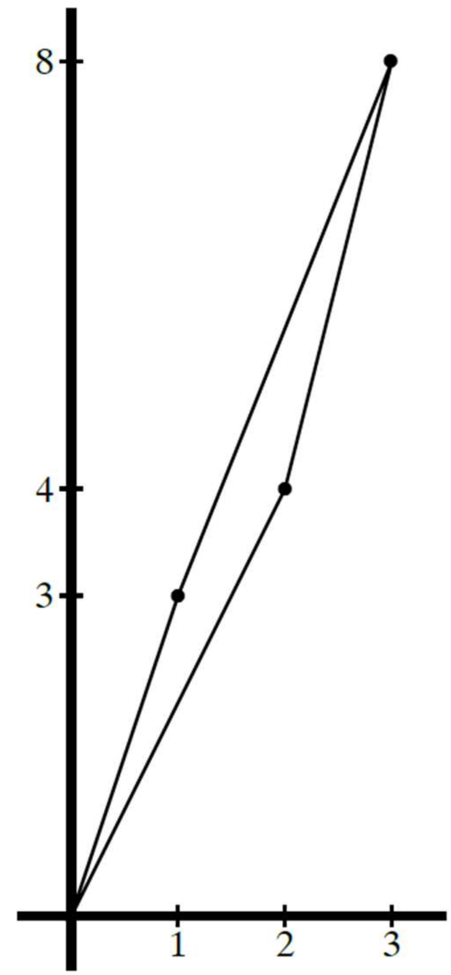
DETERMINANT

- We first need to understand determinants of a matrix.
- Notation
  - det(X) = |X|
- Interpretation:
  - Area of the parallelogram
  - More precisely manipulation of our space.

- Calculation 2x2:

$$\begin{vmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{vmatrix} = x_{11}x_{22} - x_{12}x_{21}$$

Spatial Representation of a Determinant



DETERMINANT 3X3

For a 3×3 matrix (3 rows and 3 columns):

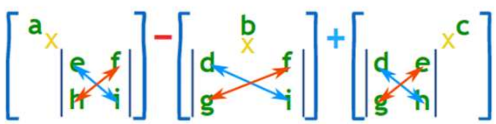
$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

The determinant is:

$$|A| = a(ei - fh) - b(di - fg) + c(dh - eg)$$

"The determinant of A equals ... etc"

It may look complicated, but **there is a pattern**:



To work out the determinant of a **3×3** matrix:

- Multiply **a** by the **determinant of the 2×2 matrix** that is **not in a**'s row or column.
- Likewise for **b**, and for **c**
- Sum them up, but remember the minus in front of the **b**

DETERMINANT 4X4

For 4×4 Matrices and Higher

The pattern continues for 4×4 matrices:

- **plus a** times the determinant of the matrix that is **not** in **a**'s row or column,
- **minus b** times the determinant of the matrix that is **not** in **b**'s row or column,
- **plus c** times the determinant of the matrix that is **not** in **c**'s row or column,
- **minus d** times the determinant of the matrix that is **not** in **d**'s row or column,

$$a \times \begin{vmatrix} f & g & h \\ j & k & l \\ n & o & p \end{vmatrix} - b \times \begin{vmatrix} e & g & h \\ i & k & l \\ m & o & p \end{vmatrix} + c \times \begin{vmatrix} e & f & h \\ i & j & l \\ m & n & p \end{vmatrix} - d \times \begin{vmatrix} e & f & g \\ i & j & k \\ m & n & o \end{vmatrix}$$

SIGN OF OPERATION.

- Leave out row 1 and col 1.
- Find determinant of the rest.
- Use the correct sign depend on the location.
- Uneven col/row (+)
- Even col/row (-)

$$\begin{bmatrix} + & - & + & - & + & - & + \\ - & + & - & + & - & + & - \\ + & - & + & - & + & - & + \\ - & + & - & + & - & + & - \\ + & - & + & - & + & - & + \\ - & + & - & + & - & + & - \\ + & - & + & - & + & - & + \end{bmatrix}$$

Onwards to **inversion..**

SF1

INVERSION

See Chris' video 4a for proof.

- Interpretation of matrix inversion.
  - $X^{-1}X = I \quad and\ XX^{-1} = I$
  - Both pre and post multiplication = Identity.
  - A matrix only has an inverse if $\det(X) \neq 0$

- 2x2 calculation.
  - Flip diagonal
  - Change the sign of the off diagonal.

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

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

Remember det(X) is a scalar so the inverse is simply

$$\det(X)^{-1} = \frac{1}{\det(X)}$$

Slide nummer 11

SF1 Sigurd Fyhn, 3/8/2022

INVERSION OF LARGER MATRICES

- **Gauss-Jordan elimination.**
  - An algorithm by which you can solve equations or find the inverse of matrices.

- Rules to follow:
  - Swap rows
  - Multiply a row
  - Subtract or add the multiple of a row to another row.

1. Swap the positions of two of the rows
2. Multiply one of the rows by a nonzero scalar.
3. Add or subtract the scalar multiple of one row to another row.

For an example of the first elementary row operation, swap the positions of the 1st and 3rd row.

$$\begin{pmatrix} 4 & 0 & -1 \\ 2 & -2 & 3 \\ 7 & 5 & 0 \end{pmatrix} \Rightarrow \begin{pmatrix} 7 & 5 & 0 \\ 2 & -2 & 3 \\ 4 & 0 & -1 \end{pmatrix}$$

For an example of the second elementary row operation, multiply the second row by 3.

$$\begin{pmatrix} 4 & 0 & -1 \\ 2 & -2 & 3 \\ 7 & 5 & 0 \end{pmatrix} \Rightarrow \begin{pmatrix} 4 & 0 & -1 \\ 6 & -6 & 9 \\ 7 & 5 & 0 \end{pmatrix}$$

For an example of the third elementary row operation, add twice the 1st row to the 2nd row.

$$\begin{pmatrix} 4 & 0 & -1 \\ 2 & -2 & 3 \\ 7 & 5 & 0 \end{pmatrix} \Rightarrow \begin{pmatrix} 4 & 0 & -1 \\ 10 & -2 & 1 \\ 7 & 5 & 0 \end{pmatrix}$$

Guide for Gauss-Jordan Elimination: https://online.stat.psu.edu/statprogram/reviews/matrix-algebra/gauss-jordan-elimination

GOOD PRINCIPLES TO FOLLOW.

1. Swap the rows so that all rows with all zero entries are on the bottom
2. Swap the rows so that the row with the largest, leftmost nonzero entry is on top.
3. Multiply the top row by a scalar so that top row's leading entry becomes 1.
4. Add/subtract multiples of the top row to the other rows so that all other entries in the column containing the top row's leading entry are all zero.
5. Repeat steps 2-4 for the next leftmost nonzero entry until all the leading entries are 1.
6. Swap the rows so that the leading entry of each nonzero row is to the right of the leading entry of the row above it.

INVERSE OF MATRICES USING GAUS-JORDAN ELIMINATION.

To obtain the inverse of a $n \times n$ matrix A :

1. Create the partitioned matrix $(A|I)$, where I is the identity matrix.
2. Perform Gauss-Jordan Elimination on the partitioned matrix with the objective of converting the first part of the matrix to reduced-row echelon form.
3. If done correctly, the resulting partitioned matrix will take the form $(I|A^{-1})$
4. Double-check your work by making sure that $AA^{-1} = I$.

The goal is to get the identity matrix on the left side by following the rules.

The matrix on the right will then be the inverse.

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 4 & 5 & 6 & 0 & 1 & 0 \\ 1 & 8 & 9 & 0 & 0 & 1 \end{array}\right]$$

EXAMPLE.

Let's perform this process on a 3×3 matrix:

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 4 & 5 & 6 & 0 & 1 & 0 \\ 1 & 8 & 9 & 0 & 0 & 1 \end{array}\right].$$

Now multiply the first row by -4, adding it to the second row, and multiply the first row by -1, adding it to the third row:

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

EXAMPLE

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

Multiply the second row by $\frac{1}{2}$, adding it to the first row, and simply add this same row to the third row:

$$\left[\begin{array}{ccc|ccc} 1 & \frac{1}{2} & 0 & -1 & \frac{1}{2} & 0 \\ 0 & -3 & -6 & -4 & 1 & 0 \\ 0 & 3 & 0 & -5 & 1 & 1 \end{array}\right].$$

EXAMPLE.

$$\left[\begin{array}{ccc|ccc} 1 & \frac{1}{2} & 0 & -1 & \frac{1}{2} & 0 \\ 0 & -3 & -6 & -4 & 1 & 0 \\ 0 & 3 & 0 & -5 & 1 & 1 \end{array}\right].$$

Multiply the third row by $-\frac{1}{6}$, adding it to the first row, and add the third row (un)multiplied to the second row:

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{6} & \frac{1}{3} & -\frac{1}{6} \\ 0 & 0 & -6 & -9 & 2 & 1 \\ 0 & 3 & 0 & -5 & 1 & 1 \end{array}\right].$$

EXAMPLE

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{6} & \frac{1}{3} & -\frac{1}{6} \\ 0 & 0 & -6 & -9 & 2 & 1 \\ 0 & 3 & 0 & -5 & 1 & 1 \end{array}\right].$$

Finally, just divide the second row by -6 and the third row by -3, and then switch their places:

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{6} & \frac{1}{3} & -\frac{1}{6} \\ 0 & 1 & 0 & -\frac{5}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 1 & \frac{3}{2} & -\frac{1}{3} & -\frac{1}{6} \end{array}\right],$$

NOW WE CAN FIND BETA.

- We now know how to transpose and find the inverse.

Therefore: $\hat{b} = (X'X)^{-1}X'y \rightarrow easy$.

- We could do everything by hand all the time.
    - But large matrices gets difficult, and we have a computer.

EXERCISES

- Check the Markdown file.

- It contains:
  - Which exercises you should do by hand.
  - Introduction to matrix manipulation in R.
  - Exercises on using matrices in R to do OLS.

EIGENVALUES, EIGENVECTORS & PCA

- Guide on Eigenvectors and Values usage to remove collinearity (PCA) :

https://builtin.com/data-science/step-step-explanation-principal-component-analysis

- Guide to PCA in R-studio:

https://www.datacamp.com/community/tutorials/pca-analysis-r