CSSS Math Camp Lecture 2

Matrix Algebra

Authored by: Laina Mercer, PhD

Erin Lipman and Jess Kunke

Department of Statistics

September 12, 2023

Lecture 2: Matrix Algebra

- Matrix Algebra
  - Definitions, notation
  - Matrix Arithmetic
  - Determinants - existence of an inverse
  - Linear equations
  - Least Squares and Regression with matrices

applications

Motivation

$$
\begin{array}{cc}
 & \text{3 colums} \\
\text{2 rows} & \begin{bmatrix} 4 & 3 & 7 \\ 8 & 2 & 6 \end{bmatrix}
\end{array}
$$

Matrix algebra provides concise notation and rules for manipulating matrices (arrays of numbers).

Matrix algebra will be important for computing linear regression estimates.

Motivation

Example dataframe:

	region	years	u5m	lower	upper
1	All	80-84	0.1691030	0.1573394	0.1815566
2	All	85-89	0.1603335	0.1490694	0.1722763
3	All	90-94	0.1208087	0.1079371	0.1349829
4	tanga	80-84	0.1810487	0.1369700	0.2354425
5	tanga	85-89	0.2230574	0.1677716	0.2902086

A *matrix* is an array of number is a rectangular form.
Examples:

$$A = \begin{bmatrix} 1 & 2 & 6 & 4 \\ 5 & 8 & 12 & 8 \\ 4 & 3 & 2 & 1 \end{bmatrix} B = \begin{bmatrix} 4 & 3 & 2 \\ 1 & 2 & 4 \end{bmatrix}$$

where A is a 3×4 matrix and B is a 2×3 matrix. Note: matrix dimensions are always listed as rows $\times$ columns.

rows x colums "roman colums"

Definitions & Notation

What is a matrix?

In mathematical notation, a matrix is written

$$X = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix}$$

Where x_{ij} is the value in the ith row and the jth column of matrix X.

A *vector* is a matrix that has n rows and 1 column (or 1 row and n columns).
Examples:

$$\begin{bmatrix} 1 & 2 & 6 & 4 \end{bmatrix} \text{ or } \begin{bmatrix} 4 \\ 5 \\ 1 \end{bmatrix}$$

A *vector* is a matrix that has n rows and 1 column (or 1 row and n columns).
Examples:

$$\begin{bmatrix} 1 & 2 & 6 & 4 \end{bmatrix} \text{ or } \begin{bmatrix} 4 \\ 5 \\ 1 \end{bmatrix}$$

A *square* matrix has the same number of rows and columns.
Example:

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

A *symmetric* matrix has elements such that $x_{ij} = x_{ji}$.
Example:

$$\begin{bmatrix} 1 & 4 & \boxed{5} \\ 4 & 2 & 3 \\ \boxed{5} & 3 & 7 \end{bmatrix}$$

A *symmetric* matrix has elements such that $x_{ij} = x_{ji}$.
Example:

$$\begin{bmatrix} 1 & 4 & 5 \\ 4 & 2 & 3 \\ 5 & 3 & 7 \end{bmatrix}$$

A symmetric matrix must also be a square matrix.

A *diagonal* matrix is a matrix that is zero everywhere except on the diagonal. Where the diagonal is defined as all elements for which the row number is equal to the column number $\{(1,1),(2,2),(3,3),...\}$.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 7 \end{bmatrix}$$

Definitions & Notation

Special Matrices

A *diagonal* matrix is a matrix that is zero everywhere except on the diagonal. Where the diagonal is defined as all elements for which the row number is equal to the column number $\{(1,1),(2,2),(3,3),...\}$.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 7 \end{bmatrix}$$

A special case of a diagonal matrix is the *identity* matrix. Its diagonal elements are all ones.

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

Clearly, the identity matrix (or any other diagonal matrix) is also symmetric.

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

Matrix Equality: Two matrices A, B are equal if and only if, for all elements, each $a_{ij} = b_{ij}$. (Note: this means they must have the same dimensions.)

Matrix Transpose: The *transpose* of a matrix is found by interchanging the corresponding rows and columns of a matrix. The first row becomes the first column, the second row becomes the second clump, etc. The dimensions are then switched and the element a_{ij} becomes the element a_{ji}. The transposed matrix is often denoted $\boxed{A^t}$ (or A').

A^T $\quad A^t$ $\quad A'$

Transponce of a symmetric matrix is itself

Matrix Equality: Two matrices A, B are equal if and only if, for all elements, each $a_{ij} = b_{ij}$. (Note: this means they must have the same dimensions.)

Matrix Transpose: The *transpose* of a matrix is found by interchanging the corresponding rows and columns of a matrix. The first row becomes the first column, the second row becomes the second clump, etc. The dimensions are then switched and the element a_{ij} becomes the element a_{ji}. The transposed matrix is often denoted A^t (or A').

$$A = \begin{bmatrix} 1 & 2 & 6 \\ 3 & 5 & 9 \end{bmatrix} \quad A^t = \begin{bmatrix} 1 & 3 \\ 2 & 5 \\ 6 & 9 \end{bmatrix}$$

2×3 $\quad$ 3×2

Two matrices can be added or subtracted only if their dimensions are the same (both rows and columns). The corresponding elements are then added or subtracted.

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \end{bmatrix}$$

Matrix Arithmetic

Addition & Subtraction

Two matrices can be added or subtracted only if their dimensions are the same (both rows and columns). The corresponding elements are then added or subtracted.

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{bmatrix}$$

Example:

$$\begin{bmatrix} 1 & 2 & \boxed{6} \\ 3 & 5 & 9 \end{bmatrix} - \begin{bmatrix} 1 & 3 & \boxed{8} \\ 6 & 9 & 6 \end{bmatrix} = \begin{bmatrix} 0 & -1 & \boxed{-2} \\ -3 & -4 & 3 \end{bmatrix}$$

number

To multiply a matrix by a $\boxed{\textit{scalar}}$ (a constant value), multiply each element by that number.

Example:

$$A = \begin{bmatrix} 1 & 3 & \boxed{8} \\ 6 & 9 & 6 \end{bmatrix} \quad 3A = \begin{bmatrix} 3 & 9 & \boxed{24} \\ 18 & 27 & 18 \end{bmatrix}$$

Matrix Arithmetic

Matrix Multiplication

Two matrices can be multiplied only if the number of columns of the first matrix equals the number of rows of the second matrix (the inside numbers). The resulting matrix has dimensions: number of rows of first matrix by number of columns of second matrix (the outside numbers).

$$A \times B$$

$$\underbrace{\boxed{r_1 \times c_1} \quad \boxed{r_2 \times c_2}}_{=}$$

ok

$$3 \times 2 \quad 2 \times 7 \quad \checkmark$$

$$3 \times 2 \quad 3 \times 2 \quad \times$$

$$\neq$$

Two matrices can be multiplied only if the number of columns of the first matrix equals the number of rows of the second matrix (the inside numbers). The resulting matrix has dimensions: number of rows of first matrix by number of columns of second matrix (the outside numbers).

Multipling a row vector by a column vector:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \end{bmatrix} \quad B = \begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \end{bmatrix}$$

A is (1×3); B is (3×1), AB is (1×1), i.e. a scalar.

$a_{11} b_{11} + a_{12} b_{21} + a_{13} b_{31}$

AB is 1×1

Two matrices can be multiplied only if the number of columns of the first matrix equals the number of rows of the second matrix (the inside numbers). The resulting matrix has dimensions: number of rows of first matrix by number of columns of second matrix (the outside numbers).

Multipling a row vector by a column vector:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \end{bmatrix} \quad B = \begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \end{bmatrix}$$

A is (1×3); B is (3×1), AB is (1×1), i.e. a scalar.

We multiply a row and column vector by taking the cross product of the numbers in A and B.

$$A \cdot B = a_{11} \cdot b_{11} + a_{12} \cdot b_{21} + a_{13} \cdot b_{31}$$

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \quad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}$$

A is (2×3); B is (3×2).
$A \cdot B$ is $(2 \times 3 \cdot 3 \times 2)$ results in 2×2.
$B \cdot A$ is $(3 \times 2 \cdot 2 \times 3)$ results in 3×3.

$$AB$$

$$\begin{bmatrix} \boxed{A_{r1} \times B_{c1}} & A_{r1} \times B_{c2} \\ A_{r2} \times B_{c1} & A_{r2} \times B_{c2} \end{bmatrix}$$

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \quad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}$$

A is (2×3); B is (3×2).
$A \cdot B$ is $(2 \times 3 \cdot 3 \times 2)$ results in 2×2.
$B \cdot A$ is $(3 \times 2 \cdot 2 \times 3)$ results in 3×3.

We find each element $(ab)_{ij}$ by summing the crossproducts of the ith row of A and the jth column of B.

$$A \cdot B = \begin{bmatrix} a_{11} \cdot b_{11} + a_{12} \cdot b_{21} + a_{13} \cdot b_{31} & a_{11} \cdot b_{12} + a_{12} \cdot b_{22} + a_{13} \cdot b_{32} \\ a_{21} \cdot b_{21} + a_{12} \cdot b_{21} + a_{23} \cdot b_{31} & a_{21} \cdot b_{12} + a_{22} \cdot b_{22} + a_{23} \cdot b_{32} \end{bmatrix}$$

$\underline{3} \times 3$ $\quad$ $3 \times \underline{2}$

Examples:

$$A = \begin{bmatrix} 1 & 3 & 8 \\ 6 & 9 & 6 \\ \boxed{2 \quad 1 \quad 3} \end{bmatrix} \quad B = \begin{bmatrix} 3 & 9 \\ 2 & 1 \\ 3 & 2 \end{bmatrix}$$

$\frac{AB}{3 \times 2}$

$$A \cdot B = \begin{bmatrix} 1\cdot3+3\cdot2+8\cdot3 & 1\cdot9+3\cdot1+8\cdot2 \\ 6\cdot3+9\cdot2+6\cdot3 & 6\cdot9+9\cdot1+6\cdot2 \\ 2\cdot3+1\cdot2+3\cdot3 & \boxed{2\cdot9+1\cdot1+3\cdot2} \end{bmatrix} = \begin{bmatrix} 33 & 28 \\ 54 & 75 \\ 17 & \boxed{25} \end{bmatrix}$$

$18 + 1 + 6 = 25$

Examples:

$$A = \begin{bmatrix} 1 & 3 & 8 \\ 6 & 9 & 6 \\ 2 & 1 & 3 \end{bmatrix} \quad B = \begin{bmatrix} 3 & 9 \\ 2 & 1 \\ 3 & 2 \end{bmatrix}$$

$$A \cdot B = \begin{bmatrix} 1 \cdot 3 + 3 \cdot 2 + 8 \cdot 3 & 1 \cdot 9 + 3 \cdot 1 + 8 \cdot 2 \\ 6 \cdot 3 + 9 \cdot 2 + 6 \cdot 3 & 6 \cdot 9 + 9 \cdot 1 + 6 \cdot 2 \\ 2 \cdot 3 + 1 \cdot 2 + 3 \cdot 3 & 2 \cdot 9 + 1 \cdot 1 + 3 \cdot 2 \end{bmatrix} = \begin{bmatrix} 33 & 28 \\ 54 & 75 \\ 17 & 25 \end{bmatrix}$$

Note: $A \cdot B$ is not necessarily equal to $B \cdot A$. For matrix multiplication, order matters. In this case $B \cdot A$ cannot be computed as the dimensions are not compatible ($3 \times 2 \cdot 3 \times 3$).

Matrix Arithmetic

Practice

$BA: (4 \times 1) \times (2 \times 4)$ $\neq$

What is the dimension of AB? Is it possible to compute BA? Find AB.

$$A = \underset{2 \times 4}{\begin{bmatrix} 3 & 1 & 4 & 1 \\ 5 & 9 & 2 & 6 \end{bmatrix}} \quad B = \underset{4 \times 1}{\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}}$$

$AB \quad (2 \times 4) \times (4 \times 1) = 2 \times 1$

$$AB = \begin{bmatrix} 3 \cdot 1 + \cancel{1 \cdot 0} + 4 \cdot 1 + \cancel{1 \cdot 0} \\ 5 \cdot 1 + \cancel{9 \cdot 0} + 2 \cdot 1 + \cancel{6 \cdot 0} \end{bmatrix} = \begin{bmatrix} 7 \\ 7 \end{bmatrix}$$

Matrix Inverse:

The inverse of a number is its reciprocal; a number multiplied by its inverse equals 1. $(4 \cdot 1/4 = 1)$ $\frac{\cancel{4}}{1} \times \frac{1}{\cancel{4}} = \frac{1}{1} = 1$

Matrix Inverse:

The inverse of a number is its reciprocal; a number multiplied by its inverse equals 1. $(4 \cdot 1/4 = 1)$

The inverse of a matrix A is the matrix A^{-1} that satisfies $A \cdot A^{-1} = I$. Where I is the identity matrix (ones along the diagonal and the rest are zeros). **Only square matrices can have inverses.**

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$AA^{-1} = I$$

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

Matrix Inverse:

The inverse of a number is its reciprocal; a number multiplied by its inverse equals 1. $(4 \cdot 1/4 = 1)$

The inverse of a matrix A is the matrix A^{-1} that satisfies $A \cdot A^{-1} = I$. Where I is the identity matrix (ones along the diagonal and the rest are zeros). **Only square matrices can have inverses.**

Remember that matrix multiplication is not just multiplying pairs of elements, so we can't just find the reciprocal of each element. So how do we find the inverse? How do we know if the inverse exists?

The *determinant* is a number that can be computed for any square matrix.

For a 2×2 matrix,

$$A = \left[\begin{array}{cc} a & b \\ c & d \end{array} \right]$$

the determinant is $D(A) = a \cdot d - b \cdot c$.

The *determinant* is a number that can be computed for any square matrix.

For a 2×2 matrix,

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

the determinant is $D(A) = a \cdot d - b \cdot c$.

If the determinant is zero, no inverse exists. If the determinant is nonzero then the inverse exists.

If no inverse exists the matrix is called *singular*.

$$A = \begin{bmatrix} 4 & 12 \\ 3 & 6 \end{bmatrix}$$

$D = 4 \cdot 6 - 12 \cdot 3 = -12$. Inverse exists.

Matrix Arithmetic

Determinant Example

$$A = \begin{bmatrix} 4 & 12 \\ 3 & 6 \end{bmatrix}$$

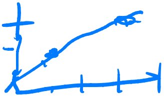
$D = 4 \cdot 6 - 12 \cdot 3 = -12$. Inverse exists.

$$A = \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}$$

$$2 \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$$

$D = 2 \cdot 2 - 4 \cdot 1 = 0$. Inverse does not exist,

side note



Matrix Arithmetic

Inverse Example

R

solve(A)
gives A^{-1}

det(A)
gives
determinant

Once we know the inverse exists, we can find it.
For a 2×2 matrix,

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$A^{-1} = \frac{1}{D(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

where again $D(A) = a \cdot d - b \cdot c$. Example:

$D(A) = 4 \cdot 6 - 3 \cdot 12 = -12$ $6/-12$ $-12/-12$

$$A = \begin{bmatrix} 4 & 12 \\ 3 & 6 \end{bmatrix}, \quad A^{-1} = \frac{1}{-12} \begin{bmatrix} 6 & -12 \\ -3 & 4 \end{bmatrix} = \begin{bmatrix} -1/2 & 1 \\ 1/4 & -1/3 \end{bmatrix}$$

$-3/-12$ $4/-12$

Note: Finding the inverse for higher dimensions involves more comp;located formulas and is usually solved by a math software.

Linear Equations

Let's go back to thinking about systems of two equations:

$$ax + by = g$$
$$cx + dy = f$$

equations of lines

Previously we solved this system by eliminating the y variable, solving for x, and then substituting back in for y.

No we can write this system in matrix notation:

A r, x z d

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad z = \begin{bmatrix} x \\ y \end{bmatrix}, \quad w = \begin{bmatrix} g \\ f \end{bmatrix}$$

ax + by

Solving our system of equations is the same as solving for z in the matrix equation:

$$A \cdot z = w$$

2x2 2x1 2x1

numbers

$$az = w$$

$$\frac{1}{a} a z = \frac{1}{a} w \rightarrow z = \frac{1}{a} w$$

Solving our system of equations is the same as solving for z in the matrix equation:

$$A \cdot z = w$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \times \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$$

So how do we solve for z?

$$\underbrace{A^{-1} A}\, z = A^{-1} w$$

$$I \cdot z = A^{-1} w$$

$$z = A^{-1} w$$

note
anything times identity is itself

Solving our system of equations is the same as solving for z in the matrix equation:

$$A \cdot z = w$$

So how do we solve for z? First, left-multiply the equation by A^{-1}:

$$A^{-1} \cdot A \cdot z = A^{-1} \cdot w$$

By definition $A^{-1} \cdot A = I$. Thus,

$$I \cdot z = A^{-1} \cdot w \text{ or } z = A^{-1} \cdot w.$$

So we can find $z = (x, y)$, the solution to our system, by finding $z = A^{-1} \cdot w$.

Linear Equations

Examples

$$2x + y = 1$$

$$4x + 3y = 8$$

ex no unique solution

$x + y = 1$

$2x + 2y = 2$

$A = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$ $\det(A) = 0$

$$\overset{A}{\begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}} \overset{Z}{\begin{bmatrix} x \\ y \end{bmatrix}} = \begin{bmatrix} 1 \\ 8 \end{bmatrix} \qquad Z = A^{-1} w$$

$$D(A) = 2 \cdot 3 - 1 \cdot 4 = 6 - 4 = 2$$

$$A^{-1} = \frac{1}{2} \begin{bmatrix} 3 & -1 \\ -4 & 2 \end{bmatrix} = \begin{bmatrix} 3/2 & -1/2 \\ -2 & 1 \end{bmatrix}$$

$D(A)$

$$\begin{bmatrix} x \\ y \end{bmatrix} = Z = \begin{bmatrix} 3/2 & -1/2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 8 \end{bmatrix}$$

$$2x + y = 1$$

$$4x + 3y = 8$$

$$A = \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}, \quad z = \begin{bmatrix} x \\ y \end{bmatrix}, \quad w = \begin{bmatrix} 1 \\ 8 \end{bmatrix}$$

$$2x + y = 1$$

$$4x + 3y = 8$$

$$A = \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}, \quad z = \begin{bmatrix} x \\ y \end{bmatrix}, \quad w = \begin{bmatrix} 1 \\ 8 \end{bmatrix}$$

$$A^{-1} = \frac{1}{2 \cdot 3 - 4 \cdot 1} \begin{bmatrix} 3 & -1 \\ -4 & 2 \end{bmatrix} = \begin{bmatrix} 3/2 & -1/2 \\ -2 & 1 \end{bmatrix}$$

Linear Equations

Examples

$$2x + y = 1$$

$$4x + 3y = 8$$

$$A = \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}, \quad z = \begin{bmatrix} x \\ y \end{bmatrix}, \quad w = \begin{bmatrix} 1 \\ 8 \end{bmatrix}$$

$$A^{-1} = \frac{1}{2 \cdot 3 - 4 \cdot 1} \begin{bmatrix} 3 & -1 \\ -4 & 2 \end{bmatrix} = \begin{bmatrix} 3/2 & -1/2 \\ -2 & 1 \end{bmatrix}$$

$$z = A^{-1} \cdot w = \begin{bmatrix} 3/2 & -1/2 \\ -2 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 8 \end{bmatrix} = \begin{bmatrix} 3/2 \cdot 1 + -1/2 \cdot 8 \\ -2 \cdot 1 + 1 \cdot 8 \end{bmatrix} = \begin{bmatrix} -5/2 \\ 6 \end{bmatrix}$$

Least Squares

Often we have much more data that pairs of points. We may have a survey where we asked n people the same p questions. We can put that data in a matrix of dimensions $n \times p$, where each row is a person and each column is one of the asked questions.

X_1	X_2	...	X_n	y
60	30		M	30K
61	23		F	70K
65	47		X	50K

Least Squares

Often we have much more data that pairs of points. We may have a survey where we asked n people the same p questions. We can put that data in a matrix of dimensions $n \times p$, where each row is a person and each column is one of the asked questions.

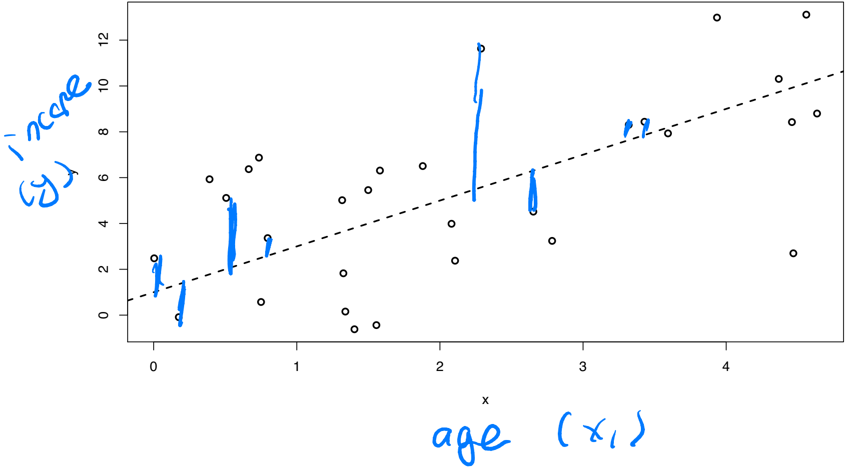
Before we saw how to put a line through two points ($y = mx + b$). What if we wanted to put a line through many points?

Least Squares

Example

$y = b + mx_1$

$y = \beta_0 + \beta_1 x_1$

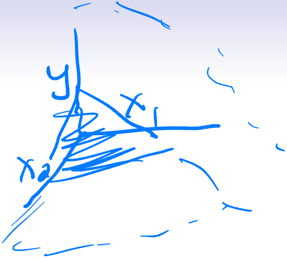


Least Squares

So how do we choose this line?

approx

So how do we choose this line?
We can write the equation:



$$y \approx \beta_0 + \beta_1 x_1 + ... + \beta_p x_p$$

where we have an intercept β_0 and then a slope β_i for each x_i (where $i = 1, .., p$). This equation has to describe the relationship as best it can for all n people we asked.

ex if $p=3$

$y \approx \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$

Least Squares

So how do we choose this line?
We can write the equation:

$$y \approx \beta_0 + \beta_1 x_1 + ... + \beta_p x_p$$

where we have an intercept β_0 and then a slope β_i for each x_i (where $i = 1, .., p$). This equation has to describe the relationship as best it can for all n people we asked. In matrix notation:

Least Squares

$y = \underset{(n \times (p+1))}{X} \underset{((p+1) \times 1)}{\beta}$, with y of size $(n \times 1)$

y	x_1	x_2	...	x_p
y_1	x_{11}	x_{21}		
y_2	x_{12}	x_{22}		

So how do we choose this line?
We can write the equation:

$$y \approx \beta_0 + \beta_1 x_1 + ... + \beta_p x_p$$

where we have an intercept β_0 and then a slope β_i for each x_i (where $i = 1, .., p$). This equation has to describe the relationship as best it can for all n people we asked. In matrix notation:

$$\begin{array}{ccccc} y & & X & & \beta \\ \begin{bmatrix} y_1 \\ ... \\ y_n \end{bmatrix} & \approx & \begin{bmatrix} 1 & x_{11} & ... & x_{1p} \\ 1 & ... & ... & ... \\ 1 & x_{n1} & ... & x_{np} \end{bmatrix} & \cdot & \begin{bmatrix} \beta_0 \\ ... \\ \beta_p \end{bmatrix} \end{array}$$

like $Az = w$

or $\boxed{y \approx X\beta.}$

ex 1 covariate 3 people

$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{21} \\ 1 & x_{12} & x_{22} \\ 1 & x_{13} & x_{23} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}$$

Least Squares

X is an $n \times (p+1)$

Example

The linear least squares procedure finds the line that minimizes the squared distance between the points and the line.

$X^T X$

$[(p+1) \times n] \times [n \times (p+1)]$

$= (p+1) \times (p+1)$

$$\boxed{\beta = \left(X^t \cdot X\right)^{-1} X^t y}$$

$$y = X\beta$$

$$X^T y = \underbrace{X^T X}\beta$$

$$(X^T X)^{-1} X^T y = (X^T X)^{-1} X^T X \beta$$

$$(X^T X)^{-1} X^T y = \beta$$

Least Squares

Example

The linear least squares procedure finds the line that minimizes the squared distance between the points and the line.

$$\beta = \left(X^t \cdot X\right)^{-1} X^t y$$

To see this note that y is $n \times 1$, X is $n \times (p+1)$, and β is $(p+1) \times 1$:

$$y \approx X\beta$$

Least Squares

Example

The linear least squares procedure finds the line that minimizes the squared distance between the points and the line.

$$\beta = \left(X^t \cdot X\right)^{-1} X^t y$$

To see this note that y is $n \times 1$, X is $n \times (p+1)$, and β is $(p+1) \times 1$:

$$\begin{aligned} y &\approx X\beta \\ X^t y &\approx X^t X\beta \end{aligned}$$

Least Squares

Example

The linear least squares procedure finds the line that minimizes the squared distance between the points and the line.

$$\beta = \left(X^t \cdot X\right)^{-1} X^t y$$

To see this note that y is $n \times 1$, X is $n \times (p+1)$, and β is $(p+1) \times 1$:

$$\begin{aligned} y &\approx X\beta \\ X^t y &\approx X^t X\beta \\ \left(X^t X\right)^{-1} X^t y &\approx \left(X^t X\right)^{-1} X^t X\beta \end{aligned}$$

Least Squares

Example

The linear least squares procedure finds the line that minimizes the squared distance between the points and the line.

$$\beta = \left(X^t \cdot X\right)^{-1} X^t y$$

To see this note that y is $n \times 1$, X is $n \times (p+1)$, and β is $(p+1) \times 1$:

$$\begin{aligned}
y &\approx X\beta \\
X^t y &\approx X^t X\beta \\
\left(X^t X\right)^{-1} X^t y &\approx \left(X^t X\right)^{-1} X^t X\beta \\
\left(X^t X\right)^{-1} X^t y &\approx I \cdot \beta
\end{aligned}$$

Least Squares

Example

The linear least squares procedure finds the line that minimizes the squared distance between the points and the line.

$$\beta = \left(X^t \cdot X\right)^{-1} X^t y$$

To see this note that y is $n \times 1$, X is $n \times (p+1)$, and β is $(p+1) \times 1$:

$$\begin{aligned}
y &\approx X\beta \\
X^t y &\approx X^t X\beta \\
\left(X^t X\right)^{-1} X^t y &\approx \left(X^t X\right)^{-1} X^t X\beta \\
\left(X^t X\right)^{-1} X^t y &\approx I \cdot \beta \\
\beta &\approx \left(X^t \cdot X\right)^{-1} X^t y
\end{aligned}$$