Mathematics for Social Scientists, II

Matrix Algebra

Howard M Thompson

University of Michigan – Flint

Table of contents

text & references

The textbook for this section of the course is

K. Namboodiri. *Matrix Algebra: An Introduction.* Sage Publications # 38, 1984.

The syllabus, CoursePack, and lecture notes for this course can be found at

http://homepages.umflint.edu/ hmthomps/icpsr/

Please read the ReadMe file.

text & references

If participants want to continue their study of matrix algebra, I advise taking a look at

G. Strang, *Introduction to Linear Algebra, Fourth Edition.* Wellesley Cambridge Press, 2009.

This is an introductory linear algebra text that might be used in a standard second year undergraduate course taken by physical science and engineering students in the United States. I prefer it because it covers Singular Value Decomposition.

Also see Strang's MIT Linear Algebra lectures (https://www.youtube.com/playlist?list=PL49CF3715CB9EF31D) and Khan Academy (https://www.khanacademy.org/math/linear-algebra).

definition and notation

A matrix is a rectangular array of numbers. The notation for a matrix A is as follows:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

This matrix has m rows and n columns. We say the matrix is m by n, written $m \times n$. The number in the ith row and jth column is a_{ij}. We also write $A = \left[a_{ij}\right]$.

See the examples in the CoursePack.

example

If

$$A = \left[a_{ij}\right] = \begin{bmatrix} 2 & -3 & 5 & -7 \\ 11 & -13 & 17 & -19 \\ 23 & -29 & 31 & -37 \end{bmatrix},$$

then $a_{23} = 17$.

some special matrices

We say two matrices are equal, $A = B$, if they have the same shape and are equal entry-wise:

$$A = B \text{ if } a_{ij} = b_{ij} \text{ for all } i \text{ and } j$$

- If $a_{ij} = 0$ for all i and j, we say $A = 0$ is "the" <u>zero</u> matrix.

some special matrices

We say two matrices are equal, $A = B$, if they have the same shape and are equal entry-wise:

$$A = B \text{ if } a_{ij} = b_{ij} \text{ for all } i \text{ and } j$$

- If $a_{ij} = 0$ for all i and j, we say $A = 0$ is "the" <u>zero</u> matrix.
- If $m = n$, we say A is <u>square</u>.

some special matrices

We say two matrices are equal, $A = B$, if they have the same shape and are equal entry-wise:

$$A = B \text{ if } a_{ij} = b_{ij} \text{ for all } i \text{ and } j$$

- If $a_{ij} = 0$ for all i and j, we say $A = 0$ is "the" zero matrix.
- If $m = n$, we say A is square.
- If A is square and $a_{ij} = 0$ when $i \neq j$, we say A is a diagonal matrix.

some special matrices

We say two matrices are equal, $A = B$, if they have the same shape and are equal entry-wise:

$$A = B \text{ if } a_{ij} = b_{ij} \text{ for all } i \text{ and } j$$

- If $a_{ij} = 0$ for all i and j, we say $A = 0$ is "the" <u>zero</u> matrix.
- If $m = n$, we say A is <u>square</u>.
- If A is square and $a_{ij} = 0$ when $i \neq j$, we say A is a <u>diagonal</u> matrix.
- If A is diagonal and $a_{ii} = 1$ for all i, we say A is "the" <u>identity</u> matrix.

addition

Addition is entry-wise.

Addition: If $A = \left[a_{ij}\right]$ and $B = \left[b_{ij}\right]$ are both $m \times n$, then $A + B = \left[a_{ij} + b_{ij}\right]$.

$$\begin{bmatrix} 2 & -3 & 5 \\ -7 & 11 & -13 \end{bmatrix} + \begin{bmatrix} 1 & 0 & -1 \\ 3 & 1 & 2 \end{bmatrix}$$

subtraction

Subtraction is entry-wise.

Subtraction: If $A = \left[a_{ij}\right]$ and $B = \left[b_{ij}\right]$ are both $m \times n$, then $A - B = \left[a_{ij} - b_{ij}\right]$.

$$\begin{bmatrix} 2 & -3 & 5 \\ -7 & 11 & -13 \end{bmatrix} - \begin{bmatrix} 1 & 0 & -1 \\ 3 & 1 & 2 \end{bmatrix}$$

scalar multiplication

Scalar multiplication is entry-wise.

Scalar Multiplication: If k is a number and $A = \left[a_{ij}\right]$, then $kA = \left[ka_{ij}\right]$.

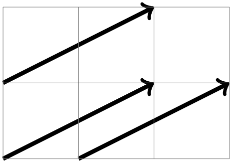
$$(-3)\begin{bmatrix} 2 & -3 & 5 \\ -7 & 11 & -13 \end{bmatrix}$$

vectors

A row vector is a matrix with only one row. A column vector (or just vector for short) is a matrix with only one column. For vectors of either type, we suppress the constant index on entries. We identity a 2×1 vector $\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ with the displacement from the origin to the point (a_1, a_2) in the Cartesian plane. Similarly, we identity a 3×1 vector $\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$ with the displacement from the origin to the point (a_1, a_2, a_3) in 3-space.

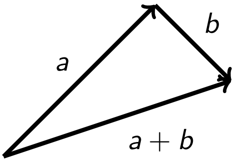
geometric vectors

All three of these arrows represent the vector $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$.



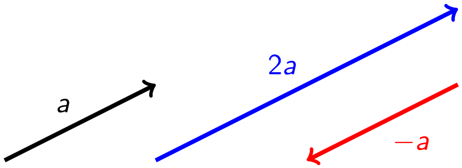
geometric vectors (cont.)

Geometrically, we sum vectors by placing them head to tail.



geometric vectors (cont.)

Geometrically, scalar multiplication is scaling.



the dot product

If $a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$ and $b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$, then

$$a \cdot b = \sum_{i=1}^{n} a_i b_i = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

The dot product is an example of an inner product. We define the product of the transpose of a, $a' = a^\top = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$, with the column vector b by the same expression.

$$a' b = a^\top b = a \cdot b = \sum_{i=1}^{n} a_i b_i = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

examples

$$\begin{bmatrix} 1 \\ -2 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

examples

$$\begin{bmatrix} 3 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} -4 \\ 3 \end{bmatrix}$$

examples

$$\begin{bmatrix} 3 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

examples

$$\begin{bmatrix} 0.6 \\ 0.8 \end{bmatrix} \cdot \begin{bmatrix} 0.6 \\ 0.8 \end{bmatrix}$$

orthogonality

We say the vectors a and b are orthogonal if $a \cdot b = 0$.

orthogonality

We say the vectors a and b are orthogonal if $a \cdot b = 0$.

$\begin{bmatrix} 3 \\ 4 \end{bmatrix}$ and $\begin{bmatrix} -4 \\ 3 \end{bmatrix}$ are orthogonal. Draw $\begin{bmatrix} 3 \\ 4 \end{bmatrix}$ and $\begin{bmatrix} -4 \\ 3 \end{bmatrix}$. What does it mean to be orthogonal?

length

The length of a is $|a| = \sqrt{a \cdot a}$.

$$\left| \begin{bmatrix} 3 \\ 4 \end{bmatrix} \right| = 5$$

Why would we call this quantity length?

length

The length of a is $|a| = \sqrt{a \cdot a}$.

$$\left| \begin{bmatrix} 3 \\ 4 \end{bmatrix} \right| = 5$$

Why would we call this quantity length?

We say a is a unit vector if $|a| = 1$. $\begin{bmatrix} 0.6 \\ 0.8 \end{bmatrix}$ is a unit vector.

matrix multiplication

If A is an $m \times p$ matrix and B is a $p \times n$ matrix, then the product $AB = C$ is the $m \times n$ matrix whose ijth entry is the product of the ith row of A and the jth column of B.

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}$$

Notice that this is the dot product of the transpose of the ith row of A and the jth column of B.

Place example here!

transpose

The transpose of a matrix $A = \left[a_{ij}\right]$ is the matrix $A' = A^\top = \left[a_{ji}\right]$ obtained by interchanging the rows and columns of A.

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}^\top$$

transpose

The transpose of a matrix $A = \left[a_{ij}\right]$ is the matrix $A' = A^\top = \left[a_{ji}\right]$ obtained by interchanging the rows and columns of A.

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}^\top$$

We say a matrix is symmetric if it is equal to its transpose.

transpose (cont.)

Let

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

Notice:

$$A^\top A = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} 35 & 44 \\ 44 & 56 \end{bmatrix}$$

is symmetric. Moreover, 35 is the sum of the squares of the entries of the first column of A. Similarly, 56 is the sum of the squares for the second column. 44 is the sum of the entry-wise products.

algebraic properties

The set of matrices equipped with the operations of addition, scalar multiplication and matrix multiplication satisfies a list of properties that is similar to those of the integers.

Addition on matrices of a particular shape has the following properties.

- $A + B = B + A$ (commutativity)

algebraic properties

The set of matrices equipped with the operations of addition, scalar multiplication and matrix multiplication satisfies a list of properties that is similar to those of the integers.

Addition on matrices of a particular shape has the following properties.

- $A + B = B + A$ (commutativity)
- $A + (B + C) = (A + B) + C$ (associativity)

algebraic properties

The set of matrices equipped with the operations of addition, scalar multiplication and matrix multiplication satisfies a list of properties that is similar to those of the integers.

Addition on matrices of a particular shape has the following properties.

- $A + B = B + A$ (commutativity)
- $A + (B + C) = (A + B) + C$ (associativity)
- $A + 0 = A$ (has an identity)

algebraic properties

The set of matrices equipped with the operations of addition, scalar multiplication and matrix multiplication satisfies a list of properties that is similar to those of the integers.

Addition on matrices of a particular shape has the following properties.

- $A + B = B + A$ (commutativity)
- $A + (B + C) = (A + B) + C$ (associativity)
- $A + 0 = A$ (has an identity)
- For every A, there is a $-A$ such that $A + (-A) = 0$ (has inverses)

algebraic properties (cont.)

Scalar multiplication on matrices of a particular shape has the following properties.

- $k_1(k_2A) = (k_1k_2)A$ (associativity)

algebraic properties (cont.)

Scalar multiplication on matrices of a particular shape has the following properties.

- $k_1(k_2A) = (k_1k_2)A$ (associativity)
- $k(A+B) = kA + kB$ and $(k_1 + k_2)A = k_1A + k_2A$ (distributes over addition)

algebraic properties (cont.)

Matrix multiplication on matrices of a particular shape has the following properties.

- $A(BC) = (AB)C$ (associativity)

algebraic properties (cont.)

Matrix multiplication on matrices of a particular shape has the following properties.

- $A(BC) = (AB)C$ (associativity)
- $k(AB) = (kA)B = A(kB)$ (commutes with scalar multiplication)

algebraic properties (cont.)

Matrix multiplication on matrices of a particular shape has the following properties.

- $A(BC) = (AB)C$ (associativity)
- $k(AB) = (kA)B = A(kB)$ (commutes with scalar multiplication)
- $IA = AI = A$ (has one-sided identities, the two Is differ when A is not square)

algebraic properties (cont.)

Matrix multiplication on matrices of a particular shape has the following properties.

- $A(BC) = (AB)C$ (associativity)
- $k(AB) = (kA)B = A(kB)$ (commutes with scalar multiplication)
- $IA = AI = A$ (has one-sided identities, the two Is differ when A is not square)
- $A(B + C) = AB + AC$ and $(A + B)C = AC + BC$ (distributes over addition)

algebraic properties (cont.)

Matrix multiplication on matrices of a particular shape has the following properties.

- $A(BC) = (AB)C$ (associativity)
- $k(AB) = (kA)B = A(kB)$ (commutes with scalar multiplication)
- $IA = AI = A$ (has one-sided identities, the two Is differ when A is not square)
- $A(B + C) = AB + AC$ and $(A + B)C = AC + BC$ (distributes over addition)

But, matrix multiplication is not commutative.

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

algebraic properties (cont.)

Matrix multiplication isn't cancellative either.

$$\begin{bmatrix} 4 & 6 \\ 6 & 9 \end{bmatrix} \begin{bmatrix} 6 & -9 \\ -4 & 6 \end{bmatrix} = \begin{bmatrix} 4 & 6 \\ 6 & 9 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

but

$$\begin{bmatrix} 6 & -9 \\ -4 & 6 \end{bmatrix} \neq \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

See the properties list in the CoursePack.

the Variance-Covariance Matrix

Recall: The variance of a single random variable is

$$\sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \mu)^2$$
$$= \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2\right]$$

the Variance-Covariance Matrix

Recall: The variance of a single random variable is

$$\sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \mu)^2$$
$$= \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2\right]$$

Now, suppose $X_1, X_2, X_3, \ldots, X_n$ are random variables (vectors with m entries) and let

$$A = \begin{bmatrix} | & | & & | \\ X_1 & X_2 & \cdots & X_n \\ | & | & & | \end{bmatrix}$$

Notice that the ijth entry of the symmetric matrix $A^\top A$ is $\sum X_i X_j$.

the Variance-Covariance Matrix (cont.)

Let U be the $m \times m$ matrix

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}$$

all of whose entries are 1. Notice that the ijth entry of the symmetric matrix $A^\top UA$ is $(\sum X_i)(\sum X_j)$.

the Variance-Covariance Matrix (cont.)

Let U be the $m \times m$ matrix

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}$$

all of whose entries are 1. Notice that the ijth entry of the symmetric matrix $A^\top UA$ is $(\sum X_i)(\sum X_j)$. Let

$$C = \frac{1}{n-1}\left(A^\top A - \frac{1}{n}A^\top UA\right)$$

C is called the Variance-Covariance Matrix. The ijth entry of C is the covariance of X_i and X_j. The sum of the diagonal of a square matrix is the trace of that matrix. trace(C) is the total variance of $X_1, X_2, X_3, \ldots, X_n$.

See the Variance-Covariance Matrix page in the CoursePack.

first steps in equation solving

We would like to solve matrix equations of the form $AX = B$. If a and b are numbers, how do we solve the equation $ax = b$ for x?

first steps in equation solving

We would like to solve matrix equations of the form $AX = B$. If a and b are numbers, how do we solve the equation $ax = b$ for x? Are there always solutions? Is there ever more than one solution?

first steps in equation solving

We would like to solve matrix equations of the form $AX = B$. If a and b are numbers, how do we solve the equation $ax = b$ for x? Are there always solutions? Is there ever more than one solution?

We need a notion for matrices that plays the role that is played by reciprocals in the number case.

The inverse of a square matrix A is a matrix A^{-1} such that $A^{-1}A = AA^{-1} = I$. If such a matrix A^{-1} exists, we say A is invertible (or nonsingular).

So, if A is invertible, $A^{-1}B$ is the solution to $AX = B$.

first example

Check that $\begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix}$ and $\begin{bmatrix} 7 & -2 \\ -3 & 1 \end{bmatrix}$ are inverses.

first example

Check that $\begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix}$ and $\begin{bmatrix} 7 & -2 \\ -3 & 1 \end{bmatrix}$ are inverses.

Let $A = \begin{bmatrix} 2 & 5 \\ 1 & 2 \end{bmatrix}$. If possible, find A^{-1}.

To do this, let $A^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and solve the pair of simultaneous equations determined by looking at the entries in the matrix equation

$$\begin{bmatrix} 2 & 5 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 2a + 5c & 2b + 5d \\ a + 2c & b + 2d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

first example (cont.)

$$\begin{bmatrix} 2a+5c & 2b+5d \\ a+2c & b+2d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Here are the systems of equations.

$$\begin{aligned} 2a+5c &= 1 \qquad & 2b+5d &= 0 \\ a+2c &= 0 \qquad & b+2d &= 1 \end{aligned}$$

Switch the equations in each system.

first example (cont.)

$$\begin{bmatrix} 2a+5c & 2b+5d \\ a+2c & b+2d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Here are the systems of equations.

$$\begin{aligned} 2a+5c &= 1 \qquad & 2b+5d &= 0 \\ a+2c &= 0 \qquad & b+2d &= 1 \end{aligned}$$

Switch the equations in each system.

$$\begin{aligned} a+2c &= 0 \qquad & b+2d &= 1 \\ 2a+5c &= 1 \qquad & 2b+5d &= 0 \end{aligned}$$

first example (cont.)

$$\begin{bmatrix} 2a+5c & 2b+5d \\ a+2c & b+2d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Here are the systems of equations.

$$\begin{aligned} 2a+5c &= 1 \qquad & 2b+5d &= 0 \\ a+2c &= 0 \qquad & b+2d &= 1 \end{aligned}$$

Switch the equations in each system.

$$\begin{aligned} a+2c &= 0 \qquad & b+2d &= 1 \\ 2a+5c &= 1 \qquad & 2b+5d &= 0 \end{aligned}$$

Subtract twice the first equation from the second in each system.

first example (cont.)

$$\begin{aligned} a + 2c &= 0 \qquad & b + 2d &= 1 \\ 2a + 5c &= 1 \qquad & 2b + 5d &= 0 \end{aligned}$$

Subtract twice the first equation from the second in each system.

first example (cont.)

$$\begin{aligned} a + 2c &= 0 \qquad & b + 2d &= 1 \\ 2a + 5c &= 1 \qquad & 2b + 5d &= 0 \end{aligned}$$

Subtract twice the first equation from the second in each system.

$$\begin{aligned} a + 2c &= 0 \qquad & b + 2d &= 1 \\ c &= 1 \qquad & d &= -2 \end{aligned}$$

Subtract twice the second equation from the first in each system.

first example (cont.)

$$\begin{aligned} a + 2c &= 0 \qquad & b + 2d &= 1 \\ 2a + 5c &= 1 \qquad & 2b + 5d &= 0 \end{aligned}$$

Subtract twice the first equation from the second in each system.

$$\begin{aligned} a + 2c &= 0 \qquad & b + 2d &= 1 \\ c &= 1 \qquad & d &= -2 \end{aligned}$$

Subtract twice the second equation from the first in each system.

$$\begin{aligned} a &= -2 \qquad & b &= 5 \\ c &= 1 \qquad & d &= -2 \end{aligned}$$

first example (cont.)

Check that $\begin{bmatrix} 2 & 5 \\ 1 & 2 \end{bmatrix}$ and $\begin{bmatrix} -2 & 5 \\ 1 & -2 \end{bmatrix}$ are inverses.

Here is the previous calculation using the notation of Gaussian Elimination.

$$\left[\begin{array}{cc|cc} 2 & 5 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{array}\right]$$

first example (cont.)

Check that $\begin{bmatrix} 2 & 5 \\ 1 & 2 \end{bmatrix}$ and $\begin{bmatrix} -2 & 5 \\ 1 & -2 \end{bmatrix}$ are inverses.

Here is the previous calculation using the notation of Gaussian Elimination.

$$\left[\begin{array}{cc|cc} 2 & 5 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_2} \left[\begin{array}{cc|cc} 1 & 2 & 0 & 1 \\ 2 & 5 & 1 & 0 \end{array}\right]$$

first example (cont.)

Check that $\begin{bmatrix} 2 & 5 \\ 1 & 2 \end{bmatrix}$ and $\begin{bmatrix} -2 & 5 \\ 1 & -2 \end{bmatrix}$ are inverses.

Here is the previous calculation using the notation of Gaussian Elimination.

$$\left[\begin{array}{cc|cc} 2 & 5 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_2} \left[\begin{array}{cc|cc} 1 & 2 & 0 & 1 \\ 2 & 5 & 1 & 0 \end{array}\right]$$

$$\xrightarrow{R_2 - 2R_1} \left[\begin{array}{cc|cc} 1 & 2 & 0 & 1 \\ 0 & 1 & 1 & -2 \end{array}\right]$$

first example (cont.)

Check that $\begin{bmatrix} 2 & 5 \\ 1 & 2 \end{bmatrix}$ and $\begin{bmatrix} -2 & 5 \\ 1 & -2 \end{bmatrix}$ are inverses.

Here is the previous calculation using the notation of Gaussian Elimination.

$$\left[\begin{array}{cc|cc} 2 & 5 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_2} \left[\begin{array}{cc|cc} 1 & 2 & 0 & 1 \\ 2 & 5 & 1 & 0 \end{array}\right]$$
$$\xrightarrow{R_2 - 2R_1} \left[\begin{array}{cc|cc} 1 & 2 & 0 & 1 \\ 0 & 1 & 1 & -2 \end{array}\right]$$
$$\xrightarrow{R_1 - 2R_2} \left[\begin{array}{cc|cc} 1 & 0 & -2 & 5 \\ 0 & 1 & 1 & -2 \end{array}\right]$$

finding inverses

To find the inverse of a $n \times n$ matrix A, we start with the augmented matrix $[A|I]$. We are allowed to do three kinds of elementary row operations.

finding inverses

To find the inverse of a $n \times n$ matrix A, we start with the augmented matrix $[A|I]$. We are allowed to do three kinds of elementary row operations.

- $R_i \leftrightarrow R_j$ (exchange row i and row j)

finding inverses

To find the inverse of a $n \times n$ matrix A, we start with the augmented matrix $[A|I]$. We are allowed to do three kinds of elementary row operations.

- $R_i \leftrightarrow R_j$ (exchange row i and row j)
- kR_i (multiply row i by a nonzero scalar k)

finding inverses

To find the inverse of a $n \times n$ matrix A, we start with the augmented matrix $[A|I]$. We are allowed to do three kinds of elementary row operations.

- $R_i \leftrightarrow R_j$ (exchange row i and row j)
- kR_i (multiply row i by a nonzero scalar k)
- $R_j + kR_i$ (add k times row i to row j)

finding inverses

To find the inverse of a $n \times n$ matrix A, we start with the augmented matrix $[A|I]$. We are allowed to do three kinds of elementary row operations.

- $R_i \leftrightarrow R_j$ (exchange row i and row j)
- kR_i (multiply row i by a nonzero scalar k)
- $R_j + kR_i$ (add k times row i to row j)

The goal is to transform the augmented matrix to $[I|A^{-1}]$ or produce a row that begins with n zeros. One of the two is always achievable. In the later case, A is singular.

other examples

- If possible, find the inverse of $\begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix}$

other examples

- If possible, find the inverse of $\begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix}$
- If possible, find the inverse of $\begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix}$

other examples

- If possible, find the inverse of $\begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix}$
- If possible, find the inverse of $\begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix}$
- If possible, find the inverse of $\begin{bmatrix} 0 & 1 & 2 \\ -1 & 3 & 1 \\ 2 & 0 & 4 \end{bmatrix}$

properties of inverses

- If all the entries in a row (respectively column) of A are zero, then A is not invertible.

properties of inverses

- If all the entries in a row (respectively column) of A are zero, then A is not invertible.
- If a row (respectively column) of A is a scalar multiple of another row (respectively column) of A, the A is not invertible.

properties of inverses

- If all the entries in a row (respectively column) of A are zero, then A is not invertible.
- If a row (respectively column) of A is a scalar multiple of another row (respectively column) of A, the A is not invertible.
- If $d_i \neq 0$ for all i,

$$D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} \quad D^{-1} = \begin{bmatrix} \frac{1}{d_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{d_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{d_n} \end{bmatrix}$$

properties of inverses

- If all the entries in a row (respectively column) of A are zero, then A is not invertible.
- If a row (respectively column) of A is a scalar multiple of another row (respectively column) of A, the A is not invertible.
- If $d_i \neq 0$ for all i,

$$D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} \quad D^{-1} = \begin{bmatrix} \frac{1}{d_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{d_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{d_n} \end{bmatrix}$$

- $(AB)^{-1} = B^{-1}A^{-1}$

properties of inverses

- If all the entries in a row (respectively column) of A are zero, then A is not invertible.
- If a row (respectively column) of A is a scalar multiple of another row (respectively column) of A, the A is not invertible.
- If $d_i \neq 0$ for all i,

$$D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} \quad D^{-1} = \begin{bmatrix} \frac{1}{d_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{d_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{d_n} \end{bmatrix}$$

- $(AB)^{-1} = B^{-1}A^{-1}$
- $(A^\top)^{-1} = (A^{-1})^\top$

properties of inverses

- If all the entries in a row (respectively column) of A are zero, then A is not invertible.
- If a row (respectively column) of A is a scalar multiple of another row (respectively column) of A, the A is not invertible.
- If $d_i \neq 0$ for all i,

$$D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} \quad D^{-1} = \begin{bmatrix} \frac{1}{d_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{d_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{d_n} \end{bmatrix}$$

- $(AB)^{-1} = B^{-1}A^{-1}$
- $(A^\top)^{-1} = (A^{-1})^\top$
- $(A^{-1})^{-1} = A$

reduced row echelon form

A matrix is said to be in reduced row echelon form if

A matrix is said to be in reduced row echelon form if

- The first nonzero entry in each row is one. We call these entries leading ones or pivots.

reduced row echelon form

A matrix is said to be in reduced row echelon form if

- The first nonzero entry in each row is one. We call these entries leading ones or pivots.
- If a column contains a leading one, the rest of the entries in that column are zero.

reduced row echelon form

A matrix is said to be in reduced row echelon form if

- The first nonzero entry in each row is one. We call these entries leading ones or pivots.
- If a column contains a leading one, the rest of the entries in that column are zero.
- If a row contains a leading one, the row above it also contains a leading one. And, the leading one in the row above is farther to the left.

reduced row echelon form (cont.)

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

is in reduced row echelon form for any choices of the entries in the $*$ positions.

We say two matrices are row equivalent if there exists a sequence of row operations that transforms one into the other. For each matrix, there is a unique row equivalent reduced row echelon form matrix.

rank

The rank of a matrix A, $r(A)$ or $\text{rank}(A)$, is the number of leading ones in its reduced row echelon form. We say a matrix is of full rank if its reduced row echelon form has a leading one in each row.

rank

The rank of a matrix A, $r(A)$ or $\text{rank}(A)$, is the number of leading ones in its reduced row echelon form. We say a matrix is of full rank if its reduced row echelon form has a leading one in each row.

A square matrix is invertible if and only if it is of full rank.

systems of simultaneous linear equations

To solve a system of simultaneous linear equations, or more generally, to solve a matrix equation

$$AX = B$$

we put the augmented matrix $[A|B]$ in reduced row echelon form.

systems of simultaneous linear equations

To solve a system of simultaneous linear equations, or more generally, to solve a matrix equation

$$AX = B$$

we put the augmented matrix $[A|B]$ in reduced row echelon form.

If the reduced row echelon form has a nonzero row with no nonzero entry to the left of the separator $|$, the system has no solutions and we say the system is overdetermined.

systems of simultaneous linear equations

To solve a system of simultaneous linear equations, or more generally, to solve a matrix equation

$$AX = B$$

we put the augmented matrix $[A|B]$ in reduced row echelon form.

If the reduced row echelon form has a nonzero row with no nonzero entry to the left of the separator $|$, the system has no solutions and we say the system is <u>overdetermined</u>.

If the system is not overdetermined and the reduced row echelon form has a column with no leading ones to the left of the separator $|$, the system has infinitely many solutions and we say the system is <u>underdetermined</u>. The variables corresponding to the columns that do not contain leading ones are called <u>parameters</u> or <u>free variables</u>.

the three dimensional case

Put some examples here!

linear combination

A linear combination of a collection vectors is a sum of scalar multiples of those vectors. The span, $\mathrm{Span}(a_1, a_2, \ldots, a_n)$, of a collection of vectors $\{a_1, a_2, \ldots, a_n\}$ is the collection of all linear combinations of those vectors.

linear combination

A <u>linear combination</u> of a collection vectors is a sum of scalar multiples of those vectors. The span, $\mathrm{Span}(a_1, a_2, \ldots, a_n)$, of a collection of vectors $\{a_1, a_2, \ldots, a_n\}$ is the collection of all linear combinations of those vectors.

For which numbers k is $\begin{bmatrix} 1 \\ 2 \\ 9 \end{bmatrix}$ a linear combination of $\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$, $\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$, and $\begin{bmatrix} 1 \\ 3 \\ k \end{bmatrix}$?

linear combination

A linear combination of a collection vectors is a sum of scalar multiples of those vectors. The span, $\mathrm{Span}(a_1, a_2, \ldots, a_n)$, of a collection of vectors $\{a_1, a_2, \ldots, a_n\}$ is the collection of all linear combinations of those vectors.

For which numbers k is $\begin{bmatrix} 1 \\ 2 \\ 9 \end{bmatrix}$ a linear combination of $\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$, $\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$, and $\begin{bmatrix} 1 \\ 3 \\ k \end{bmatrix}$?

We want to know which numbers k admit numbers a, b, and c such that

$$a \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} + b \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} + c \begin{bmatrix} 1 \\ 3 \\ k \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 9 \end{bmatrix}$$

linear combination (cont.)

Equivalently, we want to know which numbers k admit a solution to

$$\begin{bmatrix} 2 & 3 & 1 \\ 1 & 2 & 3 \\ 3 & 1 & k \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 9 \end{bmatrix}$$

linear combination (cont.)

Equivalently, we want to know which numbers k admit a solution to

$$\begin{bmatrix} 2 & 3 & 1 \\ 1 & 2 & 3 \\ 3 & 1 & k \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 9 \end{bmatrix}$$

$$\left[\begin{array}{ccc|c} 2 & 3 & 1 & 1 \\ 1 & 2 & 3 & 2 \\ 3 & 1 & k & 9 \end{array} \right]$$

linear combination (cont.)

Equivalently, we want to know which numbers k admit a solution to

$$\begin{bmatrix} 2 & 3 & 1 \\ 1 & 2 & 3 \\ 3 & 1 & k \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 9 \end{bmatrix}$$

$$\left[\begin{array}{ccc|c} 2 & 3 & 1 & 1 \\ 1 & 2 & 3 & 2 \\ 3 & 1 & k & 9 \end{array}\right] \xrightarrow{\text{(Partial) Gaussian Elimination}} \left[\begin{array}{ccr|r} 1 & 2 & 3 & 2 \\ 0 & 1 & 5 & 3 \\ 0 & 0 & k+16 & 18 \end{array}\right]$$

Which numbers k admit a solution?

linear dependence

We say a collection of vectors $\{a_1, a_2, \ldots, a_n\}$ is linearly dependent if one of the vectors can be written as a linear combination of the others.

linear dependence

We say a collection of vectors $\{a_1, a_2, \ldots, a_n\}$ is linearly dependent if one of the vectors can be written as a linear combination of the others.

Equivalently, $\{a_1, a_2, \ldots, a_n\}$ is linearly dependent if there exist scalars $k_1, k_2, \ldots, k_n$, not all zero, such that

$$k_1 a_1 + k_2 a_2 + \cdots + k_n a_n = 0$$

linear dependence

We say a collection of vectors $\{a_1, a_2, \ldots, a_n\}$ is linearly dependent if one of the vectors can be written as a linear combination of the others.

Equivalently, $\{a_1, a_2, \ldots, a_n\}$ is linearly dependent if there exist scalars $k_1, k_2, \ldots, k_n$, not all zero, such that

$$k_1 a_1 + k_2 a_2 + \cdots + k_n a_n = 0$$

Equivalently, this homogeneous system has a nonzero solution

$$\begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

has a nonzero solution.

linear independence

Equivalently, the rank of

$$\begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix}$$

is less than n.

linear independence

Equivalently, the rank of

$$\begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix}$$

is less than n.

We say a collection of vectors $\{a_1, a_2, \ldots, a_n\}$ is <u>linearly independent</u> if it is not linearly dependent.

linear independence (cont.)

For which numbers p is $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ p \end{bmatrix} \right\}$ linearly independent?

linear independence (cont.)

For which numbers p is $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ p \end{bmatrix} \right\}$ linearly independent?

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

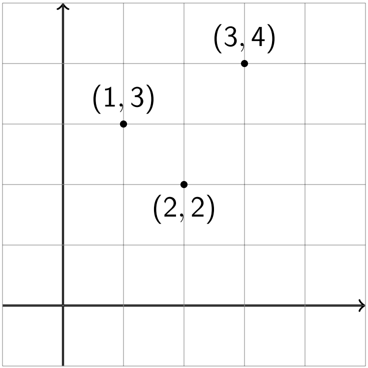
linear independence (cont.)

For which numbers p is $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ p \end{bmatrix} \right\}$ linearly independent?

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & -1 \\ 1 & 1 & p \end{bmatrix} \xrightarrow{\text{(Partial) Gaussian Elimination}} \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & p+3 \end{bmatrix}$$

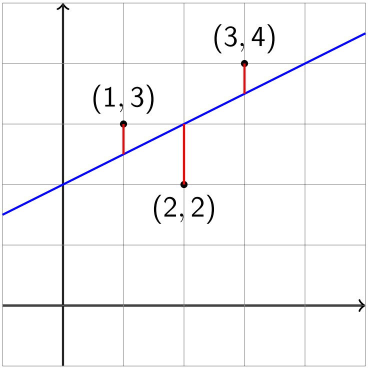
least squares example

We will find the best fit line for these three points.



least squares example

We will find the best fit line for these three points.



This line minimizes the sum of the squares of the vertical distances from the line to the points.

least squares example (cont.)

If the slope-intercept equation of this line is

$$y = mx + b$$

and the points were on the line we would have

$$\begin{aligned} 3 &= m + b \\ 2 &= 2m + b \\ 4 &= 3m + b \end{aligned}$$

least squares example (cont.)

If the slope-intercept equation of this line is

$$y = mx + b$$

and the points were on the line we would have

$$\begin{aligned} 3 &= m + b \\ 2 &= 2m + b \\ 4 &= 3m + b \end{aligned}$$

Equivalently, we would have a solution to

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

least squares example (cont.)

For now, we forget the origin of the problem.

We will use geometry to find the best approximate solution to

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

Let $\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix}$ and let $\begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix}$ be the error.

In order to make the geometric discussion less cumbersome, we identify vectors $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$ with the corresponding points (a, b, c) in 3-space.

least squares example (cont.)

Let P be the plane Span $\left(\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right)$. Note that the point (y_1, y_2, y_3) is (a general point) on P. We will check that for the best approximate solution, (y_1, y_2, y_3) is the point on P closest to $(3, 2, 4)$ in 3-space. This will complete our translation of the original question into a geometric question.

least squares example (cont.)

Let P be the plane Span $\left(\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right)$. Note that the point (y_1, y_2, y_3) is (a general point) on P. We will check that for the best approximate solution, (y_1, y_2, y_3) is the point on P closest to $(3, 2, 4)$ in 3-space. This will complete our translation of the original question into a geometric question.

In the original problem, the sum of the squares of the vertical distances from the line to the given points is

$$e_1^2 + e_2^2 + e_3^2 = \left| \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} \right|^2$$

least squares example (cont.)

Let P be the plane Span $\left(\begin{bmatrix}1\\2\\3\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix}\right)$. Note that the point (y_1, y_2, y_3) is (a general point) on P. We will check that for the best approximate solution, (y_1, y_2, y_3) is the point on P closest to $(3, 2, 4)$ in 3-space. This will complete our translation of the original question into a geometric question.

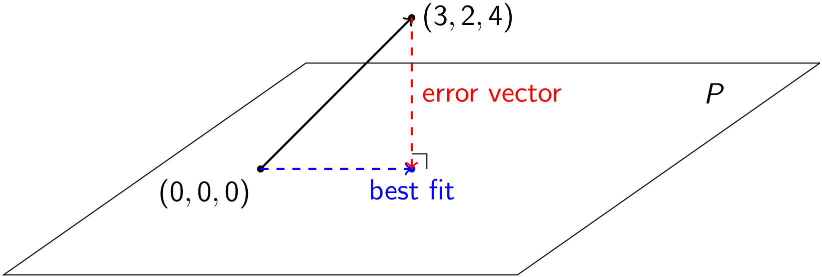
In the original problem, the sum of the squares of the vertical distances from the line to the given points is

$$e_1^2 + e_2^2 + e_3^2 = \left|\left|\begin{bmatrix}e_1\\e_2\\e_3\end{bmatrix}\right|\right|^2$$

So, we want to find the point (y_1, y_2, y_3) in the plane P closest to the point $(3, 2, 4)$.

least squares example (cont.)

If (y_1, y_2, y_3) is the closest point on the plane P to the point $(3, 2, 4)$, then the line segment from $(3, 2, 4)$ to (y_1, y_2, y_3) is perpendicular to the plane P.



least squares example (cont.)

Now, we rewrite this statement in terms of matrix algebra.

least squares example (cont.)

Now, we rewrite this statement in terms of matrix algebra.

$$\begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix} \text{ is orthogonal to every vector lying along } P.$$

least squares example (cont.)

Now, we rewrite this statement in terms of matrix algebra.

$$\begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix}$$ is orthogonal to every vector lying along P.

In particular, it is orthogonal to both $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

least squares example (cont.)

Now, we rewrite this statement in terms of matrix algebra.

$$\begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix} \text{ is orthogonal to every vector lying along } P.$$

In particular, it is orthogonal to both $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

We can package this in a single matrix equation

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \left(\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix} \right) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

least squares example (cont.)

We rewrite this as

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix}$$

least squares example (cont.)

We rewrite this as

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix}$$

Recalling that the vector $\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix}$, we get

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

least squares example (cont.)

We rewrite this as

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix}$$

Recalling that the vector $\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix}$, we get

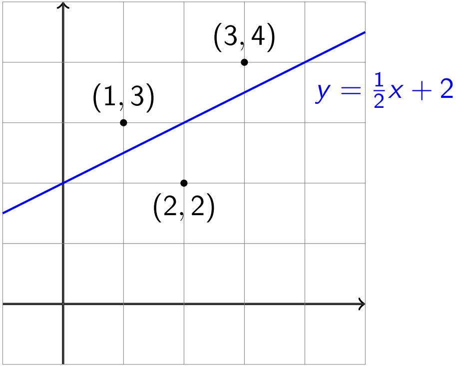
$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix}$$

or

$$\begin{bmatrix} 14 & 6 \\ 6 & 3 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 19 \\ 9 \end{bmatrix}$$

least squares example (cont.)

$$\left[\begin{array}{cc|c} 14 & 6 & 19 \\ 6 & 3 & 9 \end{array}\right] \xrightarrow{\text{Gaussian Elimination}} \left[\begin{array}{cc|c} 1 & 0 & \frac{1}{2} \\ 0 & 1 & 2 \end{array}\right]$$



determinants

More generally, solutions to $A^\top AX = A^\top B$ are best fit solutions to $AX = B$. In particular, when $A^\top A$ is invertible, $X = (A^\top A)^{-1} A^\top B$.

determinants

More generally, solutions to $A^\top AX = A^\top B$ are best fit solutions to $AX = B$. In particular, when $A^\top A$ is invertible, $X = (A^\top A)^{-1}A^\top B$.

If A is a square matrix, we define $\det(A)$, the determinant of A as follows.

determinants

More generally, solutions to $A^\top AX = A^\top B$ are best fit solutions to $AX = B$. In particular, when $A^\top A$ is invertible, $X = (A^\top A)^{-1}A^\top B$.

If A is a square matrix, we define $\det(A)$, the determinant of A as follows.

- If A is upper-triangular, then

$$\det(A) = \det\left(\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}\right) = a_{11}a_{22}\cdots a_{nn}$$

determinants

More generally, solutions to $A^\top AX = A^\top B$ are best fit solutions to $AX = B$. In particular, when $A^\top A$ is invertible, $X = (A^\top A)^{-1}A^\top B$.

If A is a square matrix, we define $\det(A)$, the determinant of A as follows.

- If A is upper-triangular, then

$$\det(A) = \det\left(\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}\right) = a_{11}a_{22}\cdots a_{nn}$$

- If $A \xrightarrow{R_i \leftrightarrow R_j} B$, then $\det(A) = -\det(B)$.

determinants

More generally, solutions to $A^\top AX = A^\top B$ are best fit solutions to $AX = B$. In particular, when $A^\top A$ is invertible, $X = (A^\top A)^{-1}A^\top B$.

If A is a square matrix, we define $\det(A)$, the determinant of A as follows.

- If A is upper-triangular, then

$$\det(A) = \det\left(\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}\right) = a_{11}a_{22}\cdots a_{nn}$$

- If $A \xrightarrow{R_i \leftrightarrow R_j} B$, then $\det(A) = -\det(B)$.
- If $A \xrightarrow{kR_i} B$, then $\det(A) = \frac{1}{k}\det(B)$.

determinants

More generally, solutions to $A^\top AX = A^\top B$ are best fit solutions to $AX = B$. In particular, when $A^\top A$ is invertible, $X = (A^\top A)^{-1}A^\top B$.

If A is a square matrix, we define $\det(A)$, the determinant of A as follows.

- If A is upper-triangular, then

$$\det(A) = \det\left(\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}\right) = a_{11}a_{22}\cdots a_{nn}$$

- If $A \xrightarrow{R_i \leftrightarrow R_j} B$, then $\det(A) = -\det(B)$.
- If $A \xrightarrow{kR_i} B$, then $\det(A) = \frac{1}{k}\det(B)$.
- If $A \xrightarrow{R_i + kR_j} B$, then $\det(A) = \det(B)$.

determinants (cont.)

In particular, $\det([a_{11}]) = a_{11}$ and

$$\det\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\right) = a_{11}a_{22} - a_{12}a_{21}$$

Find the determinant of

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 2 & -2 & 0 \\ 4 & 4 & -4 \end{bmatrix}$$

determinants (cont.)

Determinants have the following properties.

determinants (cont.)

Determinants have the following properties.

- $\det(A^\top) = \det(A)$

determinants (cont.)

Determinants have the following properties.

- $\det(A^\top) = \det(A)$
- A is invertible if and only if $\det(A) \neq 0$.

determinants (cont.)

Determinants have the following properties.

- $\det(A^\top) = \det(A)$
- A is invertible if and only if $\det(A) \neq 0$.
- If A is invertible, then $\det(A^{-1}) = \frac{1}{\det(A)}$.

determinants (cont.)

Determinants have the following properties.

- $\det(A^\top) = \det(A)$
- A is invertible if and only if $\det(A) \neq 0$.
- If A is invertible, then $\det(A^{-1}) = \frac{1}{\det(A)}$.
- If A is $n \times n$, then $\det(kA) = k^n \det(A)$.

motivating example

Consider the following products.

$$\begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 8 \\ 2 \end{bmatrix}$$

motivating example

Consider the following products.

$$\begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 8 \\ 2 \end{bmatrix}$$

$$\begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$$

motivating example

Consider the following products.

$$\begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 8 \\ 2 \end{bmatrix}$$

$$\begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$$

$$\begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 18 \\ 9 \end{bmatrix} = 9 \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

motivating example

Consider the following products.

$$\begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 8 \\ 2 \end{bmatrix}$$

$$\begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$$

$$\begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 18 \\ 9 \end{bmatrix} = 9 \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} -4 \\ 8 \end{bmatrix} = 4 \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

motivating example

Consider the following products.

$$\begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 8 \\ 2 \end{bmatrix}$$

$$\begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$$

$$\begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 18 \\ 9 \end{bmatrix} = 9 \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} -4 \\ 8 \end{bmatrix} = 4 \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

For the last two vectors, the 2×2 matrix acts by scalar multiplication.

definition

If A is a (necessarily square) matrix, we say a nonzero vector $x \neq 0$ is an eigenvector of A with eigenvalue λ if

$$Ax = \lambda x$$

(Note the the number λ can be zero.)

finding eigenvalues and eigenvectors

If x is an eigenvector of A with eigenvalue λ, we have

$$\begin{aligned} Ax &= \lambda x \\ Ax - \lambda x &= 0 \\ Ax - \lambda I x &= 0 \\ (A - \lambda I)x &= 0 \end{aligned}$$

Since $x \neq 0$, the matrix $A - \lambda I$ is not invertible. So, $\det(A - \lambda I) = 0$ if and only if λ is an eigenvalue of A.

finding eigenvalues and eigenvectors

If x is an eigenvector of A with eigenvalue λ, we have

$$\begin{aligned} Ax &= \lambda x \\ Ax - \lambda x &= 0 \\ Ax - \lambda I x &= 0 \\ (A - \lambda I)x &= 0 \end{aligned}$$

Since $x \neq 0$, the matrix $A - \lambda I$ is not invertible. So, $\det(A - \lambda I) = 0$ if and only if λ is an eigenvalue of A.

If A is $n \times n$ and λ is a variable, then $\det(\lambda I - A)$ is a polynomial of degree n called the <u>characteristic polynomial</u> of A.

finding eigenvalues and eigenvectors

If x is an eigenvector of A with eigenvalue λ, we have

$$\begin{aligned} Ax &= \lambda x \\ Ax - \lambda x &= 0 \\ Ax - \lambda I x &= 0 \\ (A - \lambda I)x &= 0 \end{aligned}$$

Since $x \neq 0$, the matrix $A - \lambda I$ is not invertible. So, $\det(A - \lambda I) = 0$ if and only if λ is an eigenvalue of A.

If A is $n \times n$ and λ is a variable, then $\det(\lambda I - A)$ is a polynomial of degree n called the characteristic polynomial of A.

For each root of $\det(A - \lambda I)$, we solve $(A - \lambda I)x = 0$ to find a corresponding eigenvector.

first example

Let's find all the eigenvalues of $\begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix}$.

first example

Let's find all the eigenvalues of $\begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix}$.

$$\begin{aligned} 0 &= \det\left(\begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) \\ &= \det\left(\begin{bmatrix} 8-\lambda & 2 \\ 2 & 5-\lambda \end{bmatrix}\right) \\ &= (8-\lambda)(5-\lambda) - 2\cdot 2 \\ &= \lambda^2 - 13\lambda + 36 \\ &= (\lambda - 9)(\lambda - 4) \end{aligned}$$

first example

Let's find all the eigenvalues of $\begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix}$.

$$\begin{aligned}
0 &= \det\left(\begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) \\
&= \det\left(\begin{bmatrix} 8-\lambda & 2 \\ 2 & 5-\lambda \end{bmatrix}\right) \\
&= (8-\lambda)(5-\lambda) - 2 \cdot 2 \\
&= \lambda^2 - 13\lambda + 36 \\
&= (\lambda - 9)(\lambda - 4)
\end{aligned}$$

The eigenvalues are 9 and 4.

first example (cont.)

Once we know the eigenvalues of $\begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix}$, we can find corresponding eigenvectors.

first example (cont.)

Once we know the eigenvalues of $\begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix}$, we can find corresponding eigenvectors.

$\underline{\lambda = 9}$

$$\left[\begin{array}{rr|r} -1 & 2 & 0 \\ 2 & -4 & 0 \end{array}\right] \xrightarrow{\text{Gaussian Elimination}} \left[\begin{array}{rr|r} 1 & -2 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

Find an eigenvector for the eigenvalue 9.

first example (cont.)

Once we know the eigenvalues of $\begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix}$, we can find corresponding eigenvectors.

$\underline{\lambda = 9}$

$$\left[\begin{array}{cc|c} -1 & 2 & 0 \\ 2 & -4 & 0 \end{array}\right] \xrightarrow{\text{Gaussian Elimination}} \left[\begin{array}{cc|c} 1 & -2 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

Find an eigenvector for the eigenvalue 9.

$\underline{\lambda = 4}$

$$\left[\begin{array}{cc|c} 4 & 2 & 0 \\ 2 & 1 & 0 \end{array}\right] \xrightarrow{\text{Gaussian Elimination}} \left[\begin{array}{cc|c} 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{array}\right]$$

Find an eigenvector for the eigenvalue 4.

a few properties

- If x is an eigenvector of A with eigenvalue λ, so is any (nonzero) scalar multiple of x.

a few properties

- If x is an eigenvector of A with eigenvalue λ, so is any (nonzero) scalar multiple of x.
- A collection of eigenvectors of A (corresponding to differing eigenvalues) is linearly independent.

a few properties

- If x is an eigenvector of A with eigenvalue λ, so is any (nonzero) scalar multiple of x.
- A collection of eigenvectors of A (corresponding to differing eigenvalues) is linearly independent.
- The product of the eigenvalues (with multiplicities) of A is $\det(A)$.

a few properties

- If x is an eigenvector of A with eigenvalue λ, so is any (nonzero) scalar multiple of x.
- A collection of eigenvectors of A (corresponding to differing eigenvalues) is linearly independent.
- The product of the eigenvalues (with multiplicities) of A is $\det(A)$.
- The sum of the eigenvalues (with multiplicities) of A is $\mathrm{trace}(A)$, the sum of the diagonal entries of A.

a few properties

- If x is an eigenvector of A with eigenvalue λ, so is any (nonzero) scalar multiple of x.
- A collection of eigenvectors of A (corresponding to differing eigenvalues) is linearly independent.
- The product of the eigenvalues (with multiplicities) of A is $\det(A)$.
- The sum of the eigenvalues (with multiplicities) of A is $\operatorname{trace}(A)$, the sum of the diagonal entries of A.
- We say an $n \times n$ matrix A is diagonalizable if it admits a collection of n linearly independent eigenvectors.

a few properties

- Symmetric matrices are diagonalizable.

a few properties

- Symmetric matrices are diagonalizable.
- A collection of eigenvectors of of a symmetric matrix A (corresponding to differing eigenvalues) are pairwise orthogonal.

a few properties

- Symmetric matrices are diagonalizable.
- A collection of eigenvectors of of a symmetric matrix A (corresponding to differing eigenvalues) are pairwise orthogonal.
- Thus, symmetric matrices always admit a full set of orthogonal unit eigenvectors.

a few properties

- Symmetric matrices are diagonalizable.
- A collection of eigenvectors of of a symmetric matrix A (corresponding to differing eigenvalues) are pairwise orthogonal.
- Thus, symmetric matrices always admit a full set of orthogonal unit eigenvectors.
- If A is a Variance-Covariance Matrix, the total variance is the sum of the eigenvalues (with multiplicities).

first example (cont.)

Let

$$P = \begin{bmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}$$

The columns of P are unit eigenvectors of $\begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix}$.

first example (cont.)

Let

$$P = \begin{bmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}$$

The columns of P are unit eigenvectors of $\begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix}$.

Since the columns of P are orthogonal unit vectors, $P^{-1} = P^{\top}$. Such matrices are called <u>orthogonal</u> matrices.

first example (cont.)

Let

$$P = \begin{bmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}$$

The columns of P are unit eigenvectors of $\begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix}$.

Since the columns of P are orthogonal unit vectors, $P^{-1} = P^{\top}$. Such matrices are called orthogonal matrices.

Let

$$\Lambda = \begin{bmatrix} 9 & 0 \\ 0 & 4 \end{bmatrix}$$

Compute $P\Lambda P^{-1}$.

first example (cont.)

Notice that if $\begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix}$ where a Variance-Covariance Matrix for two random variables X_1 and X_2, then the linear combinations $Y_1 = \frac{2}{\sqrt{5}}X_1 + \frac{1}{\sqrt{5}}X_2$ and $Y_2 = -\frac{1}{\sqrt{5}}X_1 + \frac{2}{\sqrt{5}}X_2$ are new random variables such that the Variance-Covariance Matrix of Y_1 and Y_2 is

$$\begin{bmatrix} 9 & 0 \\ 0 & 4 \end{bmatrix}$$

In particular, Y_1 and Y_2 are statistically independent. Y_1 accounts for $\frac{9}{13}$ (about 69.2%) of the total variance.