M.S. Math Bootcamp: Linear Algebra

Noah Kochanski

Department of Statistics
University of Michigan

August 19, 2024

Reminder: Resource Access

You can access the slides in the orientation materials folder shared with you.

Outline

Acknowledgements

This material was adapted from the following references

- Treil's *Linear Algebra Done Wrong*, 2017
- Strang's *Introduction to Linear Algebra*, 3rd ed.
- Golub & Van Loan's *Matrix Computations*, 4th ed.
- Previous iterations of this bootcamp offered by PhD students in the department.

Please don't distribute these slides.

Products and Conventions

Conventions - Vectors

- The most common vector space you'll work with is $\mathbb{R}^n$

Conventions - Vectors

- The most common vector space you'll work with is $\mathbb{R}^n$.
- By convention, a vector $\vec{x} \in \mathbb{R}^n$ is considered to be oriented as a column vector:

$$\vec{x} = \begin{bmatrix} | \\ \vec{x} \\ | \end{bmatrix}$$

Conventions - Vectors

- The most common vector space you'll work with is $\mathbb{R}^n$.
- By convention, a vector $\vec{x} \in \mathbb{R}^n$ is considered to be oriented as a column vector:

$$\vec{x} = \begin{bmatrix} | \\ \vec{x} \\ | \end{bmatrix}$$

- Often, vector notation is omitted and we just write x instead of $\vec{x}$.

Conventions - Vectors

- A *transpose* operation swaps rows and columns.

Conventions - Vectors

- A *transpose* operation swaps rows and columns.
- A transpose is denoted by a T in the exponent.

Conventions - Vectors

- A *transpose* operation swaps rows and columns.
- A tranpose is denoted by a T in the exponent.
- Accordingly, $x^T \in \mathbb{R}^n$ is considered to be a *row vector*

$$x^\top = \begin{bmatrix} -- & x^T & -- \end{bmatrix}$$

Dot Product

The dot product of two vectors $u, v \in \mathbb{R}^n$ is the number

$$u \cdot v = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

Dot Product

The dot product of two vectors $u, v \in \mathbb{R}^n$ is the number

$$u \cdot v = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

We may also write the dot product of two vectors $u, v \in \mathbb{R}^n$ as

$$u^\top v = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

This works for reasons we'll describe later (matrix multiplication).

Try It

Compute the dot product between vectors

$$u = [-2, 4, 5]^\top$$
$$v = [5, 10, -6]^\top.$$

Do the same for the dot product between u and itself.

Try It

Compute the dot product between vectors

$$u = [-2, 4, 5]^\top$$
$$v = [5, 10, -6]^\top.$$

The dot product is $-10 + 40 - 30 = 0$.

Do the same for the dot product between u and itself.

Try It

Compute the dot product between vectors

$$u = [-2, 4, 5]^\top$$
$$v = [5, 10, -6]^\top.$$

The dot product is $-10 + 40 - 30 = 0$. When the dot product between two vectors is 0, the vectors are said to be *orthogonal*.
Do the same for the dot product between u and itself.

Try It

Compute the dot product between vectors

$$u = [-2, 4, 5]^\top$$
$$v = [5, 10, -6]^\top.$$

Do the same for the dot product between u and itself.
The dot product is $4 + 16 + 25 = 45$.

Try It

Compute the dot product between vectors

$$u = [-2, 4, 5]^\top$$
$$v = [5, 10, -6]^\top.$$

Do the same for the dot product between u and itself.
The dot product is $4 + 16 + 25 = 45$. It's fact that $u \cdot u = ||u||_2^2$, so the norm of the vector u is $\sqrt{45}$.

Matrices

An $m \times n$ *matrix* is a rectangular array with m rows and n columns.

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

The notation a_{ij} is used to denote the element in row i, column j of the matrix A.

Matrices

An $m \times n$ *matrix* is a rectangular array with m rows and n columns.

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

The notation a_{ij} is used to denote the element in row i, column j of the matrix A.

The set of all $m \times n$ real-valued matrices is denoted $\mathbb{R}^{m \times n}$.

Matrices

It usually is counterproductive to write down every single element of n $m \times n$ matrix A.

We often will choose (depending on the situation) to write A in terms of its rows or columns.

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

Matrices

It usually is counterproductive to write down every single element of n $m \times n$ matrix C.

We often will choose (depending on the situation) to write C in terms of its rows or columns.

$$C = \begin{bmatrix} -- & c_1^T & -- \\ -- & c_2^T & -- \\ \vdots & \vdots & \vdots \\ -- & c_m^T & -- \end{bmatrix}$$

Matrix Transpose

- The transpose of a matrix swaps rows and columns, so $A \in \mathbb{R}^{m \times n}$ implies $A^T \in \mathbb{R}^{n \times m}$.

Matrix Transpose

- The transpose of a matrix swaps rows and columns, so $A \in \mathbb{R}^{m \times n}$ implies $A^T \in \mathbb{R}^{n \times m}$.
- More explicitly, we have

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

and

$$A^T = \begin{bmatrix} -- & a_1^T & -- \\ -- & a_2^T & -- \\ \vdots & \vdots & \vdots \\ -- & a_n^T & -- \end{bmatrix}$$

Matrix-Vector Products

An $m \times n$ matrix A and a vector $b \in \mathbb{R}^n$ can be multiplied together. The multiplication operation is defined as

$$A \cdot b = A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} b_1 a_{11} \\ b_1 a_{21} \\ \vdots \\ b_1 a_{m1} \end{bmatrix} + \begin{bmatrix} b_2 a_{12} \\ b_2 a_{22} \\ \vdots \\ b_2 a_{m2} \end{bmatrix} + \begin{bmatrix} b_n a_{1n} \\ b_n a_{2n} \\ \vdots \\ b_n a_{mn} \end{bmatrix}$$

Matrix-Vector Products

We can write (more succinctly)

$$A \cdot b = \begin{bmatrix} | & | & \cdots & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & \vdots & | \end{bmatrix} \cdot \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = b_1\vec{a}_1 + b_2\vec{a}_2 + \cdots + b_n\vec{a}_n$$

Try It

Compute

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

Try It

Compute

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

$$= 1 \begin{bmatrix} 4 \\ 0 \\ -2 \end{bmatrix} + 2 \begin{bmatrix} -1 \\ 1 \\ 10 \end{bmatrix} + 3 \begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix} + 4 \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$$

Try It

Compute

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

$$= 1 \begin{bmatrix} 4 \\ 0 \\ -2 \end{bmatrix} + 2 \begin{bmatrix} -1 \\ 1 \\ 10 \end{bmatrix} + 3 \begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix} + 4 \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 19 \\ 6 \\ 26 \end{bmatrix}$$

Try It

What's wrong with

$$\begin{bmatrix} 1 & -9 & 3 \\ -3 & 1 & \\ 9 & -2 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ -3 \\ 1 \\ 5 \end{bmatrix}$$

Matrix-Vector Product

- We've already seen

$$A \cdot b = \begin{bmatrix} | & | & \cdots & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & \vdots & | \end{bmatrix} \cdot \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = b_1\vec{a}_1 + b_2\vec{a}_2 + \cdots + b_n\vec{a}_n$$

Matrix-Vector Product

- We've already seen

$$A \cdot b = \begin{bmatrix} | & | & \cdots & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & \vdots & | \end{bmatrix} \cdot \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = b_1\vec{a}_1 + b_2\vec{a}_2 + \cdots + b_n\vec{a}_n$$

- We can also compute

$$A^T \cdot c = \begin{bmatrix} -- & a_1^T & -- \\ -- & a_2^T & -- \\ \vdots & \vdots & \vdots \\ -- & a_n^T & -- \end{bmatrix} \cdot \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} a_1^T c \\ a_2^T c \\ \vdots \\ a_n^T c \end{bmatrix}$$

Try It

Compute

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

using the dot product method.

Try It

Compute

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

using the dot product method.

We get

$$\begin{bmatrix} 8+2-6 \\ 10+6-1 \\ -6-4+1 \end{bmatrix} = \begin{bmatrix} 4 \\ 15 \\ -9 \end{bmatrix}$$

Matrix-Matrix Products

Suppose $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times k}$. Then $AB \in \mathbb{R}^{m\times k}$ is given by

$$A \cdot \begin{bmatrix} | & | & \cdots & | \\ \vec{b}_1 & \vec{b}_2 & \cdots & \vec{b}_k \\ | & | & \vdots & | \end{bmatrix} = \begin{bmatrix} | & | & \cdots & | \\ A\vec{b}_1 & A\vec{b}_2 & \cdots & A\vec{b}_k \\ | & | & \vdots & | \end{bmatrix}$$

Matrix-Matrix Products

Suppose $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times k}$. Then $AB \in \mathbb{R}^{m\times k}$ is given by

$$A \cdot \begin{bmatrix} | & | & \cdots & | \\ \vec{b}_1 & \vec{b}_2 & \cdots & \vec{b}_k \\ | & | & \vdots & | \end{bmatrix} = \begin{bmatrix} | & | & \cdots & | \\ A\vec{b}_1 & A\vec{b}_2 & \cdots & A\vec{b}_k \\ | & | & \vdots & | \end{bmatrix}$$

If we have A given to us in row form, then

$$A \cdot B = \begin{bmatrix} -- & a_1^T & -- \\ -- & a_2^T & -- \\ \vdots & \vdots & \vdots \\ -- & a_n^T & -- \end{bmatrix} \begin{bmatrix} | & | & \cdots & | \\ \vec{b}_1 & \vec{b}_2 & \cdots & \vec{b}_k \\ | & | & \vdots & | \end{bmatrix} = \begin{bmatrix} a_1^T b_1 & a_1^T b_2 & \cdots & a_1^T b_k \\ a_2^T b_1 & a_2^T b_2 & \cdots & a_2^T b_k \\ a_m^T b_1 & a_1^T b_2 & \cdots & a_m^T b_k \end{bmatrix}$$

Special Matrices

Identity Matrix

The *identity matrix* I_n (or sometimes just I when the dimension is clear) is

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

Invertible Matrices

Definition: A matrix $A \in \mathbb{R}^{n \times n}$ is invertible if there exists a matrix A^{-1} such that $AA^{-1} = A^{-1}A = I$.

Try It

Let $A \in \mathbb{R}^{2\times 3}$ be given by

$$\begin{bmatrix} -2 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

Find a matrix $B \in \mathbb{R}^{3\times 2}$ such that $AB = I_2$.

Try It

Let $A \in \mathbb{R}^{2\times 3}$ be given by

$$\begin{bmatrix} -2 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

Find a matrix $B \in \mathbb{R}^{3\times 2}$ such that $AB = I_2$.
The matrix

$$B = \begin{bmatrix} -2 & -2 \\ -3 & -4 \\ 5 & 7 \end{bmatrix}$$

works.

Try It

Let $A \in \mathbb{R}^{2\times 3}$ be given by

$$\begin{bmatrix} -2 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

Find a matrix $B \in \mathbb{R}^{3\times 2}$ such that $AB = I_2$.
The matrix

$$B = \begin{bmatrix} -2 & -2 \\ -3 & -4 \\ 5 & 7 \end{bmatrix}$$

works. In this example, the matrix B is a *right inverse* of A.

Permutation Matrices

Matrices of the form

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

are *permutation matrices*.

Permutation Matrices

Matrices of the form

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

are *permutation matrices*.

Such matrices can permute the rows or columns of a matrix.

Try It

Let

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

and

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

.
Compute the products AP and PA to examine the effect of multiplying by a permutation matrix.

Try It

Let

$$A = \begin{bmatrix} -4 & 1 & 6 \\ -5 & 3 & 1 \\ 3 & -2 & -1 \end{bmatrix}, P = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

.

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

Try It

Let

$$A = \begin{bmatrix} -4 & 1 & 6 \\ -5 & 3 & 1 \\ 3 & -2 & -1 \end{bmatrix}, P = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

.

$$AP = \begin{bmatrix} 1 & -4 & 6 \\ 3 & -5 & 1 \\ -2 & 3 & -1 \end{bmatrix}, PA = \begin{bmatrix} -5 & 3 & 1 \\ -4 & 1 & 6 \\ 3 & -2 & -1 \end{bmatrix}$$

Banded Matrices

- Some matrices have special structure, e.g. zeros in patterns.
- *Banded matrices* only have nonzero entries in their diagonals and several bands away from the diagonal.

Banded Matrices

- Some matrices have special structure, e.g. zeros in patterns.
- *Banded matrices* only have nonzero entries in their diagonals and several bands away from the diagonal.
- A *tridiagonal matrix* looks like

$$\begin{bmatrix} a & b & 0 & 0 & 0 & \cdots & 0 \\ c & d & e & 0 & 0 & \cdots & 0 \\ 0 & f & g & h & 0 & \cdots & 0 \\ 0 & 0 & i & j & k & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & v & w & x \\ 0 & 0 & \cdots & 0 & 0 & y & z \end{bmatrix}$$

Try It

$$\begin{bmatrix} a & b & 0 & 0 & 0 & \cdots & 0 \\ c & d & e & 0 & 0 & \cdots & 0 \\ 0 & f & g & h & 0 & \cdots & 0 \\ 0 & 0 & i & j & k & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & v & w & x \\ 0 & 0 & \cdots & 0 & 0 & y & z \end{bmatrix}$$

How many operations are required to multiply an $n \times n$ tridiagonal matrix and an n-dimensional vector?

Try It

$$\begin{bmatrix} a & b & 0 & 0 & 0 & \cdots & 0 \\ c & d & e & 0 & 0 & \cdots & 0 \\ 0 & f & g & h & 0 & \cdots & 0 \\ 0 & 0 & i & j & k & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & v & w & x \\ 0 & 0 & \cdots & 0 & 0 & y & z \end{bmatrix}$$

How many operations are required to multiply an $n \times n$ tridiagonal matrix and an n-dimensional vector? $2 + 2 + 3(\text{n-}2)$

Orthogonal Matrices

An orthogonal matrix $Q \in \mathbb{R}^{n\times n}$ satisfies

$$Q^T Q = QQ^T = I_n$$

Orthogonal Matrices

An orthogonal matrix $Q \in \mathbb{R}^{n \times n}$ satisfies

$$Q^T Q = QQ^T = I_n$$

Question: Is an orthogonal matrix invertible? Why or why not?

Matrix and Vector Norms

- We measure the "size" of real numbers with absolute value.

Matrix and Vector Norms

- We measure the "size" of real numbers with absolute value.
- A *norm* generalizes this idea to higher-dimensional sets. Precisely, a norm is a function
$$d : V \to [0, \infty)$$
that associates a nonnegative number to every element in a set V (the function d must satisfy certain conditions).

Matrix and Vector Norms

We have the following norms for $\mathbb{R}^n$:

- $||x||_2 = \sqrt{x_1^2 + \cdots + x_n^2}$

Matrix and Vector Norms

We have the following norms for $\mathbb{R}^n$:

- $||x||_2 = \sqrt{x_1^2 + \cdots + x_n^2}$
- $||x||_1 = |x_1| + \cdots + |x_n|$

Matrix and Vector Norms

We have the following norms for $\mathbb{R}^n$:

- $||x||_2 = \sqrt{x_1^2 + \cdots + x_n^2}$
- $||x||_1 = |x_1| + \cdots + |x_n|$
- $||x||_\infty = \max_{i=1,\ldots,n} |x_i|$

Matrix and Vector Norms

We have the following norms for matrices in $\mathbb{R}^{m\times n}$:

- $||A||_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2}$, the Frobenius norm.

Matrix and Vector Norms

We have the following norms for matrices in $\mathbb{R}^{m\times n}$:

- $||A||_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2}$, the Frobenius norm.
- $||A||_1 = \max_{1\leq j\leq n}\sum_i |a_{ij}|$, the maximum (absolute) column sum.
- $||A||_\infty = \max_{1\leq i\leq m}\sum_j |a_{ij}|$, the maximum (absolute) row sum.
- There are a variety of *operator norms* for matrices that are outside our scope for now.

Practice: Matrices

Vector Spaces

What Is Linear Algebra?

- The subject is usually involves an *abstract algebra*, which (informally) is a set of objects which is closed under some operations.
- The subject is *linear* in that we usually consider linear expressions and functions.

Vector Space

Definition A vector space V on $\mathbb{R}$ is a set of objects (called *vectors*) endowed with two operations '$+$' and '$\cdot$' that satisfy the following:

1. For two elements $v, w \in$, the "sum" $v + w \in V$.

Vector Space

Definition A vector space V on $\mathbb{R}$ is a set of objects (called *vectors*) endowed with two operations '$+$' and '$\cdot$' that satisfy the following:

1. For two elements $v, w \in$, the "sum" $v + w \in V$.
2. For any $\alpha \in \mathbb{R}$, the "product" $\alpha v \in V$.

Vector Space

Definition Additionally, the operations on the vector space V must satisfy the following *axioms* for every $u, v, w \in V$ and $\alpha, \beta \in \mathbb{R}$:

1. Commutativity: $v + w = w + v$ for all $v, w \in V$.
2. Associativity: $(u + v) + w = u + (v + w)$ for all $u, v, w \in V$.
3. Zero: There exists a vector, denoted $\vec{0}$, such that $v + \vec{0} = v$ for all $v \in V$.
4. Additive inverse: For every $v \in V$ there exists a vector $w \in V$ such that $v + w = 0$. We denote w by $-v$.

Vector Space (cont'd)

Definition Additionally, the operations on the vector space V must satisfy the following *axioms* for every $u, v, w \in V$ and $\alpha, \beta \in \mathbb{R}$:

5. $1 \cdot v = v$ for all $v \in V$.
6. $(\alpha\beta) \cdot v = \alpha \cdot (\beta \cdot v)$ for all $\alpha, \beta \in \mathbb{R}$, $v \in V$.
7. $\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v$ for all $\alpha \in \mathbb{R}, v \in V$
8. $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$ for all $\alpha, \beta \in \mathbb{R}, v \in V$.

Vector Space (cont'd)

Definition Additionally, the operations on the vector space V must satisfy the following *axioms* for every $u, v, w \in V$ and $\alpha, \beta \in \mathbb{R}$:

5. $1 \cdot v = v$ for all $v \in V$.
6. $(\alpha\beta) \cdot v = \alpha \cdot (\beta \cdot v)$ for all $\alpha, \beta \in \mathbb{R}$, $v \in V$.
7. $\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v$ for all $\alpha \in \mathbb{R}$, $v \in V$
8. $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$ for all $\alpha, \beta \in \mathbb{R}$, $v \in V$.

So Abstract!

What are these "vectors" and "operations"? Can we get a real example?

The Intuitive Example

An intuitive first example is the set $\mathbb{R}$, endowed with the usual $+, \cdot$ operations between real numbers.

Is it clear to you that all the properties stated above hold?

A Real Example

The set $\mathbb{R}^n$, endowed with the operations $+, \cdot$, is a vector space, with the operations defined as:

$$u + v = \begin{bmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n \end{bmatrix}$$

and

$$\alpha \cdot v = \begin{bmatrix} \alpha v_1 \\ \vdots \\ \alpha v_n \end{bmatrix}$$

A Tough Example

The set of strictly positive real numbers $\mathbb{R}_+$ is a vector space, with operations defined as

$$u + v = uv$$

and

$$\alpha v = v^{\alpha}$$

Try It

The set of strictly positive real numbers $\mathbb{R}_+$ is a vector space, with operations defined as

$$u + v = uv$$

and

$$\alpha v = v^{\alpha}$$

Try It: Verify that the above is a vector space.

Try It

Let u, v be any members of the set $\mathbb{R}_+$.

1. Commutativity: $\mathrm{u} + \mathrm{v} = \mathrm{uv} = \mathrm{vu} = \mathrm{v}{+}\mathrm{u}$
2. Associativity: $(\mathrm{u} + \mathrm{v}) + \mathrm{w} = \mathrm{uvw} = \mathrm{u} + (\mathrm{v} + \mathrm{w})$
3. Zero Vector: The number $1 \in \mathbb{R}_+$ works.
4. For every $v \in \mathbb{R}^n$, define $-v$ as $\frac{1}{v}$.

Try It

5. $1 \cdot v = v^1 = v$
6. $(\alpha\beta) \cdot v = v^{\alpha\beta} = (v^\beta)^\alpha = \alpha \cdot (\beta \cdot v)$.
7. $\alpha \cdot (u + v) = (uv)^\alpha = u^\alpha v^\alpha = \alpha \cdot u + \alpha \cdot v$.
8. $(\alpha + \beta) \cdot v = v^{\alpha+\beta} = v\alpha v^\beta = \alpha \cdot v + \beta \cdot v$.

Other Examples of Vector Spaces

The set of continuous functions on an interval $[a, b]$ is a vector space with, with operations

$$f + g, \alpha \cdot f$$

defined as usual.

Linear Combinations & Bases

Linear Combination

Let V be a vector space on $\mathbb{R}$ with operations $+, \cdot$, and let $v_1, \ldots, v_p \in V$.
A *linear combination* of the vectors $v_1, \ldots, v_p$ is a sum

$$\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2 + \cdots + \alpha_p v_p$$

for any scalars $\alpha_1, \ldots, \alpha_p \in \mathbb{R}$.

Span

The *span* of a set of vectors is the set of all possible linear combinations of those vectors.

$$\mathrm{span}(v_1, \ldots, v_p) = \{\alpha_1 v_1 + \cdots + \alpha_p v_p \mid \alpha_1, \ldots, \alpha_p \in \mathbb{R}\}$$

Try It

1. In words, what is

$$\mathrm{span}\left(\begin{bmatrix}1\\1\end{bmatrix}\right)?$$

2. Show that any vector y in

$$\mathrm{span}\left(\begin{bmatrix}1\\-2\\3\end{bmatrix},\begin{bmatrix}0\\5\\-1\end{bmatrix},\begin{bmatrix}3\\1\\0\end{bmatrix}\right)$$

can be written as $y = Ax$ for some matrix A and some vector x.

Try It

1. In words, what is
$$\mathrm{span}(\begin{bmatrix}1\\1\end{bmatrix})?$$

Using the definition, we have

$$\mathrm{span}(\begin{bmatrix}1\\1\end{bmatrix}) = \left\{\alpha \cdot \begin{bmatrix}1\\1\end{bmatrix} \mid \alpha \in \mathbb{R}\right\}$$

Try It

1. In words, what is
$$\operatorname{span}(\begin{bmatrix}1\\1\end{bmatrix})?$$

Using the definition, we have

$$\begin{aligned}\operatorname{span}(\begin{bmatrix}1\\1\end{bmatrix}) &= \left\{\alpha\cdot\begin{bmatrix}1\\1\end{bmatrix} \mid \alpha\in\mathbb{R}\right\}\\ &= \left\{\begin{bmatrix}\alpha\\\alpha\end{bmatrix} \mid \alpha\in\mathbb{R}\right\}\end{aligned}$$

Try It

1. In words, what is
$$\operatorname{span}\left(\begin{bmatrix}1\\1\end{bmatrix}\right)?$$

Using the definition, we have

$$\begin{aligned}\operatorname{span}\left(\begin{bmatrix}1\\1\end{bmatrix}\right) &= \left\{\alpha \cdot \begin{bmatrix}1\\1\end{bmatrix} \mid \alpha \in \mathbb{R}\right\}\\ &= \left\{\begin{bmatrix}\alpha\\\alpha\end{bmatrix} \mid \alpha \in \mathbb{R}\right\}\\ &= \{\text{all vectors in } \mathbb{R}^2 \text{ with the same first and second entry}\}\end{aligned}$$

Try It

2. Show that any vector y in

$$\operatorname{span}\left(\begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 5 \\ -1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}\right)$$

can be written as $y = Ax$ for some matrix A and some vector x.

Try It

2. Show that any vector y in

$$\operatorname{span}(\begin{bmatrix}1\\-2\\3\end{bmatrix}, \begin{bmatrix}0\\5\\-1\end{bmatrix}, \begin{bmatrix}3\\1\\0\end{bmatrix})$$

can be written as $y = Ax$ for some matrix A and some vector x.

We have

$$\operatorname{span}(\begin{bmatrix}1\\-2\\3\end{bmatrix}, \begin{bmatrix}0\\5\\-1\end{bmatrix}, \begin{bmatrix}3\\1\\0\end{bmatrix}) = \left\{\alpha_1 \begin{bmatrix}1\\-2\\3\end{bmatrix} + \alpha_2 \begin{bmatrix}0\\5\\-1\end{bmatrix} + \alpha_3 \begin{bmatrix}3\\1\\0\end{bmatrix} \middle| \alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}\right\}$$

Try It

2. Show that any vector y in

$$\mathrm{span}(\begin{bmatrix}1\\-2\\3\end{bmatrix}, \begin{bmatrix}0\\5\\-1\end{bmatrix}, \begin{bmatrix}3\\1\\0\end{bmatrix})$$

can be written as $y = Ax$ for some matrix A and some vector x.

We have

$$\mathrm{span}(\begin{bmatrix}1\\-2\\3\end{bmatrix}, \begin{bmatrix}0\\5\\-1\end{bmatrix}, \begin{bmatrix}3\\1\\0\end{bmatrix}) = \left\{\alpha_1 \begin{bmatrix}1\\-2\\3\end{bmatrix} + \alpha_2 \begin{bmatrix}0\\5\\-1\end{bmatrix} + \alpha_3 \begin{bmatrix}3\\1\\0\end{bmatrix} \middle| \alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}\right\}$$

$$= \left\{\begin{bmatrix}1 & 0 & 3\\-2 & 5 & 1\\3 & -1 & 0\end{bmatrix} \begin{bmatrix}\alpha_1\\\alpha_2\\\alpha_3\end{bmatrix} \middle| \alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}\right\}$$

Basis

A set of vectors $v_1, \ldots, v_p \in V$ is a *basis* for V if the following holds.

For every $v \in V$, there exists a unique set of coefficients $\alpha_1, \ldots, \alpha_p$ such that

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_p v_p$$

Basis

A set of vectors $v_1, \ldots, v_p \in V$ is a *basis* for V if the following holds.

For every $v \in V$, there exists a unique set of coefficients $\alpha_1, \ldots, \alpha_p$ such that

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_p v_p$$

Basis

A set of vectors $v_1, \dots, v_p \in V$ is a *basis* for V if the following holds.

For every $v \in V$, there exists a unique set of coefficients $\alpha_1, \dots, \alpha_p$ such that

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_p v_p$$

The numbers $\alpha_1, \dots, \alpha_p$ are the *coordinates* of v with respect to basis $\{v_1, \dots, v_p\}$.

Basis

A set of vectors $v_1, \ldots, v_p \in V$ is a *basis* for V if the following holds.

For every $v \in V$, there exists a unique set of coefficients $\alpha_1, \ldots, \alpha_p$ such that

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_p v_p$$

The numbers $\alpha_1, \ldots, \alpha_p$ are the coordinates of v with respect to basis $\{v_1, \ldots, v_p\}$.

If a set of vectors $v_1, \ldots, v_p \in V$ is a *basis* for V, then clearly $\mathrm{span}(v_1, \ldots, v_p) = V$.

Example

- Consider the vector space $\mathbb{R}$ with the usual operations.

Example

- Consider the vector space $\mathbb{R}$ with the usual operations.
- The number $1 \in \mathbb{R}$ is a basis.

Example

- Consider the vector space $\mathbb{R}$ with the usual operations.
- The number $1 \in \mathbb{R}$ is a basis.
- *Proof:* For every $v \in \mathbb{R}$, we can write $v = \alpha \cdot 1$ by taking $\alpha = v$. Moreover, this representation is unique.

Try It

Is the set of vectors $\{1, 2\}$ a basis for $\mathbb{R}$?

Is the set of vectors $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$ a basis for $\mathbb{R}^2$?

Is the set of vectors $\left\{ \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$ a basis for $\mathbb{R}^2$?

Standard Basis Vectors

The *standard basis vectors* for $\mathbb{R}^n$ are denoted $e_1, \ldots, e_n$ with

$$e_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

the vector of all 0's except for a 1 in the ith position.

Linear Independence & Dependence

A set of vectors $v_1, \ldots, v_p \in V$ is *linearly dependent* if and only if one of the vectors can be written as a linear combination of the other vectors, i.e.

$$v_k = \beta_1 \cdot v_1 + \cdots \beta_{k-1} \cdot v_{k-1} + \beta_{k+1} \cdot v_{k+1} + \cdots \beta_p \cdot v_p$$

for some k and some cofficients $\beta_1, \ldots, \beta_{k-1}, \beta_{k+1}, \cdots \beta_p$.

A set of vectors $v_1, \ldots, v_p$ is linearly independent if and only if the set is not linearly dependent.

Why Do We Care?

Linear (in)dependence is an important concept because it informs us if there is any **redundancy** in a set of vectors.

A linearly dependent set is redundant. It contains at least one vector we don't "need" (informally).

Try It

Is the set of vectors

$$\left\{ \begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}2\\1\end{bmatrix}, \begin{bmatrix}2\\4\end{bmatrix} \right\}$$

linearly independent?

Try It

Is the set of vectors

$$\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \end{bmatrix} \right\}$$

linearly independent?

No. We clearly have

$$\begin{bmatrix} 2 \\ 4 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 0 \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

This reveals to us that

$$\operatorname{span}\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \end{bmatrix} \right) = \operatorname{span}\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right)$$

Practice: Vector Spaces

Linear Transformations

Functions on Vector Spaces

We're now going to study functions on vector spaces! This is the heart of linear algebra.

Let V, W be vector spaces. A *transformation* is a function $T : V \to W$.

Linear Transformations

A transformation $T : V \to W$ is *linear* if the following two conditions hold:

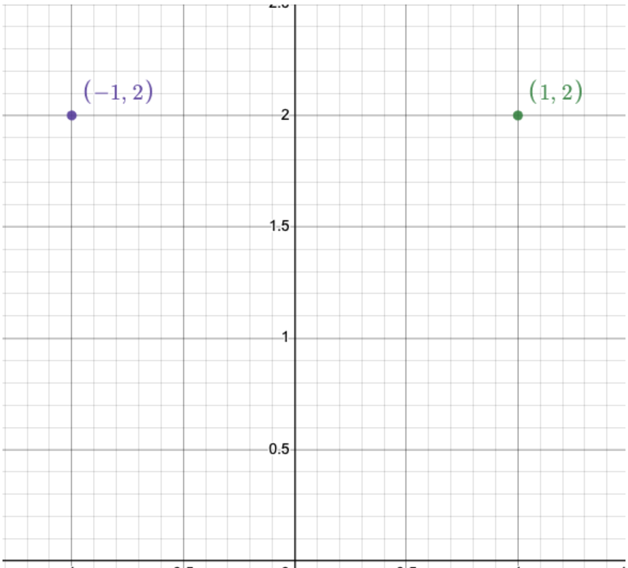
- $T(u + v) = T(u) + T(v)$ for all $u, v \in V$.
- $T(\alpha v) = \alpha T(v)$ for all $\alpha \in \mathbb{R}, v \in V$.

Example: Reflection

Let $V = \mathbb{R}^2$, $W = \mathbb{R}^2$, and

$$T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} -x \\ y \end{bmatrix}$$

Example: Reflection



Example: Reflection

Let $V = \mathbb{R}^2$, $W = \mathbb{R}^2$, and

$$T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} -x \\ y \end{bmatrix}$$

Show that T is a linear transformation.
We have

$$T(u+v) = T(\begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \end{bmatrix}) = \begin{bmatrix} -u_1 - v_1 \\ u_2 + v_2 \end{bmatrix} = \begin{bmatrix} -u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} -v_1 \\ v_2 \end{bmatrix} = T(u) + T(v)$$

Example: Reflection

Let $V = \mathbb{R}^2$, $W = \mathbb{R}^2$, and

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} -x \\ y \end{bmatrix}$$

Show that T is a linear transformation.
We have

$$T(u+v) = T\left(\begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \end{bmatrix}\right) = \begin{bmatrix} -u_1 - v_1 \\ u_2 + v_2 \end{bmatrix} = \begin{bmatrix} -u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} -v_1 \\ v_2 \end{bmatrix} = T(u) + T(v)$$

$$T(\alpha v) = \begin{bmatrix} -\alpha v_1 \\ \alpha v_2 \end{bmatrix} = \alpha \begin{bmatrix} -v_1 \\ v_2 \end{bmatrix} = \alpha T(v)$$

Try It

Suppose that $T : \mathbb{R}^2 \to \mathbb{R}^2$ acts on the standard basis vectors via

$$T(e_1) = \begin{bmatrix} 3 \\ 7 \end{bmatrix}, T(e_2) = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$$

Compute T($\begin{bmatrix} 2 \\ 0 \end{bmatrix}$).

Compute T($\begin{bmatrix} -3 \\ 2 \end{bmatrix}$).

Try It

Suppose that $T : \mathbb{R}^2 \to \mathbb{R}^2$ acts on the standard basis vectors via

$$T(e_1) = \begin{bmatrix} 3 \\ 7 \end{bmatrix}, T(e_2) = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$$

Compute T($\begin{bmatrix} 2 \\ 0 \end{bmatrix}$).

Compute T($\begin{bmatrix} -3 \\ 2 \end{bmatrix}$).

Hint: rewrite $\begin{bmatrix} 2 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} -3 \\ 2 \end{bmatrix}$ as linear combinations of e_1 and e_2 before applying T.

Try It

Suppose that $T : \mathbb{R}^2 \to \mathbb{R}^2$ acts on the standard basis vectors via

$$T(e_1) = \begin{bmatrix} 3 \\ 7 \end{bmatrix}, T(e_2) = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$$

Compute T($\begin{bmatrix} 2 \\ 0 \end{bmatrix}$).

We know $\begin{bmatrix} 2 \\ 0 \end{bmatrix} = 2e_1 + 0e_2$. By linearity, we have

$$T(2e_1 + 0e_2) = 2T(e_1) + 0T(e_2) = \begin{bmatrix} 6 \\ 14 \end{bmatrix}$$

Try It

Suppose that $T : \mathbb{R}^2 \to \mathbb{R}^2$ acts on the standard basis vectors via

$$T(e_1) = \begin{bmatrix} 3 \\ 7 \end{bmatrix}, T(e_2) = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$$

Compute T($\begin{bmatrix} x \\ y \end{bmatrix}$) for arbitrary x, y.

Solution

Suppose that $T : \mathbb{R}^2 \to \mathbb{R}^2$ acts on the standard basis vectors via

$$T(e_1) = \begin{bmatrix} 3 \\ 7 \end{bmatrix}, T(e_2) = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$$

Compute T($\begin{bmatrix} x \\ y \end{bmatrix}$) for arbitrary x, y.

Solution

Suppose that $T : \mathbb{R}^2 \to \mathbb{R}^2$ acts on the standard basis vectors via

$$T(e_1) = \begin{bmatrix} 3 \\ 7 \end{bmatrix}, T(e_2) = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$$

Compute T($\begin{bmatrix} x \\ y \end{bmatrix}$) for arbitrary x, y. Clearly, we have

$$T(\begin{bmatrix} x \\ y \end{bmatrix}) = T(xe_1 + ye_2)$$

Solution

Suppose that $T : \mathbb{R}^2 \to \mathbb{R}^2$ acts on the standard basis vectors via

$$T(e_1) = \begin{bmatrix} 3 \\ 7 \end{bmatrix}, T(e_2) = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$$

Compute T($\begin{bmatrix} x \\ y \end{bmatrix}$) for arbitrary x, y. Clearly, we have

$$T(\begin{bmatrix} x \\ y \end{bmatrix}) = T(xe_1 + ye_2) = xT(e_1) + yT(e_2)$$

Solution

Suppose that $T : \mathbb{R}^2 \to \mathbb{R}^2$ acts on the standard basis vectors via

$$T(e_1) = \begin{bmatrix} 3 \\ 7 \end{bmatrix}, T(e_2) = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$$

Compute T($\begin{bmatrix} x \\ y \end{bmatrix}$) for arbitrary x, y. Clearly, we have

$$T(\begin{bmatrix} x \\ y \end{bmatrix}) = T(xe_1 + ye_2) = xT(e_1) + yT(e_2) = x \begin{bmatrix} 3 \\ 7 \end{bmatrix} + y \begin{bmatrix} -2 \\ 4 \end{bmatrix}$$

Solution

Suppose that $T : \mathbb{R}^2 \to \mathbb{R}^2$ acts on the standard basis vectors via

$$T(e_1) = \begin{bmatrix} 3 \\ 7 \end{bmatrix}, T(e_2) = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$$

Compute T($\begin{bmatrix} x \\ y \end{bmatrix}$) for arbitrary x, y. Clearly, we have

$$T(\begin{bmatrix} x \\ y \end{bmatrix}) = T(xe_1 + ye_2) = xT(e_1) + yT(e_2) = x \begin{bmatrix} 3 \\ 7 \end{bmatrix} + y \begin{bmatrix} -2 \\ 4 \end{bmatrix}$$

Using what we know about matrix multiplication, this is the same as
Clearly, we have

$$\begin{bmatrix} 3 & -2 \\ 7 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Discussion

Something really deep just happened...we turned a linear transformation into a matrix!

Matrices Are Linear Transformations

Suppose we have a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$, and we know

$$\begin{aligned} T(e_1) &= a_1 \\ T(e_2) &= a_2 \\ &\vdots \\ T(e_n) &= a_n \end{aligned}$$

Then for an $x \in \mathbb{R}^n$, we know $T(x)$.

Matrices Are Linear Transformations

The result of $T(x)$ will be given by

$$
\begin{aligned}
T(x_1 e_1 + \cdots x_n e_n) &= x_1 \begin{bmatrix} | \\ a_1 \\ | \end{bmatrix} + \cdots + x_n \begin{bmatrix} | \\ a_n \\ | \end{bmatrix} \\
&= \begin{bmatrix} | & \cdots & | \\ a_1 & \cdots & a_n \\ | & \cdots & | \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \\
&= Ax
\end{aligned}
$$

Change of Basis

Consider two different bases for $\mathbb{R}^2$, given by

$$B_1 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

$$B_2 = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

Change of Basis - Try It

Consider two different bases for $\mathbb{R}^2$, given by

$$B_1 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

$$B_2 = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

Consider a linear transformation $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 2x \\ -3y \end{bmatrix}$.

Find the matrix for this linear transformation with respect to the bases B_1, B_2

Change of Basis - Try It

Consider two different bases for $\mathbb{R}^2$, given by

$$B_1 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

$$B_2 = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

Consider a linear transformation $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 2x \\ -3y \end{bmatrix}$.

Find the matrix for this linear transformation with respect to the bases B_1, B_2.

The two matrices should be

$$\begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ -3 & -3 \end{bmatrix}$$

Special Matrix = Special Function

- We've learned that a matrix represents a linear transformation (with respect to some basis).
- It shouldn't surprise you to learn that matrices with "special" properties are really just representing special types of functions.

Square Matrices

A matrix $A \in \mathbb{R}^{m \times n}$ represents a mapping $T : \mathbb{R}^n \to \mathbb{R}^m$.

Accordingly, square matrices $A \in \mathbb{R}^{n \times n}$ simply represent a mapping $T : \mathbb{R}^n \to \mathbb{R}^n$

Invertible Matrices

Definition: A matrix $A \in \mathbb{R}^{n\times n}$ is *invertible* if there exists a matrix A^{-1} such that $AA^{-1} = A^{-1}A = I$.

Intuition: A matrix A is invertible if the linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ it represents is invertible (bijective)

Orthogonal Matrices

Definition: An orthogonal matrix $O \in \mathbb{R}^n$ satisfies $O^\top O = OO^\top = I_n$.

Orthogonal Matrices

Definition: An orthogonal matrix $O \in \mathbb{R}^n$ satisfies $O^\top O = OO^\top = I_n$. In other words, $O^T = O^{-1}$.

Geometric Intuition: the linear transformation O corresponds to a rotation or reflection. In these cases, it turns out that the transpose "undoes" the rotation or reflection.

Try It

Consider the transformation

$$T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} 0 \\ y \end{bmatrix}$$

.

Would you expect the matrix that represents this matrix to be invertible? Why or why not?

Thoughts on Invertibility

- Hopefully you've been convinced that the *columns* of a matrix are important. This is why we started writing matrix in column form early.

Thoughts on Invertibility

- Hopefully you've been convinced that the *columns* of a matrix are important. This is why we started writing matrix in column form early.
- In particular, the columns of a square matrix being *linearly independent* or *linearly dependent* is directly tied to the invertibility of a matrix. You'll show this on the worksheet

Some (Informal) Vocabulary

- The *dimension* of a vector space V is the number of items in a basis of the V. We denote the $\dim V$.

Some (Informal) Vocabulary

- The *dimension* of a vector space V is the number of items in a basis of the V. We denote the $\dim V$.
- The *range* of a linear transformation $T : V \to W$ is $\{ Tv \mid v \in V\} \subseteq W$.

Some (Informal) Vocabulary

- The *dimension* of a vector space V is the number of items in a basis of the V. We denote the $\dim V$.
- The *range* of a linear transformation $T : V \to W$ is $\{\, Tv \mid v \in V\} \subseteq W$.
- The *null space* of a linear transformation $T : V \to W$ is $\{v \in V | Tv = 0\} \subseteq V$.

Some (Informal) Observations

Again, now let T be linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ for ease.

- We have $\mathrm{range}(T) \subseteq \mathbb{R}^n$ is a vector space, and $\mathrm{null}(T) \subseteq \mathbb{R}^n$ is a vector space.

Some (Informal) Observations

Again, now let T be linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ for ease.

- We have $\mathrm{range}(T) \subseteq \mathbb{R}^n$ is a vector space, and $\mathrm{null}(T) \subseteq \mathbb{R}^n$ is a vector space.
- A matrix A (corresponding to T) is invertible if and only if $\dim(\mathrm{null}(T)) = 0$.

Some (Informal) Observations

Again, now let T be linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ for ease.

- We have $\mathrm{range}(T) \subseteq \mathbb{R}^n$ is a vector space, and $\mathrm{null}(T) \subseteq \mathbb{R}^n$ is a vector space.
- A matrix A (corresponding to T) is invertible if and only if $\dim(\mathrm{null}(T)) = 0$.
- A matrix A (corresponding to T) is invertible if and only if $\dim(\mathrm{range}(T)) = n$.

Try It

1. Consider the invertible matrix

$$A = I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

. How many vectors are in the null space of A? What are they?

2. Consider the non-invertible matrix

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

. How many vectors are in the null space of A? What are they?

Try It

1. Consider the invertible matrix

$$A = I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

. How many vectors are in the null space of A? What are they?
The null space of A is exactly

$$\mathrm{null}(A) = \{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \}$$

Try It

2. Consider the non-invertible matrix

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

. How many vectors are in the null space of A? What are they?
The null space of A is exactly

$$\mathrm{null}(A) = \{ \begin{bmatrix} 0 \\ \alpha \end{bmatrix} | \alpha \in \mathbb{R} \}$$

so there are infinitely many vectors in the null space.

Some Famous Decompositions

Determinant

The determinant is a function

$$\det : \mathbb{R}^{n\times n} \to \mathbb{R}$$

Determinant

The determinant is a function

$$\det : \mathbb{R}^{n\times n} \to \mathbb{R}$$

The function $\det$ is uniqueley defined by three key geometric properties.

Determinant

1. We have $\det(I_n) = 1$.

Determinant

1. We have $\det(I_n) = 1$.
2. The function $\det$ is *multilinear*, i.e. linear in every column argument. This means

$$\det\left(\begin{bmatrix} | & \cdots & | & \cdots & | \\ a_1 & \cdots & \alpha b + c & \cdots & a_n \\ | & \cdots & | & \cdots & | \end{bmatrix}\right) = \alpha \det\left(\begin{bmatrix} | & \cdots & | & \cdots & | \\ a_1 & \cdots & b & \cdots & a_n \\ | & \cdots & | & \cdots & | \end{bmatrix}\right) + \det\left(\begin{bmatrix} | & \cdots & | & \cdots & | \\ a_1 & \cdots & c & \cdots & a_n \\ | & \cdots & | & \cdots & | \end{bmatrix}\right)$$

Determinant

1. We have $\det(I_n) = 1$.

2. The function $\det$ is *multilinear*, i.e. linear in every column argument. This means

$$\det\left(\begin{bmatrix} | & \cdots & | & \cdots & | \\ a_1 & \cdots & \alpha b + c & \cdots & a_n \\ | & \cdots & | & \cdots & | \end{bmatrix}\right) = \alpha \det\left(\begin{bmatrix} | & \cdots & | & \cdots & | \\ a_1 & \cdots & b & \cdots & a_n \\ | & \cdots & | & \cdots & | \end{bmatrix}\right) + \det\left(\begin{bmatrix} | & \cdots & | & \cdots & | \\ a_1 & \cdots & c & \cdots & a_n \\ | & \cdots & | & \cdots & | \end{bmatrix}\right)$$

3. If any two columns of the matrix are identical, its determinant is zero.

Calculating The Determinant

- Determinants for 2×2 matrices have an easy formula,

$$\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \right) = ad - bc$$

Calculating The Determinant

- Determinants for 2×2 matrices have an easy formula,

$$\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \right) = ad - bc$$

- For 3×3 matrices a formula exists; for 4×4 and beyond you need to use the cofactor method.

Calculating The Determinant

$$\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{matrix} =$$

$$\underline{a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}} - (\underline{a_{31}a_{22}a_{13} + a_{32}a_{23}a_{11} + a_{33}a_{21}a_{12}})$$

Eigenvectors

A *eigenvector* of a matrix $A \in \mathbb{R}^{n \times n}$ with associated *eigenvalue* λ satisfies the equation

$$Ax = \lambda x$$

Eigenvectors

A *eigenvector* of a matrix $A \in \mathbb{R}^{n \times n}$ with associated *eigenvalue* λ satisfies the equation

$$Ax = \lambda x$$

Note: the zero vector with eigenvalue 0 always satisfies this equation; so typically, when we speak of eigenvectors and eigenvalues, we insist that the eigenvector is a non-zero vector.

Eigenvectors

A *eigenvector* of a matrix $A \in \mathbb{R}^{n \times n}$ with associated *eigenvalue* λ satisfies the equation

$$Ax = \lambda x$$

Suppose $q_1, \ldots, q_n$ are eigenvectors of A with eigenvalues $\lambda_1, \ldots, \lambda_n$.

Eigenvectors

A *eigenvector* of a matrix $A \in \mathbb{R}^{n \times n}$ with associated *eigenvalue* λ satisfies the equation

$$Ax = \lambda x$$

Suppose $q_1, \ldots, q_n$ are eigenvectors of A with eigenvalues $\lambda_1, \ldots, \lambda_n$.

Then we have

$$A \begin{bmatrix} | & | & \cdots & | \\ q_1 & q_2 & \cdots & q_n \\ | & | & \vdots & | \end{bmatrix}$$

Eigenvectors

A *eigenvector* of a matrix $A \in \mathbb{R}^{n \times n}$ with associated *eigenvalue* λ satisfies the equation

$$Ax = \lambda x$$

Suppose $q_1, \ldots, q_n$ are eigenvectors of A with eigenvalues $\lambda_1, \ldots, \lambda_n$.

Then we have

$$A \begin{bmatrix} | & | & \cdots & | \\ q_1 & q_2 & \cdots & q_n \\ | & | & \vdots & | \end{bmatrix} = \begin{bmatrix} | & | & \cdots & | \\ Aq_1 & Aq_2 & \cdots & Aq_n \\ | & | & \vdots & | \end{bmatrix} =$$

Eigenvectors

A *eigenvector* of a matrix $A \in \mathbb{R}^{n \times n}$ with associated *eigenvalue* λ satisfies the equation

$$Ax = \lambda x$$

Suppose $q_1, \ldots, q_n$ are eigenvectors of A with eigenvalues $\lambda_1, \ldots, \lambda_n$.

Then we have

$$A \begin{bmatrix} | & | & \cdots & | \\ q_1 & q_2 & \cdots & q_n \\ | & | & \vdots & | \end{bmatrix} = \begin{bmatrix} | & | & \cdots & | \\ Aq_1 & Aq_2 & \cdots & Aq_n \\ | & | & \vdots & | \end{bmatrix} = \begin{bmatrix} | & | & \cdots & | \\ \lambda_1 q_1 & \lambda_2 q_2 & \cdots & \lambda_n q_n \\ | & | & \vdots & | \end{bmatrix}$$

Eigenvectors

A *eigenvector* of a matrix $A \in \mathbb{R}^{n \times n}$ with associated *eigenvalue* λ satisfies the equation

$$Ax = \lambda x$$

Suppose $q_1, \ldots, q_n$ are eigenvectors of A with eigenvalues $\lambda_1, \ldots, \lambda_n$.

Then we have

$$\begin{bmatrix} | & | & \cdots & | \\ \lambda_1 q_1 & \lambda_2 q_2 & \cdots & \lambda_n q_n \\ | & | & \vdots & | \end{bmatrix} = \begin{bmatrix} | & | & \cdots & | \\ q_1 & q_2 & \cdots & q_n \\ | & | & \vdots & | \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

Eigenvector Decomposition

If we let

$$Q = \begin{bmatrix} | & | & \cdots & | \\ q_1 & q_2 & \cdots & q_n \\ | & | & \vdots & | \end{bmatrix}, \Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix} \tag{1}$$

then the above work implies

$$AQ = Q\Lambda$$

and so

$$A = Q\Lambda Q^{-1}$$

Eigenvector Decomposition

$$AQ = Q\Lambda$$

and so

$$A = Q\Lambda Q^{-1}$$

The above is the *eigenvector decomposition* of the matrix A. This only holds if Q is invertible, i.e. if the eigenvectors are linearly independent.

Finding Eigenvalues/Eigenvectors

- For all of the above, we assumed we knew eigenvalues/eigenvectors. How do we find them in the first place?

Finding Eigenvalues/Eigenvectors

- For all of the above, we assumed we knew eigenvalues/eigenvectors. How do we find them in the first place?
- We'll use the following facts for a non-zero eigenvector v:

$$\begin{aligned} Av = \lambda v &\implies Av - \lambda v = 0 \\ &\implies (A - \lambda I)v = 0 \end{aligned}$$

and so the matrix $A - \lambda I$ is not invertible.

Finding Eigenvalues/Eigenvectors

- For all of the above, we assumed we knew eigenvalues/eigenvectors. How do we find them in the first place?
- We'll use the following facts for a non-zero eigenvector v:

$$\begin{aligned} Av = \lambda v &\implies Av - \lambda v = 0 \\ &\implies (A - \lambda I)v = 0 \end{aligned}$$

  and so the matrix $A - \lambda I$ is not invertible.
- Why? It has a nonzero vector v in its null space.

Finding Eigenvalues/Eigenvectors

- For all of the above, we assumed we knew eigenvalues/eigenvectors. How do we find them in the first place?
- We'll use the following facts for a non-zero eigenvector v:

$$\begin{aligned} Av = \lambda v &\implies Av - \lambda v = 0 \\ &\implies (A - \lambda I)v = 0 \end{aligned}$$

  and so the matrix $A - \lambda I$ is not invertible.
- Why? It has a nonzero vector v in its null space.
- Accordingly, $\det(A - \lambda I) = 0$.

The Characteristic Polynomial

The characteristic polynomial of a matrix A is defined to be

$$p_C(\lambda) = \det(A - \lambda I).$$

Finding the roots of this polynomial (i.e. the numbers λ such that $p_C(\lambda) = 0$ yields the eigenvalues of A.

Singular Value Decomposition

If $A \in \mathbb{R}^{m \times n}$, then there exists orthogonal

$$U = \begin{bmatrix} | & | & \cdots & | \\ u_1 & u_2 & \cdots & u_m \\ | & | & \vdots & | \end{bmatrix} \in \mathbb{R}^{m \times m}, V = \begin{bmatrix} | & | & \cdots & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & \vdots & | \end{bmatrix} \in \mathbb{R}^{n \times n}$$

and values $\sigma_1, \ldots, \sigma_p$ (where $p = \min\{n, m\}$) such that

$$A = U \begin{bmatrix} \sigma_1 & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \sigma_p \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} V^T = U \Sigma V^T$$

Challenge Problem

Given an SVD of the matrix $A \in \mathbb{R}^{n \times n}$, $A = U\Sigma V^T$, find an eigenvector decomposition of the matrix $A^T A$. That is, find matrices Q, Λ such that

$$A^T A = Q \Lambda Q^{-1}$$

You can use the fact that for an orthogonal matrix, the transpose is the same as the inverse.

Application: Least Squares

Problem Setup

- We are given n observations of *predictors* $x_i \in \mathbb{R}^p$, and *response* $y_i \in \mathbb{R}$ for $i = 1, \ldots, n$.

Problem Setup

- We are given n observations of *predictors* $x_i \in \mathbb{R}^p$, and *response* $y_i \in \mathbb{R}$ for $i = 1, \ldots, n$.
- Goal: find $\beta \in \mathbb{R}^p$ that minimizes the quantity

$$\sum_{i=1}^{n} |\beta^\top x_i - y_i|^2$$

Matrix Notation

- Rewrite the n response values $y_1, \ldots, y_n$ into a single vector $y \in \mathbb{R}^n$.

Matrix Notation

- Rewrite the n response values $y_1, \ldots, y_n$ into a single vector $y \in \mathbb{R}^n$.
- Rewrite all of the predictors x_i as rows in a matrix X given by

$$X = \begin{bmatrix} - - - & x_1^\top & - - - \\ - - - & x_2^\top & - - - \\ \vdots & \vdots & \vdots \\ - - - & x_n^\top & - - - \end{bmatrix} \in \mathbb{R}^{n \times p}$$

Matrix Notation

- Rewrite the n response values $y_1, \ldots, y_n$ into a single vector $y \in \mathbb{R}^n$.
- Rewrite all of the predictors x_i as rows in a matrix X given by

$$X = \begin{bmatrix} --- & x_1^\top & --- \\ --- & x_2^\top & --- \\ \vdots & \vdots & \vdots \\ --- & x_n^\top & --- \end{bmatrix} \in \mathbb{R}^{n \times p}$$

- Try It: Write the the problem on the previous slide in matrix notation.

Matrix Notation

- Rewrite the n response values $y_1, \ldots, y_n$ into a single vector $y \in \mathbb{R}^n$.
- Rewrite all of the predictors x_i as rows in a matrix X given by

$$X = \begin{bmatrix} --- & x_1^\top & --- \\ --- & x_2^\top & --- \\ \vdots & \vdots & \vdots \\ --- & x_n^\top & --- \end{bmatrix} \in \mathbb{R}^{n \times p}$$

- The problem can be rewritten as

$$\min_{\beta \in \mathbb{R}^p} ||X\beta - y||_2^2 = (X\beta - y)^\top (X\beta - y)$$

Take Gradient

- We want to solve the problem

$$\min_{\beta \in \mathbb{R}^p} (X\beta - y)^\top (X\beta - y)$$

Take Gradient

- We want to solve the problem

$$\min_{\beta \in \mathbb{R}^p} (X\beta - y)^\top (X\beta - y)$$

- This quantity is

$$\beta^\top X^\top X\beta - 2y^\top X\beta + y^\top y.$$

Take Gradient

- We want to solve the problem

$$\min_{\beta \in \mathbb{R}^p} (X\beta - y)^\top (X\beta - y)$$

- This quantity is

$$\beta^\top X^\top X\beta - 2y^\top X\beta + y^\top y.$$

- Try It: compute the gradient of this quantity with respect to β.

Take Gradient

- We want to solve the problem

$$\min_{\beta \in \mathbb{R}^p} (X\beta - y)^\top (X\beta - y)$$

- This quantity is

$$\beta^\top X^\top X\beta - 2y^\top X\beta + y^\top y.$$

- The gradient with respect to β is

$$2X^\top X\beta - 2X^\top y.$$

Set To Zero, Solve

- We set the gradient to zero, and solve for β:

$$\begin{aligned} 0 &= 2X^\top X\beta - 2X^\top y \\ \cancel{2}X^\top y &= \cancel{2}X^\top X\beta \\ (X^\top X)^{-1}X^\top y &= (X^\top X)^{-1}X^\top X\beta \end{aligned}$$

which implies

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

The End