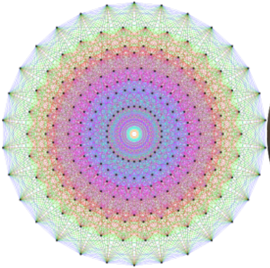
Further topics in Linear Algebra

4433LALG3: Linear Algebra

Week 3, Lecture 12, Valente Ramírez

Mathematical & Statistical Methods group — Biometris, Wageningen University & Research



Overview

- Abstract vector spaces
- Norms
- Inner products

References:

- 3Blue1Brown Ch.16

Section 1

Abstract vector spaces

Vector spaces

In this course, we have made extensive use of the vector notation $\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$.

But we have also emphasized that the "coordinates" are just a name.

What the notation *really* means is that we should use the numbers v_i to combine certain basis vectors:

$$\mathbf{v} = v_1\mathbf{e}_1 + \ldots + v_n\mathbf{e}_n$$

or

$$\mathbf{v} = v_1\mathbf{b}_1 + \ldots + v_n\mathbf{b}_n$$

or

. . . something else?

Vector spaces

Definition

A **vector space** is a set V of "objects" (we will call them vectors) such that:

- We can add them: if $\mathbf{v}, \mathbf{w} \in V$, the operation $\mathbf{v} + \mathbf{w}$ makes sense
- We can re-scale them: if $\mathbf{v} \in V$ and $\alpha \in \mathbb{R}$, then $\alpha\mathbf{v}$ makes sense
- There is a special vector $\mathbf{0}$ that serves as "zero-vector"
- The above operations satisfy *"the usual algebraic properties"*

Vector spaces

Rules for vectors addition and scaling

1. $\vec{\mathrm{u}} + (\vec{\mathrm{v}} + \vec{\mathrm{w}}) = (\vec{\mathrm{u}} + \vec{\mathrm{v}}) + \vec{\mathrm{w}}$
2. $\vec{\mathrm{v}} + \vec{\mathrm{w}} = \vec{\mathrm{w}} + \vec{\mathrm{v}}$
3. There is a vector $\mathbf{0}$ such that $\mathbf{0} + \vec{\mathrm{v}} = \vec{\mathrm{v}}$ for all $\vec{\mathrm{v}}$
4. For every vector $\vec{\mathrm{v}}$ there is a vector $-\vec{\mathrm{v}}$ so that $\vec{\mathrm{v}} + (-\vec{\mathrm{v}}) = \mathbf{0}$
5. $a(b\vec{\mathrm{v}}) = (ab)\vec{\mathrm{v}}$
6. $1\vec{\mathrm{v}} = \vec{\mathrm{v}}$
7. $a(\vec{\mathrm{v}} + \vec{\mathrm{w}}) = a\vec{\mathrm{v}} + a\vec{\mathrm{w}}$
8. $(a + b)\vec{\mathrm{v}} = a\vec{\mathrm{v}} + b\vec{\mathrm{v}}$

"Axioms"

Examples

1. We can add and re-scale matrices.
   - The space $\mathrm{Mat}(m, n)$ of all $m \times n$ matrices is a vector space.

2. We can add and re-scale functions.
   - The space $\mathcal{C}^0(\mathbb{R})$ of all continuous functions $f\colon \mathbb{R} \to \mathbb{R}$ is a vector space.

3. We can add and re-scale random variables.
   - The space $\mathbb{X}$ of all random variables is a vector space.

Linear transformations

Definition

A transformation between vector spaces $T\colon V \to W$ is **linear** if:

- $T(\mathbf{v}_1 + \mathbf{v}_2) = T(\mathbf{v}_1) + T(\mathbf{v}_2)$ for all $\mathbf{v}_1, \mathbf{v}_2 \in V$,
- $T(\alpha\mathbf{v}) = \alpha T(\mathbf{v})$ for all $\alpha \in \mathbb{R}$ and $\mathbf{v} \in V$.

Examples:

1. Transposition $\top\colon \mathrm{Mat}(m,n) \to \mathrm{Mat}(n,m)$ is a linear transformation.
2. Integration $\int_a^b\colon \mathcal{C}^0(\mathbb{R}) \to \mathbb{R}$ is a linear transformation.
3. Expectation $\mathbb{E}\colon \mathbb{X} \to \mathbb{R}$ is a linear transformation.

Basis and dimension

- The following matrices form a **basis** for $\mathrm{Mat}(2,2)$:

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

  Therefore the **dimension** of $\mathrm{Mat}(2,2)$ is 4.

- A **basis** for the space of *polynomial* functions is:

$$1, \quad x, \quad x^2, \quad x^3, \quad \ldots$$

  This space is **infinite dimensional**.

Subspaces

Definition

A **linear subspace** of a vector space V is a (non-empty) subset $S \subseteq V$ that is a vector space itself (i.e. with the operations inherited from V).

Examples:

1. The space of diagonal matrices is a subspace of $\mathrm{Mat}(n, n)$
2. The space of differentiable functions is a subspace of $\mathcal{C}^0(\mathbb{R})$.
3. The space of variables X for which $\mathbb{E}(X) = 0$ is a subspace of $\mathbb{X}$.

Kernel and image

Consider the following linear transformation on $\mathcal{C}(\mathbb{R})$:

$$T\colon \mathcal{C}(\mathbb{R}) \to \mathcal{C}(\mathbb{R}) \quad \text{such that} \quad T(f(x)) = \frac{f(x) + f(-x)}{2}.$$

- The **kernel** (null space) of T is the space of all *odd functions*:

  $$f(x) \text{ is odd if } f(-x) = -f(x).$$

- The **image** (column space) of T is the space of all *even functions*:

  $$g(x) \text{ is even if } g(-x) = g(x).$$

Eigenvalues and eigenvectors

Consider the derivative $\frac{d}{dx}$ as a linear transformation on the space of differentiable functions.

What are the eigenvalues and eigenvectors?

Set the usual equation: $\frac{d}{dx}f(x) = \lambda f(x)$.

- Any real number is an **eigenvalue**.
- The corresponding **eigenfunction** is $f(x) = e^{\lambda x}$.

Section 2

Norms

Norms

A **norm** is a way to measure the length (the size) of a vector.

This can be on $\mathbb{R}^n$ or for an abstract vector space.

Depending on the situation, a particular way of measuring length might be more convenient than another.

If you are on $\mathbb{R}^2$ and can only move horizontally and/or vertically, you might want to measure length using:

$$\|\mathbf{x}\|_1 = |x_1| + |x_2| \qquad \text{instead of} \qquad \|\mathbf{x}\|_2 = \sqrt{x_1^2 + x_2^2}.$$

If you are dealing with observations from a $\mathcal{N}(\mathbf{0}, \Sigma)$–distribution, you should measure length using:

$$d_\Sigma(\mathbf{x}) = \sqrt{\mathbf{x}^\top \Sigma^{-1} \mathbf{x}}.$$

Norms

It's OK to measure length in some alternative way, as long as a few *axioms* are satisfied.

Definition

A **norm** in a vector space V is a function $F\colon V \to \mathbb{R}$ satisfying:

- $F(\mathbf{x}) \geq 0$
- $F(\mathbf{x}) = 0$ only if $\mathbf{x} = \mathbf{0}$
- $F(\alpha\mathbf{x}) = |\alpha| F(\mathbf{x})$
- $F(\mathbf{x}+\mathbf{y}) \leq F(\mathbf{x}) + F(\mathbf{y})$

In practice, people write $\|\mathbf{x}\|_F$, or $\|\mathbf{x}\|_V$, or something similar. Sometimes we just write $\|\mathbf{x}\|$, if it is clear from the context **which** norm we are talking about.

Examples

There are many ways to measure how big a matrix is.

For example, consider $M \in \mathrm{Mat}(m, n)$.

- The *operator norm* is given by:

$$\|M\|_O = \max\left\{ \frac{\|M\mathbf{x}\|}{\|\mathbf{x}\|} \;\middle|\; \mathbf{x} \neq \mathbf{0} \right\}.$$

- The *spectral norm* is given by:

$$\|M\|_S = \sigma_{\mathsf{max}}(A), \quad \text{i.e. the largest singular value of } A.$$

Examples

There are many ways to measure how big a continuous function is.

Consider $\mathcal{C}^0([0,1])$, the space of continuous functions $f\colon [0,1] \to \mathbb{R}$.

- The *uniform norm* is given by:

$$\|M\|_U = \max \{ \, |f(x)| \ \text{ with } \ x \in [0,1] \, \} \, .$$

- The L^1–*norm* is given by:

$$\|f\|_1 = \int_0^1 |f(x)| \, \mathrm{d}x.$$

- The L^2–*norm* is given by:

$$\|f\|_2 = \sqrt{\int_0^1 (f(x))^2 \, \mathrm{d}x}.$$

Section 3

Inner products

Inner products

What about "alternative dot products"?

They also exist, and are called *inner products*.

Definition

An **inner product** on a vector space V is a map $\langle\ ,\ \rangle \colon V \times V \to \mathbb{R}$ satisfying the following axioms:

1. It is symmetric
   - That is, $\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle$
2. It is positive-definite
   - $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$ for all $\mathbf{v}$
   - $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ only if $\mathbf{v} = \mathbf{0}$
3. It is bilinear
   - $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$
   - $\langle \alpha\mathbf{v}, \mathbf{w} \rangle = \alpha\langle \mathbf{v}, \mathbf{w} \rangle$

Examples

- Let Q be a **positive-definite** symmetric matrix. Then Q defines an inner product on $\mathbb{R}^n$ via:

$$\langle \mathbf{x}, \mathbf{y} \rangle_Q = \mathbf{x}^\top Q \, \mathbf{y}.$$

- On $\mathrm{Mat}(m, n)$ we have the inner product:

$$\langle A, B \rangle = \mathrm{tr}(A^\top B).$$

- On $\mathcal{C}^0([0, 1])$ we have the inner product:

$$\langle f, g \rangle = \int_0^1 f(x) g(x) \, \mathrm{d}x.$$

This last can be thought of as **covariance between functions**.

Induced norms

Remark

Every inner product induces a norm.

Indeed, if you have an inner product $\langle \ , \ \rangle$ defined on V, you can set

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}.$$

If the axioms for inner products are satisfied, the axioms for norms are automatically satisfied as well.

However, not all norms come from an inner product!

Conclusion:

Inner products can provide a lot of additional structure on V, which can be extremely useful.

Orthogonality

Consider $\mathcal{C}^0([0,1])$ again, with the inner product $\langle f, g\rangle = \displaystyle\int_0^1 f(x)g(x)\,\mathrm{d}x$.

The functions $\cos(2\pi x)$ and $\sin(2\pi x)$ are orthogonal, because

$$\int_0^1 \cos(2\pi x)\sin(2\pi x)\,\mathrm{d}x = 0.$$

In fact, the following functions form an orthogonal set:

- 1
- $\cos(2\pi n x)$, for $n = 1, 2, 3, \ldots$
- $\sin(2\pi n x)$, for $n = 1, 2, 3, \ldots$

The **expansion** of a function $f \in \mathcal{C}^0([0,1])$ in terms of this orthogonal basis is called the **Fourier series** of f:

$$f(x) = a_0 + \sum_{n=1}^{\infty} \big(a_n \cos(2\pi n x) + b_n \sin(2\pi n x)\big).$$

Section 4

Supplementary material

Examination material

From this lecture, the only exam material is:

- The concept of an abstract vector space, and its subspaces

You are expected to be able to identify whether a given subset S of a vector space V is a linear subspace or not.

Examples

Example

Let $\mathbf{x}$ be a fixed vector in $\mathbb{R}^n$. Consider $S = \{A \in \mathrm{Mat}(m, n) \mid A\mathbf{x} = \mathbf{0}\}$.
Is this a linear subspace?

Let's see.

Consider two matrices $A, B \in S$, and $\alpha \in \mathbb{R}$.

We need to check whether $A + B$ and αA are also in S.

- $(A + B)\mathbf{x} = A\mathbf{x} + B\mathbf{x} = \mathbf{0} + \mathbf{0} = \mathbf{0}$. Therefore, $A + B \in S$.
- $(\alpha A)\mathbf{x} = \alpha(A\mathbf{x}) = \alpha\mathbf{0} = \mathbf{0}$. So $\alpha A \in S$.

We conclude that S is indeed a subspace.

Examples

Example

Let $\lambda \neq 0$ and $\mathbf{x} \in \mathbb{R}^n$ be fixed. Consider $S = \{A \in \mathrm{Mat}(m,n) \mid A\mathbf{x} = \lambda\mathbf{x}\}$. Is this a linear subspace?

Let's see.

Consider two matrices $A, B \in S$, and $\alpha \in \mathbb{R}$.

We need to check whether $A + B$ and αA are also in S.

- $(A+B)\mathbf{x} = A\mathbf{x} + B\mathbf{x} = \lambda\mathbf{x} + \lambda\mathbf{x} = 2\lambda\mathbf{x}$.

  Because $\lambda \neq 0$, we get that $2\lambda \neq \lambda$.

  Therefore, $(A+B)\mathbf{x} \neq \lambda\mathbf{x}$.

We conclude that S is <u>not</u> a subspace.

Exercises

In each of the following cases, you are given a vector space V and $S \subseteq V$.

Is S a linear subspace of V?

1. $V = \mathrm{Mat}(n, n)$ and S is the subset of upper triangular matrices.
2. $V = \mathrm{Mat}(n, n)$ and S is the subset of matrices with zero determinant.
3. $V = \mathbb{X}$ and $S = \{X \in \mathbb{X} \mid \mathrm{Cov}(X, Y) = 0\}$, Y a *fixed* random variable.
4. $V = \mathcal{C}^0([0, 1])$ and $S = \{f \in V \mid f(0) = 0\}$.
5. $V = \mathcal{C}^0([0, 1])$ and $S = \{f \in V \mid \int_0^1 f(x)\,\mathrm{d}x = 1\}$.