Overview

Last time we studied the evolution of a discrete linear dynamical system, and today we begin the final topic of the course (loosely speaking).

Today we'll recall the definition and properties of the dot product. In the next two weeks we'll try to answer the following questions:

Question

What is the relationship between diagonalisable matrices and vector projection? How can we use this to study linear systems without exact solutions?

From Lay, §6.1, 6.2

Motivation for the inner product

- A linear system $A\mathbf{x} = \mathbf{b}$ that arises from experimental data often has no solution. Sometimes an acceptable substitute for a solution is a vector $\hat{\mathbf{x}}$ that makes the distance between $A\hat{\mathbf{x}}$ and $\mathbf{b}$ as small as possible (you can see this $\hat{\mathbf{x}}$ as a good approximation of an actual solution). As the definition for distance involves a sum of squares, the desired $\hat{\mathbf{x}}$ is called a *least squares solution*.
- Just as the dot product on $\mathbb{R}^n$ helps us understand the geometry of Euclidean space with tools to detect angles and distances, the inner product can be used to understand the geometry of abstract vector spaces.

In this section we begin the development of the concepts of orthogonality and orthogonal projections; these will play an important role in finding $\hat{\mathbf{x}}$.

Recall the definition of the dot product:

Definition

The *dot* (or *scalar* or *inner*) *product of two vectors* $\mathbf{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}, \mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$ *in* $\mathbb{R}^n$ *is the scalar*

$$\begin{aligned}(\mathbf{u}, \mathbf{v}) &= \mathbf{u}\cdot\mathbf{v} = \mathbf{u}^T\mathbf{v} \\ &= \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + \cdots + u_n v_n .\end{aligned}$$

The following properties are immediate:

(a) $\mathbf{u}\cdot\mathbf{v} = \mathbf{v}\cdot\mathbf{u}$

(b) $\mathbf{u}\cdot(\mathbf{v} + \mathbf{w}) = \mathbf{u}\cdot\mathbf{v} + \mathbf{u}\cdot\mathbf{w}$

(c) $k(\mathbf{u}\cdot\mathbf{v}) = (k\mathbf{u})\cdot\mathbf{v} = \mathbf{u}\cdot(k\mathbf{v})$, $k \in \mathbb{R}$

(d) $\mathbf{u}\cdot\mathbf{u} \geq 0$, $\mathbf{u}\cdot\mathbf{u} = 0$ if and only if $\mathbf{u} = \mathbf{0}$.

Example 1

Consider the vectors

$$\mathbf{u} = \begin{bmatrix} 1 \\ 3 \\ -2 \\ 4 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} -1 \\ 0 \\ 3 \\ -2 \end{bmatrix}$$

Then

$$\begin{aligned} \mathbf{u} \cdot \mathbf{v} &= \mathbf{u}^T \mathbf{v} \\ &= \begin{bmatrix} 1 & 3 & -2 & 4 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ 3 \\ -2 \end{bmatrix} \\ &= (1)(-1) + (3)(0) + (-2)(3) + (4)(-2) \\ &= -15 \end{aligned}$$

The length of a vector

For vectors in $\mathbb{R}^3$, the dot product recovers the length of the vector:

$$\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \sqrt{u_1^2 + u_2^2 + u_3^2}.$$

We can use the dot product to define the length of a vector in an arbitrary Euclidean space.

Definition

For $\mathbf{u} \in \mathbb{R}^n$, the *length* of $\mathbf{u}$ is

$$\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \sqrt{u_1^2 + \cdots + u_n^2}.$$

It follows that for any scalar c, the length of $c\mathbf{v}$ is $|c|$ times the length of $\mathbf{v}$:

$$\|c\mathbf{v}\| = |c|\|\mathbf{v}\|.$$

Unit Vectors

A vector whose length is 1 is called a **unit vector**
If $\mathbf{v}$ is a non-zero vector, then

$$\mathbf{u} = \frac{\mathbf{v}}{||\mathbf{v}||}$$

is a unit vector in the direction of $\mathbf{v}$. To see this, compute

$$\begin{aligned} ||\mathbf{u}||^2 &= \mathbf{u} \cdot \mathbf{u} \\ &= \frac{\mathbf{v}}{||\mathbf{v}||} \cdot \frac{\mathbf{v}}{||\mathbf{v}||} \\ &= \frac{1}{||\mathbf{v}||^2} \mathbf{v} \cdot \mathbf{v} \\ &= \frac{1}{||\mathbf{v}||^2} ||\mathbf{v}||^2 \\ &= 1 \end{aligned} \tag{1}$$

Replacing $\mathbf{v}$ by the unit vector $\frac{\mathbf{v}}{||\mathbf{v}||}$ is called *normalising* $\mathbf{v}$.

Example 2

Find the length of $\mathbf{u} = \begin{bmatrix} 1 \\ -3 \\ 0 \\ 2 \end{bmatrix}$.

$$||\mathbf{u}|| = \sqrt{\mathbf{u}\cdot\mathbf{u}} = \sqrt{\left(\begin{bmatrix} 1 \\ -3 \\ 0 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -3 \\ 0 \\ 2 \end{bmatrix}\right)} = \sqrt{1+9+4} = \sqrt{14}.$$

Orthogonal vectors

The concept of perpendicularity is fundamental to geometry. The dot product generalises the idea of perpendicularity to vectors in $\mathbb{R}^n$.

Definition

The vectors $\mathbf{u}$ *and* $\mathbf{v}$ *are* **orthogonal** *to each other if* $\mathbf{u}\cdot\mathbf{v} = 0$.

Since $\mathbf{0}\cdot\mathbf{v} = 0$ for every vector $\mathbf{v}$ in $\mathbb{R}^n$, the zero vector is orthogonal to every vector.

Orthogonal complements

Definition

Suppose W is a subspace of $\mathbb{R}^n$. If the vector $\mathbf{z}$ is orthogonal to every $\mathbf{w}$ in W, then $\mathbf{z}$ is *orthogonal to* W.

Example 3

The vector $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ is orthogonal to $W = \text{Span}\left\{\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right\}$.

Example 4

We can also see that $\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ is orthogonal to $\text{Nul}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}$.

Definition

The set of all vectors $\mathbf{x}$ that are orthogonal to W is called the *orthogonal complement* of W and is denoted by $W^\perp$.

$$W^\perp = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \cdot \mathbf{y} = 0 \text{ for all } \mathbf{y} \in W\}$$

From the basic properties of the inner product it follows that

- A vector $\mathbf{x}$ is in $W^\perp$ if and only if $\mathbf{x}$ is orthogonal to every vector in a set that spans W.
- $W^\perp$ is a subspace
- $W \cap W^\perp = \mathbf{0}$ since $\mathbf{0}$ is the only vector orthogonal to itself.

Example 5

Let $W = \text{Span}\left\{\begin{bmatrix}1\\2\\-1\end{bmatrix}\right\}$. Find a basis for $W^\perp$, the orthogonal complement of W.

$W^\perp$ consists of all the vectors $\begin{bmatrix}x\\y\\z\end{bmatrix}$ for which

$$\begin{bmatrix}1\\2\\-1\end{bmatrix} \cdot \begin{bmatrix}x\\y\\z\end{bmatrix} = 0.$$

For this we must have $x + 2y - z = 0$, which gives $x = -2y + z$.

Thus

$$\begin{bmatrix}x\\y\\z\end{bmatrix} = \begin{bmatrix}-2y+z\\y\\z\end{bmatrix} = y\begin{bmatrix}-2\\1\\0\end{bmatrix} + z\begin{bmatrix}1\\0\\1\end{bmatrix}.$$

So a basis for $W^\perp$ is given by

$$\left\{\begin{bmatrix}-2\\1\\0\end{bmatrix}, \begin{bmatrix}1\\0\\1\end{bmatrix}\right\}.$$

Since $W = \text{Span}\left\{\begin{bmatrix}1\\2\\-1\end{bmatrix}\right\}$, we can check that every vector in $W^\perp$ is orthogonal to every vector in W.

Example 6

Let $V = \text{Span}\left\{ \begin{bmatrix} 1 \\ 3 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ -1 \\ -1 \\ 3 \end{bmatrix} \right\}$. Find a basis for $V^\perp$.

$V^\perp$ consists of all the vectors $\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$ in $\mathbb{R}^4$ that satisfy the two conditions

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 3 \\ 3 \\ 1 \end{bmatrix} = 0 \quad \text{and} \quad \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \cdot \begin{bmatrix} 3 \\ -1 \\ -1 \\ 3 \end{bmatrix} = 0$$

This gives a homogeneous system of two equations in four variables:

$$\begin{array}{rrrrl} a & +3b & +3c & +d & = 0 \\ 3a & -b & -c & +3d & = 0 \end{array}$$

Row reducing the augmented matrix we get

$$\left[\begin{array}{rrrr|r} 1 & 3 & 3 & 1 & 0 \\ 3 & -1 & -1 & 3 & 0 \end{array}\right] \to \left[\begin{array}{rrrr|r} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{array}\right]$$

So c and d are free variables and the general solution is

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} -d \\ -c \\ c \\ d \end{bmatrix} = d \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix} + c \begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}$$

The two vectors in the parametrisation above are linearly independent, so a basis for $V^\perp$ is

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

Notice that in the previous example (and also in the one before it) we found the orthogonal complement as the null space of a matrix.
We have

$$V^\perp = \text{Nul } A$$

where

$$A = \begin{bmatrix} 1 & 3 & 3 & 1 \\ 3 & -1 & -1 & 3 \end{bmatrix}$$

is the matrix whose ROWS are the transpose of the column vectors in the spanning set for V.
To find a basis for the null space of this matrix we just proceeded as usual by bringing the augmented matrix for $A\mathbf{x} = \mathbf{0}$ to reduced row echelon form.

Theorem

Let A be an $m \times n$ matrix.
The orthogonal complement of the row space of A is the null space of A.
The orthogonal complement of the column space of A is the null space of A^T.

$$(Row\ A)^\perp = Nul\ A \quad \text{and} \quad (Col\ A)^\perp = Nul\ A^T.$$

(Remember, Row A is the span of the rows of A.)
Proof The calculation for computing $A\mathbf{x}$ (multiply each row of A by the column vector $\mathbf{x}$) shows that if $\mathbf{x}$ is in Nul A, then $\mathbf{x}$ is orthogonal to each row of A. Since the rows of A span the row space, $\mathbf{x}$ is orthogonal to every vector in Row$A$.
Conversely, if $\mathbf{x}$ is orthogonal to Row A, then $\mathbf{x}$ is orthogonal to each row of A, and hence $A\mathbf{x} = \mathbf{0}$.
The second statement follows since Row A^T = Col A.

Example 7

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

- Then Row A = Span $\left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \right\}$.
- Nul A = Span $\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$

Hence $(\text{Row } A)^\perp = \text{Nul } A$.

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

- Col A = Span $\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$.
- Nul A^T = Span $\left\{ \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right\}$.

Clearly, $(\text{Col } A)^\perp = \text{Nul } A^T$.

An important consequence of the previous theorem.

Theorem

If W is a subspace of $\mathbb{R}^n$, then $\dim W + \dim W^\perp = n$

Choose vectors $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_p$ such that $W = \text{Span}\{\mathbf{w}_1, \dots, \mathbf{w}_p\}$. Let

$$A = \begin{bmatrix} \mathbf{w}_1^T \\ \mathbf{w}_2^T \\ \vdots \\ \mathbf{w}_p^T \end{bmatrix}$$

be the matrix whose rows are $\mathbf{w}_1^T, \dots, \mathbf{w}_p^T$.
Then $W = \text{Row}\, A$ and $W^\perp = (\text{Row}\, A)^\perp = \text{Nul}\ A$. Thus

$$\dim W = \dim(\text{Row}\, A) = \text{Rank}\, A$$

$$\dim W^\perp = \dim(\text{Nul}\ A)$$

and the Rank Theorem implies

$$\dim W + \dim W^\perp = \text{Rank}\, A + \dim(\text{Nul}\ A) = n$$

Example 8

Let $W = \text{Span}\left\{\begin{bmatrix}1\\4\\3\end{bmatrix}\right\}$. Describe $W^\perp$.

We see first that $\dim W = 1$ and W is a *line* through the origin in $\mathbb{R}^3$.
Since we must have $\dim W + \dim W^\perp = 3$, we can then deduce that $\dim W^\perp = 2$: $W^\perp$ is a *plane* through the origin.
In fact, $W^\perp$ is the set of all solutions to the homogeneous equation coming from this equation:

$$\begin{bmatrix}x\\y\\z\end{bmatrix} \cdot \begin{bmatrix}1\\4\\3\end{bmatrix} = 0.$$

That is,

$$x + 4y + 3z = 0\,.$$

We recognise this as the equation of the plane through the origin in $\mathbb{R}^3$ with normal vector $\langle 1, 4, 3\rangle = \mathbf{w}$.

Basis Theorem

Theorem

If $\mathcal{B} = \{\mathbf{b_1}, \dots, \mathbf{b_m}\}$ is a basis for W and $\mathcal{C} = \{\mathbf{c_1}, \dots, \mathbf{c_r}\}$ is a basis for $W^\perp$, then $\{\mathbf{b_1}, \dots, \mathbf{b_m}, \mathbf{c_1}, \dots, \mathbf{c_r}\}$ is a basis for $\mathbb{R}^{m+r}$.

It follows that if W is a subspace of $\mathbb{R}^n$, then for any vector $\mathbf{v}$, we can write

$$\mathbf{v} = \mathbf{w} + \mathbf{u},$$

where $\mathbf{w} \in W$ and $\mathbf{u} \in W^\perp$.

If W is the span of a nonzero vector in $\mathbb{R}^3$, then w is just the vector projection of v onto this spanning vector.

Example 9

Let $W = \text{Span}\left\{\begin{bmatrix}1\\1\\0\\1\end{bmatrix}, \begin{bmatrix}1\\1\\1\\0\end{bmatrix}\right\}$. Decompose $\mathbf{v} = \begin{bmatrix}2\\1\\1\\3\end{bmatrix}$ as a sum of vectors in W and $W^\perp$.

To start, we find a basis for $W^\perp$ and then write $\mathbf{v}$ in terms of the bases for W and $W^\perp$.

We're given a basis for W in the problem, and

$$W^\perp = \text{Span}\left\{\begin{bmatrix}1\\-1\\0\\0\end{bmatrix}, \begin{bmatrix}1\\0\\-1\\-1\end{bmatrix}\right\}$$

Therefore $\mathbf{v} = 2\left(\begin{bmatrix}1\\1\\0\\1\end{bmatrix}\right) + \left(\begin{bmatrix}1\\-1\\0\\0\end{bmatrix} - \begin{bmatrix}1\\0\\-1\\-1\end{bmatrix}\right) = \begin{bmatrix}2\\2\\0\\2\end{bmatrix} + \begin{bmatrix}0\\-1\\1\\1\end{bmatrix}$.