5.1: Orthogonal Projections & Orthonormal Bases

Class summary to date.

Three main motivations for linear algebra

(1) applications

(2) challenge common conceptions in math.

(3) Intro to mathematical abstraction & reasoning.

This outline skips ch 6 & 7

Course outline to date

(1) Linear Equations

(2) Linear Transformations.

(3) Subspaces of $\mathbb{R}^n$ and their dimension.

→ image & kernel of a L.T.

→ bases & L.I.

→ Dimension

→ coordinates.

So where are we going? In our next chapter we will focus on a special type/class of bases (orthonormal), how to find them, and their applications (least squares).

Basic vector concepts.

(a) perpendicular/orthogonal vectors.

(b) length/magnitude/norm

(c) unit vectors

(d) finding unit vectors.

(e) orthonormal vectors.

$\vec{u}_1, \dots, \vec{u}_m \in \mathbb{R}^n$ are <u>orthonormal</u> if

$$\vec{u}_i \cdot \vec{u}_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

<u>ex1</u>: The standard vecs.

<u>ex2</u>: The cols of the 2D rotation matrix

<u>ex3</u>: $\begin{bmatrix} 2/3 \\ 2/3 \\ 1/3 \end{bmatrix}, \begin{bmatrix} -2/3 \\ 1/3 \\ 2/3 \end{bmatrix}, \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix}$ (show they are orthonormal)

<u>Thm</u>: properties of orthonormal vecs.

(a) orthonormal vecs are L.I.

(b) orthonormal vecs $\vec{u}_1, \dots, \vec{u}_n$ in $\mathbb{R}^n$ form a basis for $\mathbb{R}^n$.

□ proof.

Suppose $\vec{u}_1, \dots, \vec{u}_m \in \mathbb{R}^n$ are orthonormal vecs. To show that these are L.I. we must show

$$c_1\vec{u}_1 + \dots + c_m\vec{u}_m = \vec{0}$$ has only the trivial sol.

□ proof

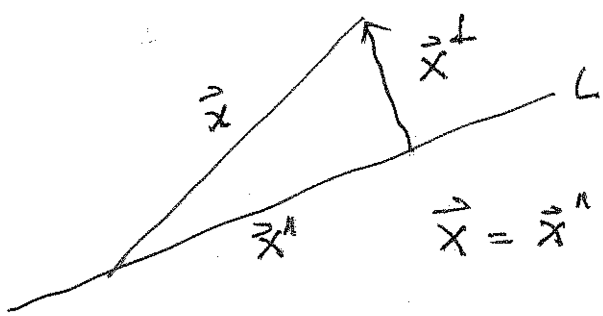
consider $u_i \cdot (c_1\vec{u}_1 + \dots + c_m\vec{u}_m) = \vec{u}_i \cdot \vec{0}$ for $i = 1, \dots, m$

$$\Rightarrow c_1 u_i \cdot u_1 + \dots + c_{i-1} u_i \cdot u_{i-1} + c_i u_i \cdot u_i + c_{i+1} u_i \cdot u_{i+1} + \dots + c_m u_i \cdot u_m = \vec{0}$$

$\Rightarrow c_i = 0$ for $i = 1, \dots, m$

Hence $\vec{u}_1, \dots, \vec{u}_m$ are L.I. QED ■

orthogonal projections



$$\vec{x} = \vec{x}^\parallel + \vec{x}^\perp$$

Thm: consider a vector $\vec{x} \in \mathbb{R}^n$ and a subspace V of $\mathbb{R}^n$. Then we can write $\vec{x} = \vec{x}^\parallel + \vec{x}^\perp$ where $\vec{x}^\parallel \in V$ and $\vec{x}^\perp$ is orthogonal to V. This representation is unique.

□ proof.

Consider an orthonormal basis $\vec{u}_1, \ldots, \vec{u}_m$ of V.

If $\vec{x} = \vec{x}^{\parallel} + \vec{x}^{\perp}$ exists, then

$$x^{\parallel} = c_1 \vec{u}_1 + \ldots + c_i \vec{u}_i + \ldots + c_m \vec{u}_m$$

for yet to be determined coefficients.

$\Rightarrow \quad \vec{x}^{\perp} = \vec{x} - \vec{x}^{\parallel} = \vec{x} - c_1 \vec{u}_1 - \ldots - c_i \vec{u}_i - \ldots - c_m \vec{u}_m$

is orthogonal to all $\vec{u}_i \in V$.

$\Rightarrow \quad 0 = \vec{u}_i \cdot (\vec{x} - c_1 \vec{u}_1 - \ldots - c_i \vec{u}_i - \ldots - c_m \vec{u}_m)$

$\Rightarrow \quad 0 = \vec{u}_i \cdot \vec{x} - c_i \quad$ for $i = 1, 2, 3, \ldots, m$

$\Rightarrow \quad c_i = \vec{u}_i \cdot \vec{x} \quad$ for $i = 1, 2, 3, \ldots, m$

$\Rightarrow \quad \vec{x}^{\parallel} = (\vec{u}_1 \cdot \vec{x}) \vec{u}_1 + \ldots + (\vec{u}_m \cdot \vec{x}) \vec{u}_m$

and $\vec{x}^{\perp} = \vec{x} - \vec{x}^{\parallel}$.

It is unique by construction. ■

Thus if V is a subspace of $\mathbb{R}^n$ w/ orthonormal basis $\vec{u}_1, \ldots, \vec{u}_m$ then $\mathrm{proj}_V(\vec{x}) = \vec{x}^{\parallel} = (\vec{u}_1 \cdot \vec{x}) \vec{u}_1 + \ldots + (\vec{u}_m \cdot \vec{x}) \vec{u}_m$ for all $\vec{x}$ in $\mathbb{R}^n$.

ex4: consider the subspace $V = \mathrm{im}(A)$ of $\mathbb{R}^3$ where

$A = \begin{bmatrix} 2 & -2 \\ 2 & 1 \\ 1 & 2 \end{bmatrix}$. Find $\mathrm{proj}_V(\vec{x})$ for $\vec{x} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}$

$\vec{u}_1 = \begin{bmatrix} 2/3 \\ 2/3 \\ 1/3 \end{bmatrix}$ and $\vec{u}_2 = \begin{bmatrix} -2/3 \\ 1/3 \\ 2/3 \end{bmatrix}$ $\quad x^{\parallel} = (5) \begin{bmatrix} 2/3 \\ 2/3 \\ 1/3 \end{bmatrix} + (1) \begin{bmatrix} -2/3 \\ 1/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 8/3 \\ 11/3 \\ 7/3 \end{bmatrix}$

check that $\vec{x}^{\perp}$ is $\perp$ to $\vec{u}_1$ & $\vec{u}_2$

Thm: Consider an orthonormal basis $\vec{u}_1, \ldots, \vec{u}_n \in \mathbb{R}^n$.
Then $\vec{x} = (\vec{u}_1 \cdot \vec{x})\vec{u}_1 + \ldots + (\vec{u}_n \cdot \vec{x})\vec{u}_n$ for all $\vec{x} \in \mathbb{R}^n$.

Recall, if $\vec{v}_1, \ldots, \vec{v}_n \in \mathbb{R}^n$ is a basis for $\mathbb{R}^n$, Then $c_1, \ldots, c_n$ s.t. $\vec{x} = c_1\vec{v}_1 + \ldots + c_n\vec{v}_n$ are the coordinates of $\vec{x}$.

If $\vec{u}_1, \ldots, \vec{u}_n \in \mathbb{R}^n$ are an orthonormal basis, then the coords are easy to find: $c_i = \vec{u}_i \cdot \vec{x}$

Dfn: Consider a subspace V of $\mathbb{R}^n$. The <u>orthogonal complement</u> $V^\perp$ of V is the set of those vectors $\vec{x} \in \mathbb{R}^n$ that are orthogonal to all vecs in V.

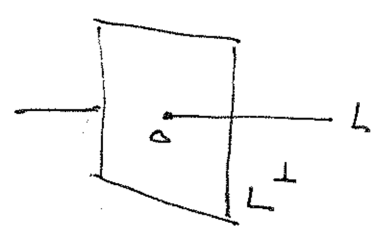
$$V^\perp = \{\vec{x} \in \mathbb{R} \mid \vec{v} \cdot \vec{x} = 0 \text{ for all } \vec{v} \in V\}$$

<u>Note</u>: $V^\perp$ is the kernel of the orthogonal projection onto V.

Thm: Consider a subspace V of $\mathbb{R}^n$.

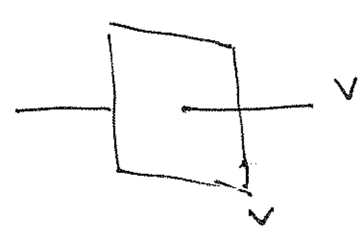
(a) the orthogonal complement $V^\perp$ of V is a subspace of $\mathbb{R}^n$.

(b) The intersection $V \cap V^\perp = \{\vec{0}\}$



(c) $\dim(V) + \dim(V^\perp) = n$ (see pics).

(d) $(V^\perp)^\perp = V$



Derivation of the Cauchy-Schwarz Inequality.

Let $\vec{y}$ be a vector in the direction of the line L. and $\vec{u} = \frac{\vec{y}}{\|\vec{y}\|}$.

$$\|\vec{x}\| \geq \|proj_L \vec{x}\|$$

$$= \|(\vec{x} \cdot \vec{u})\vec{u}\|$$

$$= |\vec{x} \cdot \vec{u}| \, \|\vec{u}\|$$

$$= |\vec{x} \cdot \vec{u}|$$

$$= \left| \vec{x} \cdot \frac{\vec{y}}{\|\vec{y}\|} \right|$$

$$= \frac{1}{\|\vec{y}\|} |\vec{x} \cdot \vec{y}|$$

Triangle Inequality

$$\|x+y\| \leq \|x\| + \|y\|$$

$\Rightarrow \quad \|\vec{x}\| \|\vec{y}\| \geq |\vec{x} \cdot \vec{y}|$ (equal only if $\vec{x}$ & $\vec{y}$ are parallel).

The angle between vectors. $(\vec{x}, \vec{y} \neq \vec{0})$

$$\vec{x} \cdot \vec{y} = \|\vec{x}\| \, \|\vec{y}\| \cos\theta$$

$$\Rightarrow \quad \theta = \arccos\left[\frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\| \, \|\vec{y}\|}\right]$$

This is always defined by Cauchy-Schwarz

$$\uparrow$$

$$-1 \leq \frac{x \cdot y}{\|\vec{x}\|\|\vec{y}\|} \leq 1$$