6.2 Orthogonal Sets

Defn $S=\{u_1,\dots,u_p\}$ a set of vectors in $\mathbb{R}^n$ are said to be an orthogonal set if every pair of distinct vectors is orthogonal, ie $u_i\cdot u_j=0\ \forall i\neq j$

EX1 $u_1=\begin{bmatrix}1\\0\\1\end{bmatrix}$ $u_2=\begin{bmatrix}-1\\4\\1\end{bmatrix}$ $u_3=\begin{bmatrix}2\\1\\-2\end{bmatrix}$ is an orthogonal set see fig. 1 on pg 384

$u_1\cdot u_2=-1+1=0$

$u_1\cdot u_3=2-2=0$

$u_2\cdot u_3=-2+4-2=0$

Thm 4
If $S=\{u_1,\dots,u_p\}$ is an orthogonal set of nonzero vectors in $\mathbb{R}^n$, then S is lin. indep., and thus a basis for the set spanned by S.

pf suppose dependent, then $0=c_1u_1+\dots+c_pu_p$ some $c_1,\dots,c_p$ not all zero, taking a dot prod. w/ u_1

$0=0\cdot u_1=c_1u_1\cdot u_1+\dots+c_pu_p\cdot u_1=c_1u_1\cdot u_1$ b/c $u_i\cdot u_j=0\ \forall i\neq j$

but if $0=c_1u_1\cdot u_1$ thus $c_1=0$ (b/c u_1 is not zero)

taking dot prod. w/ u_2 shows $c_2=0$ and so on for all c_i ↯ not all were zero. so S is lin. indep.

Defn An orthogonal basis for a subspace W of $\mathbb{R}^n$ is a basis for W that is also an orthogonal set.

Compute wts in linear combinations easily for orthogonal bases.

Thm 5 let $\{u_1,\dots,u_p\}$ be an orthogonal basis for W subspace of $\mathbb{R}^n$.
then $\forall y\in W$, $y=c_1u_1+\dots+c_pu_p$ w/ wts. $c_1,\dots,c_p$ given by $c_j=\dfrac{y\cdot u_j}{u_j\cdot u_j}$ $j=1,\dots,p$.

pf by orthogonality (as in last pf) $y\cdot u_j=c_ju_j\cdot u_j \Rightarrow c_j=\dfrac{y\cdot u_j}{u_j\cdot u_j}$ b/c $u_j\cdot u_j\neq 0$ (b/c orthogonal set basis can't have zero vector)

EX2 set $S=\{u_1,u_2,u_3\}$ from ex1 is an orthogonal basis for $\mathbb{R}^3$. express $y=\begin{bmatrix}2\\9\\2\end{bmatrix}$ as a lin comb. of vectors in S.

$y\cdot u_1=4$ $\quad y\cdot u_2=36$ $\quad y\cdot u_3=9$

$u_1\cdot u_1=2$ $\quad u_2\cdot u_2=18$ $\quad u_3\cdot u_3=9$

so $y=\frac{4}{2}u_1+\frac{36}{18}u_2+\frac{9}{9}u_3$

$y=2u_1+2u_2+u_3$

If basis is not orthogonal, then need to solve system.

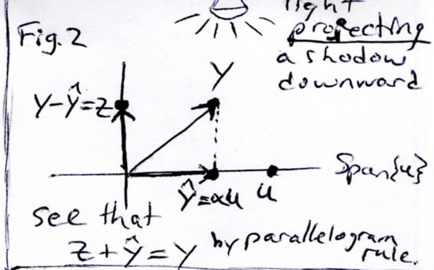
check work $2\begin{bmatrix}1\\0\\1\end{bmatrix}+2\begin{bmatrix}-1\\4\\1\end{bmatrix}+\begin{bmatrix}2\\1\\-2\end{bmatrix}=\begin{bmatrix}2\\9\\2\end{bmatrix}$

Bonus
let $\mathcal{B}=\{u_1,\dots,u_n\}$ be an orthogonal basis for $\mathbb{R}^n$ for any $x\in\mathbb{R}^n$ what is $[x]_{\mathcal{B}}$, the $\mathcal{B}$-coords of x?

$[x]_{\mathcal{B}}=\begin{bmatrix}c_1\\\vdots\\c_n\end{bmatrix}$ where $x=c_1u_1+\dots+c_nu_n$ and by thm 5 we know $c_j=\dfrac{x\cdot u_j}{u_j\cdot u_j}$ b/c $\mathcal{B}$ is an orthogonal basis.

so $[x]_{\mathcal{B}}=\begin{bmatrix}\frac{x\cdot u_1}{u_1\cdot u_1}\\\vdots\\\frac{x\cdot u_n}{u_n\cdot u_n}\end{bmatrix}$

Orthogonal Projections



given nonzero $u \in \mathbb{R}^n$, Span$\{u\}$ is a line

the orthogonal complement is an $(n-1)$ dim'l subspace consisting of all vectors orthogonal to u.

For any vector $y \in \mathbb{R}^n$, we want to decompose y into the sum of 2 vectors, one on the line spanned by u, and the other in the orthogonal complement.

$y = \hat{y} + z$ where $\hat{y} = \alpha u$. thus $z = y - \hat{y} = y - \alpha u$

and z should be orthogonal to u, so $(y - \alpha u)\cdot u = 0$

to find $\hat{y}$ and z (the decomposition of y, we seek) we need to solve for α.

$$(y-\alpha u)\cdot u = y\cdot u - \alpha\, u\cdot u = 0 \quad \Rightarrow \quad \frac{y\cdot u}{u\cdot u} = \alpha$$

and thus $\hat{y} = \dfrac{y\cdot u}{u\cdot u}u$ and $z = y - \dfrac{y\cdot u}{u\cdot u}u$

defn $\hat{y}$ is called the orthogonal projection of y onto u.
z is called the component of y orthogonal to u.

Note that if we used cu instead of u (for any nonzero scalar $c \in \mathbb{R}$), we'd get the same $\hat{y}$ and z.
So really we are projecting y onto the line $L = \text{Span}\{u\}$
$\therefore$ sometimes we write $\text{proj}_L y$ instead of $\hat{y}$ and call the orthogonal projection of $\hat{y}$ onto u the orthogonal projection of y onto L.

Formula to remember
$$\boxed{\hat{y} = \text{proj}_L y = \frac{y\cdot u}{u\cdot u}u}$$
where $L = \text{Span}\{u\}$ is the orthogonal projection of y onto L or u.

EX3 consider $\mathbb{R}^2$

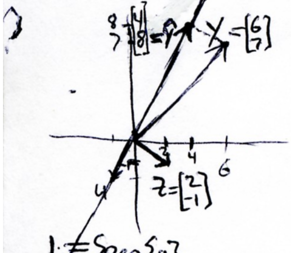
$y = \begin{bmatrix} 6 \\ 7 \end{bmatrix}$ $u = \begin{bmatrix} -1 \\ -2 \end{bmatrix}$ Find the orthogonal projection of y onto u.
Rewrite (or decompose) y into the sum of 2 vectors, one in Span$\{u\}$ and the other orthogonal to u.

$y\cdot u = -20$

$u \cdot u = 5$

on $L=\text{Span}\{u\}$: $\hat{y} = \alpha u = -4u$; orthogonal to u $(u\cdot z = 0)$

so $\hat{y} = \dfrac{-20}{5}u = \begin{bmatrix} 4 \\ 8 \end{bmatrix}$ and $z = \begin{bmatrix} 6 \\ 7 \end{bmatrix} - \begin{bmatrix} 4 \\ 8 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, thus $\underbrace{\begin{bmatrix} 6 \\ 7 \end{bmatrix}}_{y} = \underbrace{\begin{bmatrix} 4 \\ 8 \end{bmatrix}}_{\hat{y}} + \underbrace{\begin{bmatrix} 2 \\ -1 \end{bmatrix}}_{z = y-\hat{y}}$



as a check see that $\{\hat{y}, y-\hat{y}\} = \{\hat{y}, z\}$ is an orthogonal set
i.e. $\hat{y}\cdot z = \begin{bmatrix} 4 \\ 8 \end{bmatrix}\cdot\begin{bmatrix} 2 \\ -1 \end{bmatrix} = 4(2) + 8(-1) = 0$ ✓

Note that b/c of geometry the point $\hat{y}$ is the closest point on L to y.
This is the important concept we will use orthogonal projections to find least-squares solns.

EX 4 What is the distance from y to $L = \text{Span}\{u\}$.

From geometry, this is the length of the perpendicular line segment from y to $\hat{y}$, the orthogonal projection of y onto L.

This is the length of $z = y - \hat{y}$, thus $\text{dist}(y, \hat{y}) = \|y - \hat{y}\| = \sqrt{2^2 + (-1)^2} = \sqrt{5}$.

Recall Thm 5 says w/ an orthogonal basis $\{u_1, \dots, u_p\}$ for a subspace W in $\mathbb{R}^n$

every $y \in W$, is a linear combination $y = c_1 u_1 + \dots + c_p u_p$

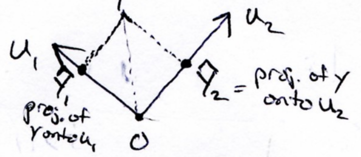
where the weights are $c_j = \dfrac{y \cdot u_j}{u_j \cdot u_j}$ for $j = 1, \dots, p$

but then $c_j u_j = \dfrac{y \cdot u_j}{u_j \cdot u_j} u_j$ is the orthogonal projection of y onto u_j.

So the linear combination in thm 5 decomposes y into a sum of orthogonal projections onto 1-dimensional subspaces, the basis elements (or lines they span).

EX Consider an orthogonal basis $\{u_1, u_2\}$ for $\mathbb{R}^2$. Every $y \in \mathbb{R}^2$ can be written

$$y = \underbrace{\frac{y \cdot u_1}{u_1 \cdot u_1} u_1}_{\text{projection of } y \text{ onto } u_1} + \underbrace{\frac{y \cdot u_2}{u_2 \cdot u_2} u_2}_{\text{projection of } y \text{ onto } u_2}$$



This easily generalizes to $\mathbb{R}^n$ and even subspaces W of $\mathbb{R}^n$ w/ orthogonal bases $\{u_1, \dots, u_p\}$ where $\forall y \in W$ $\hat{y}_j = \dfrac{y \cdot u_j}{u_j \cdot u_j} u_j$ and $y = \hat{y}_1 + \dots + \hat{y}_p$ by thm. 5.

other common examples are with forces in physics.

Orthonormal Sets

$\{u_1, \dots, u_p\}$ is an orthonormal set if it is an orthogonal set of unit vectors. That is $u_i \cdot u_j = 0$ $\forall i, j \in \{1, \dots, p\}$ and $\|u_\ell\| = 1$

if $W = \text{Span}\{u_1, \dots, u_p\}$, we say the orthonormal set is an orthonormal basis for W. Independent b/c orthogonal set + theorem 4.

The standard basis $\mathcal{E}_n = \{e_1, \dots, e_n\}$ is an orthonormal basis for $\mathbb{R}^n$.

any non empty subset of $\mathcal{E}_n$ is an orthonormal set as well.

EX 5 a non-~~standard~~ standard example of an orthonormal basis. Show $\{v_1, v_2, v_3\}$ is an orthonormal basis for $\mathbb{R}^3$

if $v_1 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$ $v_2 = \begin{bmatrix} -2/\sqrt{6} \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$ $v_3 = \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$

soln $v_1 \cdot v_2 = -2/\sqrt{18} + 1/\sqrt{18} + 1/\sqrt{18} = 0$

$v_1 \cdot v_3 = 0 + 1/\sqrt{6} - 1/\sqrt{6} = 0$

$v_2 \cdot v_3 = 0 + 1/\sqrt{12} - 1/\sqrt{12} = 0$

so orthogonal set and

$\|v_1\| = v_1 \cdot v_1 = \frac{1}{3} + \frac{1}{3} + \frac{1}{3} = 1$

$\|v_2\|^2 = v_2 \cdot v_2 = \frac{4}{6} + \frac{1}{6} + \frac{1}{6} = 1$ ✓

$\|v_3\|^2 = v_3 \cdot v_3 = 0 + \frac{1}{2} + \frac{1}{2} = 1$

thus $\{v_1, v_2, v_3\}$ is an orthonormal basis!

EX 5.5

Given an orthogonal set, we can get an orthonormal set by normalizing.

(normalization will not effect orthogonality)

given the orthogonal set (basis) in example 1, construct an orthonormal set (basis) by normalizing each vector.

recall from ex1

$u_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ $u_2 = \begin{bmatrix} -1 \\ 4 \\ 1 \end{bmatrix}$ $u_3 = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$

so $v_1 = \dfrac{u_1}{\|u_1\|} = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}$ $v_2 = \dfrac{u_2}{\|u_2\|} = \begin{bmatrix} -1/\sqrt{18} \\ 4/\sqrt{18} \\ 1/\sqrt{18} \end{bmatrix}$ $v_3 = \dfrac{u_3}{\|u_3\|} = \begin{bmatrix} 2/3 \\ 1/3 \\ -2/3 \end{bmatrix}$

Check your result as in ex 5.

$v_2 = \begin{bmatrix} -1/3\sqrt{2} \\ 4/3\sqrt{2} \\ 1/3\sqrt{2} \end{bmatrix}$ * we could rationalize denominators (but not necessary)

$v_1 = \begin{bmatrix} \sqrt{2}/2 \\ 0 \\ \sqrt{2}/2 \end{bmatrix}$ $v_2 = \begin{bmatrix} -\sqrt{2}/6 \\ 2\sqrt{2}/3 \\ \sqrt{2}/6 \end{bmatrix}$ $v_3 = \begin{bmatrix} 2/3 \\ 1/3 \\ -2/3 \end{bmatrix}$

check

$v_1 \cdot v_2 = -1/\sqrt{36} + 1/\sqrt{36} = 0$

$v_1 \cdot v_3 = 2/3\sqrt{2} - 2/3\sqrt{2} = 0$

$v_2 \cdot v_3 = -2/\sqrt{54} + 4/\sqrt{54} - 2/\sqrt{54}$

$\|v_1\|^2 = v_1 \cdot v_1 = \frac{1}{2} + \frac{1}{2} = 1$

$\|v_2\|^2 = v_2 \cdot v_2 = \frac{1}{18} + \frac{16}{18} + \frac{1}{18} = 1$

$\|v_3\|^2 = v_3 \cdot v_3 = \frac{4}{9} + \frac{1}{9} + \frac{4}{9} = 1$

before we move on, can we prove (for practice) that normalizing doesn't effect orthogonality?

If we prove any scaling doesn't effect orthogonality, then we are done since normalization is a (special) type of rescaling.

So show if u and v are orthogonal then αu and βv are as well.

So $u \cdot v = 0$ and $(\alpha u)\cdot(\beta v) \underset{\text{by thm. 1}}{=} (\alpha\beta)(u\cdot v) = (\alpha\beta)\cdot(0) = 0$

so αu and βv are orthogonal

Thus, if $\alpha = \frac{1}{\|u\|}$ and $\beta = \frac{1}{\|w\|}$

We have shown that if 2 vectors are orthogonal, then their normalizations are as well.

Given an orthonormal set $\{u_1, \ldots, u_n\}$ of vectors in $\mathbb{R}^m$

let $U = [u_1 \cdots u_n]$, an $m\times n$ matrix w/ orthonormal columns

$\boxed{\text{Thm 6} \quad U^T U = I}$

pf $U^T U = \begin{bmatrix} u_1^T \\ \vdots \\ u_n^T \end{bmatrix} [u_1 \cdots u_n] = \begin{bmatrix} u_1^T u_1 & \cdots & u_1^T u_n \\ \vdots & u_i^T u_j & \vdots \\ u_n^T u_1 & \cdots & u_n^T u_n \end{bmatrix}$

row vectors; col vectors; $n\times n$ matrix

by the row-column rule for matrix mult. the (i,j) entry of U^TU is the ith row of U^T times the jth col of U, $u_i^T u_j = u_i \cdot u_j$ by def. of dot prod.

Since $\{u_1, \ldots, u_n\}$ is an orthonormal set

$u_i \cdot u_j = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i\neq j \end{cases}$ so $U^TU = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \end{bmatrix} = I_n$ the $n\times n$ identity matrix

Note: U^T is $n\times m$, U is $m\times n$ so U^TU is $n\times n$

[redacted]

$\boxed{\text{Warning: } UU^T \text{ is } m\times m\text{, but } \underline{\text{NOT}}\ I_m}$ b/c U^T does not (necessarily) have orthonormal columns (only if $m=n$ is this true for sure).

Thm 7 let U be $m\times n$ w/ orthonormal columns, let $x, y \in \mathbb{R}^n$

then a) $\|Ux\| = \|x\|$ — pf given by exercise 25 (HW)

b) $(Ux)\cdot(Uy) = x\cdot y$

c) $(Ux)\cdot(Uy) = 0$ iff $x\cdot y = 0$

Consider the linear transformation $x \mapsto Ux$

property a) says this map is length preserving, $\forall x \in \mathbb{R}^n$ on the entire domain!

property c) says the mapping also preserves orthogonality, meaning if 2 vectors are orthogonal in the domain, their images are also orthogonal.

EX 6 pick 2 vectors from ex 5

say v_2 and v_3, then $U = [v_2\ v_3] = \begin{bmatrix} -2/\sqrt{6} & 0 \\ 1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{6} & -1/\sqrt{2} \end{bmatrix}$ has orthonormal columns.

let $x = \begin{bmatrix} -\sqrt{6} \\ 2\sqrt{2} \end{bmatrix}$

Verify that $U^TU = I_2$ and $UU^T \neq I_3$ and that $\|Ux\| = \|x\|$

$U^TU = \begin{bmatrix} -2/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6} \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} -2/\sqrt{6} & 0 \\ 1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{6} & -1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

and $UU^T = \begin{bmatrix} -2/\sqrt{6} & 0 \\ 1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{6} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} -2/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6} \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 2/3 & \cancel{1/3} & -1/3 \\ -1/3 & 2/3 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{bmatrix}$ not I_3

$Ux = \begin{bmatrix} -2/\sqrt{6} & 0 \\ 1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{6} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} -\sqrt{6} \\ 2\sqrt{2} \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}$

$\|x\| = \sqrt{6 + 4(2)} = \sqrt{14}$ ✓

$\|Ux\| = \sqrt{2^2 + 1^2 + (-3)^2} = \sqrt{14}$

If $m=n$, i.e. n orthonormal vectors in $\mathbb{R}^n$ (an orthonormal basis) make up the columns of U, $n\times n$

then $U^TU = I_n \Rightarrow U^T = U^{-1}$!

and so $UU^T = I_n$ as well.

And so the rows are orthonormal as well (why?)

the book says this is surprising, but it shouldn't be hard to prove.

defn Any square invertible matrix, U w/ the property $U^T = U^{-1}$ is called an <u>orthogonal matrix</u>.

Thm 6 says that every orthogonal matrix has orthonormal columns. In fact it is not hard to see that invertibility here means that the columns actually form an orthonormal basis for $\mathbb{R}^n$.

let $U = [v_1\ v_3\ v_2]$ (from ex 5) $= \begin{bmatrix} 1/\sqrt{3} & 0 & -2/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \end{bmatrix}$ EX 7 verify that the rows are orthonormal.

order unimportant

also check that $UU^T = I$

Orthogonal matrices play a central role in a relatively new field of mathematics called Random Matrix Theory (RMT)

As well as in Chapter 7 of our book (which we don't cover but may be of potential interest to many of you CS?)