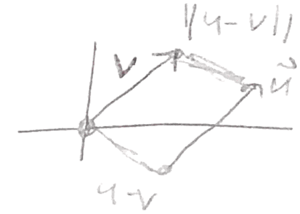


$\boxed{6.1}$ inner product, length, and orthogonality.

→ u, v vectors then they are $n \times 1$ matrices
→ u^T is $1 \times n$ matrix

(inner product)
dot product by $u \cdot v$ which is actually like $u^T v$

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

$$u \cdot v = u^T v = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + \cdots + u_n v_n$$

→ $u \cdot v = v \cdot u$

length of vector: $||v|| = \sqrt{v \cdot v} = \sqrt{v_1^2 + \cdots + v_n^2}$
magnitude

$||v||$ ↓ distance from origin to point

also like ↓

unit vector: $\dfrac{v}{||v||}$ and $||cv|| = |c| \cdot ||v||$

distance between 2 vectors: $\text{dist}(u, v) = ||u - v||$
↳ $= \sqrt{(u-v)\cdot(u-v)}$

two vectors are orthogonal if $u \cdot v = 0$

$||u+v||^2 = ||u||^2 + ||v||^2$ only if $u \perp v$

← can only use pythag on right triangles

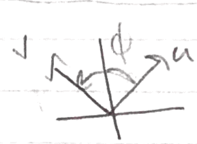
---

if z is $\perp$ to every vector in W then z orthogonal to W
$z \perp W$

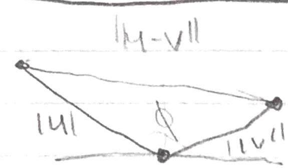
set of all $z \perp W$ is orthogonal complement of W denoted by $W^\perp$

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

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



$$\cos(\phi) = \frac{u \cdot v}{||u|| \, ||v||}$$



do ... of cos ... & ... get ϕ eventually

6.2

$u \cdot v = 0$ means they are orthogonal

orthogonal set if all combos of vectors are $\perp$

ex v_1, v_2, v_3

say $v_1 \cdot v_2 = 0$
$v_2 \cdot v_3 = 0$ Orthogonal set
$v_3 \cdot v_1 = 0$

-> if you have orthogonal set, of nonzero vectors, then the set is lin ind

Orthogonal Basis, for a subspace, W, is a basis for W that is also an orthogonal set.

ex: $\{e_1, e_2, e_3\}$ basis & orthogonal

For a basis, B, of W, finding $[X]_B$ of a vector x is not simple
-> solve $X = c_1 b_1 + c_2 b_2 + \dots + c_k b_k$
$B = \{b_1 \dots b_k\}$ and $[X]_B = \begin{bmatrix} c_1 \\ \vdots \\ c_k \end{bmatrix}$
really simple if B is an orthogonal basis:

orthog basis
Let $B = \{b_1 \dots b_p\}$ be a basis for W, for each x in W ->

$\vec{X} = c_1 b_1 + \dots c_p b_p$, $\vec{x} \cdot \vec{b_1} = c_1 b_1 \cdot b_1 + \cancel{c_2 b_1 \cdot b_2}^{0} + \cancel{c_3 b_1 \cdot b_3}^{0} + \dots + c_p b_p \cdot b_1$

$c_1 = \frac{\vec{x} \cdot b_1}{b_1 \cdot b_1}$ to find c_2 or c_j do $c_j = \frac{\vec{x} \cdot b_j}{b_j \cdot b_j}$

$[X]_B = \begin{bmatrix} c_1 \\ \vdots \\ c_j \end{bmatrix}$

$$P = \frac{LL^T}{L^TL}$$

has projection matrix

projection

$\hat{y} = \text{proj}_L y$ = projection of y onto $L = \frac{y \cdot u}{u \cdot u} u$

ex let $y = \begin{bmatrix} 7 \\ 6 \end{bmatrix}$, $u = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$ find the orthogonal proj of y onto u

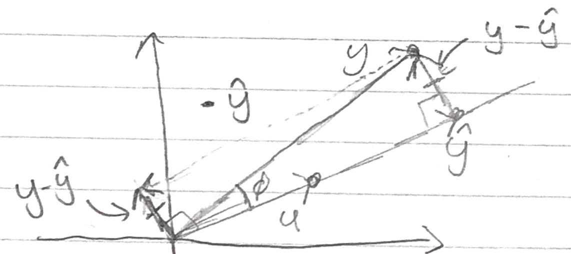
$y \cdot u = \begin{bmatrix} 7 \\ 6 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 40$, $u \cdot u = \begin{bmatrix} 4 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 20$

$\text{proj}_u y = \hat{y} = \frac{y \cdot u}{u \cdot u} u = \frac{40}{20} u = 2 \begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$

component of $y \perp u$, $y - \hat{y} = \begin{bmatrix} 7 \\ 6 \end{bmatrix} - \begin{bmatrix} 8 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$

$\begin{bmatrix} 7 \\ 6 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix} + \begin{bmatrix} -1 \\ 2 \end{bmatrix}$

$y = \hat{y} + (y - \hat{y})$



distance from y to $\hat{y}$

$\|y - \hat{y}\| = \sqrt{(-1)^2 + 2^2} = \sqrt{5}$

→ U has $\perp$ columns iff $U^T U = I$ → $u_1^T u_1 = 1$ if columns of u have unit length

$U^T U = \begin{bmatrix} u_1^T \\ u_2^T \\ u_3^T \end{bmatrix} \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix} = \begin{bmatrix} u_1^T u_1 & u_1^T u_2 & u_1^T u_3 \\ u_2^T u_1 & u_2^T u_2 & u_2^T u_3 \\ u_3^T u_1 & u_3^T u_2 & u_3^T u_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ are orthogonal iff, $u_1^T u_2 = 0$

→ U w/ orthogonal columns, $\|Ux\| = \|x\|$, $(Ux) \cdot (Uy) = x \cdot y$, $= 0$ iff $x \cdot y = 0$

6.3

Let $\{u_1, u_2, u_3, u_4, u_5\}$ be an orthogonal basis of $\mathbb{R}^5$

and $\vec{y} = c_1\vec{u}_1 + c_2\vec{u}_2 + c_3\vec{u}_3 + c_4\vec{u}_4 + c_5\vec{u}_5$

$W = \text{Span}\{u_1, u_2\}$ subspace of $\mathbb{R}^5$

Write $\vec{y}$ as a sum of vector z_1 (inside W) and z_2 (inside $W^\perp$)

$$\vec{y} = \underbrace{(c_1u_1 + c_2u_2)}_{z_1} + \underbrace{(c_3u_3 + c_4u_4 + c_5u_5)}_{z_2}$$

show: $z_2 \perp W$ or z_2 is in $W^\perp$

$z_2 \cdot u_1 = u_1 \cdot (c_3u_3 + c_4u_4 + c_5u_5) = c_3u_3 \cdot u_1 + c_4 u_4 \cdot u_1 + c_5 u_5 \cdot u_1$

$\rightarrow z_2 \cdot u_1 = 0$ orthogonal basis

similarly $z_2 \perp u_1$

$z_2 \cdot u_2 = 0$

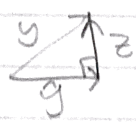
$\rightarrow$ so z_2 is orthogonal to all of $W = \{u_1, u_2\}$

Generally:

$$y = \hat{y} + z$$

"projection"

$\uparrow$ $\hat{y}$ in W $\qquad$ $\uparrow$ z in $W^\perp$

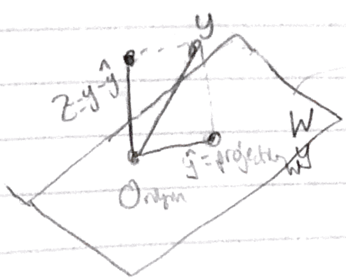


only works for orthogonal basis

---

if $\{u_1, \dots, u_p\}$ is any orthogonal basis of W then

$$\hat{y} = \frac{y \cdot u_1}{u_1 \cdot u_1} u_1 + \dots + \frac{y \cdot u_p}{u_p \cdot u_p} u_p$$ (c_1, c_p) and $z = y - \hat{y}$



proj of y onto W doesn't depend on choice of orthogonal basis.

ex $u_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}$, $u_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$, $y = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

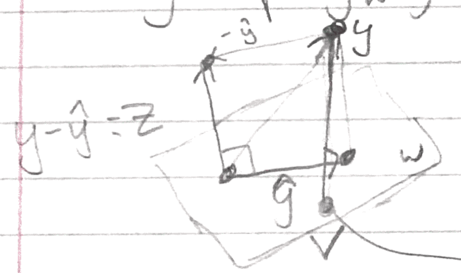
find $\text{proj}_W y$, $W = \text{Span}\{u_1, u_2\}$ ← show true

$u_1 \cdot u_2 = 0$

$$\text{proj}_W y = \hat{y} = \frac{y \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{y \cdot u_2}{u_2 \cdot u_2} u_2$$

$$= \frac{9}{30} \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + \frac{3}{6} \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix}$$

$\hat{y} = \text{proj}_W y$ is the closest point in W to y



point coming off will always be longer

point coming from $z = y - \hat{y}$

$$\|y - \hat{y}\| \leq \|y - v\| \quad (v \text{ is in } W)$$

$\boxed{\text{Distance between a vector } y \text{ in } \mathbb{R}^n \text{ and a subspace of } \mathbb{R}^n\ (W) \text{ is distance } y \text{ to nearest point}}$ → $\text{dist}(y, W) = \text{dist}(y, \hat{y}) = \|y - \hat{y}\|$

y in $\mathbb{R}^n$, $\hat{y}$ on W

★ remember $\hat{y}$ for W is $\hat{y} = \frac{y u_1}{u_1 u_1} u_1 + \frac{y u_2}{u_2 u_2} u_2 + \dots \frac{y u_p}{u_p u_p} u_p$

if ~~DAX~~ $\{u_1, \dots, u_p\}$ is orthonormal (unit vector) → $U = [u_1 \dots u_p]$

$\text{proj}_W y = \frac{y u_1}{u_1 u_1} u_1 + \dots = \frac{y u_1}{\|u_1\|^2} u_1 + \dots = \frac{(y \cdot u_1)}{1} u_1 + \dots$ → then $\text{proj}_W y = (U U^T) y$

orthonormal so length = 1

→ transformation onto W — not $U^T U = I$

Gram schmidt process

How to find an orthogonal basis of a nontrivial subspace of $\mathbb{R}^n$?

let $W = \text{Span}\{x_1, x_2\}$ $\quad x_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}$ $\quad x_2 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$

→ construct an orthogonal basis $\{v_1, v_2\}$ of W

check $x_1 \cdot x_2 \neq 0 = 15$

first set $v_1 = x_1$ now find v_2 thats in W and $v_2 \perp (v_1 = x_1)$

let $V = \text{Span}\{v_1\} = \text{Span}\{x_1\}$

we know x_2 is not on V,

$$\underbrace{x_2 - \hat{x}_2}_{\perp V} = x_2 - \text{proj}_V\, x_2 \ \overset{\text{on } V}{=} v_2 \qquad \text{thus } v_2 \perp (v_1 = x_1)$$

call this

projection formula of x_2 onto v
and v is spaned by x_1

$$v_2 = x_2 - \hat{x}_2 = x_2 - \overbrace{\frac{x_2 \cdot x_1}{x_1 \cdot x_1}\, x_1}$$

$$v_2 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \frac{15}{45}\begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} = v_2$$

$\{v_1, v_2\}$, an orthogonal subspa of W

ex 2

$$X_1 = \begin{bmatrix}1\\1\\1\\1\end{bmatrix} \quad X_2 = \begin{bmatrix}0\\1\\1\\1\end{bmatrix} \quad X_3 = \begin{bmatrix}0\\0\\1\\1\end{bmatrix}$$

they are lin ind and $\{X_1, X_2, X_3\}$ is a basis for W & spans,

find orthogonal basis for W

V_1, V_2, V_3 form an orthogonal basis for W

Let $V_1 = X_1$ $\quad W_1 = \text{Span}\{V_1\} = \text{Span}\{X_1\}$ ← $V_1 = X_1$

$$V_2 = X_2 - \text{proj}_{W_2} X_2 = X_2 - \frac{X_2 \cdot V_1}{V_1 \cdot V_1} \cdot V_1$$

lets clean up 1/4 mess by replacing $V_2 = 4V_2$ doesn't change orthogonalness so its ok

$$V_2 = \begin{bmatrix}0\\1\\1\\1\end{bmatrix} - \frac{3}{4}\begin{bmatrix}1\\1\\1\\1\end{bmatrix} = \begin{bmatrix}-3/4\\1/4\\1/4\\1/4\end{bmatrix}$$

$$V_2 = \begin{bmatrix}-3\\1\\1\\1\end{bmatrix}$$ and $V_1 \perp V_2$

Now find $V_3 \perp (V_1, V_2)$. Same thing. $W_2 = \text{Span}\{V_1, V_2\}$

$$V_3 = X_3 - \text{proj}_{W_2} X_3$$ ← $\perp W_2$

$$V_3 = X_3 - \frac{X_3 \cdot V_1}{V_1 \cdot V_1} \cdot V_1 - \frac{X_3 \cdot V_2}{V_2 \cdot V_2} \cdot V_2$$

$$V_3 = \begin{bmatrix}0\\0\\1\\1\end{bmatrix} - \frac{2}{4}\begin{bmatrix}1\\1\\1\\1\end{bmatrix} - \frac{2}{12}\begin{bmatrix}-3\\1\\1\\1\end{bmatrix} = \begin{bmatrix}0\\-2/3\\1/3\\1/3\end{bmatrix} = V_3$$

$W_3 = \text{Span}\{V_1, V_2, V_3\}$ forms a basis for W

Overall

$V_1 = X_1$

$V_2 = X_2 - \frac{X_2 \cdot V_1}{V_1 \cdot V_1} \cdot V_1$

$V_3 = X_3 - \frac{X_3 \cdot V_1}{V_1 \cdot V_1} \cdot V_1 - \frac{X_3 \cdot V_2}{V_2 \cdot V_2} \cdot V_2$

$V_p = X_p - \frac{X_p \cdot V_1}{V_1 \cdot V_1} \cdot V_1 - \dots - \frac{X_p \cdot V_{p-1}}{V_{p-1} \cdot V_{p-1}} \cdot V_{p-1}$

---

can now make a othernormal basis

$\{V_1, V_2, \dots V_k\} \to \left\{ \frac{V_1}{\|V_1\|}, \dots \frac{V_k}{\|V_k\|} \right\}$

$\boxed{\text{QR Factorization}}$

$\underset{m\times n}{A} = \underset{m\times n}{Q} \cdot R$, A's columns are lin ind.

cols of Q = orthonormal basis of Col A

R upper triangular w/ positive diagonal entries

ex

$A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ Col A = Span $\{X_1, X_2, X_3\}$

← from earlier $A = [X_1\ X_2\ X_3]$ and thus

V_1, V_2, V_3 form orthogonal basis for W = Col A

Q = orthonorm of V_1, V_2, V_3 = $\left[\frac{V_1}{\|V_1\|}\ \frac{V_2}{\|V_2\|}\ \frac{V_3}{\|V_3\|} \right] = \begin{bmatrix} 1/2 & -3/\sqrt{12} & 0 \\ 1/2 & 1/\sqrt{12} & -2/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \end{bmatrix}$

Since $A = QR \Rightarrow Q^T A = \underbrace{Q^T Q}_{\text{recall w/ orthonormal stuff} = I} R = R$

thus $R = \begin{bmatrix} 2 & 3/2 & 1 \\ 0 & 3/\sqrt{12} & 2/\sqrt{12} \\ 0 & 0 & 2/\sqrt{6} \end{bmatrix}$

$Ax=b$ sometimes has no solution

When this happens we would like to find an x that makes Ax as close as possible to b

A is $m\times n$ and b is in $\mathbb{R}^m$

we want an $\hat{x}$ such that

$$\|b-A\hat{x}\| \le \|b-Ax\|$$

Let $A=[a_1\ a_2 \cdots a_n]$ $\quad x=\begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix}$

$Ax=r_1a_1+r_2a_2+\cdots+r_na_n$

Ax is in Col A for any x

$A\hat{x}$ should be the closest point in Col A to b

$A\hat{x}=\hat{b}=\text{proj}_{\text{Col}A}\, b$ since $\hat{b}$ is in Col A, $A\hat{x}=\hat{b}$ is consistent

Solution of ~~of~~ this problem is $A\hat{x}=\hat{b}$, time consuming if not in ...

second approach

Suppose $\hat{x}$ satisfies $A\hat{x}=\hat{b}$, $(b-\hat{b}) \perp \text{Col}A$

$b-A\hat{x}=(b-\hat{b}) \leftarrow$ this is $\perp$ Col A and $b-A\hat{x}$ is $=$ to it

... Col A

$a_i\cdot(b-A\hat{x})=0$

$a_i^T\cdot(b-A\hat{x})=0 \to A^T(b-A\hat{x})=0$

$\to A^Tb-A^TA\hat{x}=0 \to A^Tb=A^TAx$ — normal eq for $Ax=b$

set of least squares sol's of $Ax = b$
equals the (non-empty) sol set of normal eq

$$A^T A \hat{x} = A^T b$$

(ex) find a least squares sol for the inconsistent sys $Ax = b$ where

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

$$A^T A = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}$$

$$A^T b = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$ then $A^T A x = A^T b$

$$\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix} \longrightarrow \left[\begin{array}{cc|c} 17 & 1 & 19 \\ 1 & 5 & 11 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 2 \end{array}\right]$$

$$\hat{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Theorem

a) $Ax = b$ has unique least squares sol for
b) Cols A are l.in ind
c) $A^T A$ is invertible $\rightarrow$ $\hat{x} = (A^T A)^{-1} A^T b$

$\rightarrow \|b - A\hat{x}\|$ the least squares error in this apprx.

(ex) First ex

$$\left\| \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} - \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\| = \left\| \begin{bmatrix} -2 \\ -4 \\ 8 \end{bmatrix} \right\| = \sqrt{84}$$