Chapter 6

Orthogonality

Section 6.1

Dot Products and Orthogonality

Orientation

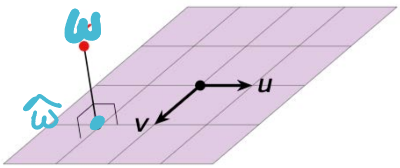
Recall: This course is about learning to:

- Solve the matrix equation $Ax = b$ ← idea row reduce $[A|b]$ & get parametric vector form.
- Solve the matrix equation $Ax = \lambda x$ ← eigenvalues & eigenvectors solving $(A - \lambda I)x = 0$.
- Almost solve the equation $Ax = b$

We are now aiming at the last topic. ← inconsistent systems ??

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

Idea: In the real world, data is imperfect. Suppose you measure a data point w which you know for theoretical reasons must lie on a plane spanned by two vectors u and v.



Goal: given x in R^n, find the closest vector in Span{u,v} to the vector x

The Dot Product

Definition

The **dot product** of two vectors x, y in $\mathbf{R}^n$ is

$$x \cdot y = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \stackrel{\text{def}}{=} x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

Thinking of x, y as column vectors, this is the same as $x^T y$.

$$x \cdot y = x^T y = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = 32$$

Example $x \cdot y$

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = (1)(4) + (2)(5) + (3)(6) = 4 + 10 + 18 = 32$$

Example

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = 1^2 + 2^2 + 3^2 = 14$$

Properties of the Dot Product

Many usual arithmetic rules hold, as long as you remember you can only dot two vectors together, and that the result is a scalar.

- $x \cdot y = y \cdot x$
- $(x + y) \cdot z = x \cdot z + y \cdot z$
- $(cx) \cdot y = c(x \cdot y)$

Dotting a vector with itself is special:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1^2 + x_2^2 + \cdots + x_n^2.$$

Hence:

- $x \cdot x \geq 0$
- $x \cdot x = 0$ if and only if $x = 0$.

Important: $x \cdot y = 0$ does *not* imply $x = 0$ or $y = 0$. For example, $\binom{1}{0} \cdot \binom{0}{1} = 0$.

The Dot Product and Length

Definition

The **length** or **norm** of a vector x in $\mathbf{R}^n$ is

$$\|x\| = \sqrt{x \cdot x} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

Fact

If x is a vector and c is a scalar, then $\|cx\| = |c| \cdot \|x\|$.

Example: find the length of 2*[3;4]

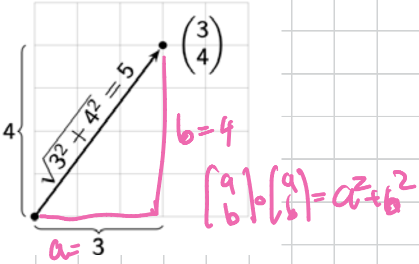
$\begin{bmatrix}3\\4\end{bmatrix}$ has length 5.

$2 * \begin{bmatrix}3\\4\end{bmatrix}$ has length 10

$$\left\|2\begin{bmatrix}3\\4\end{bmatrix}\right\| = \left\|\begin{bmatrix}6\\8\end{bmatrix}\right\| = \sqrt{6^2+8^2} = \sqrt{100} = \boxed{10}$$

$\left\|\begin{bmatrix}3\\4\end{bmatrix}\right\|$ length of $\begin{bmatrix}3\\4\end{bmatrix}$

$$\begin{bmatrix}3\\4\end{bmatrix}\cdot\begin{bmatrix}3\\4\end{bmatrix} = 3^2+4^2 = 25$$



$$\left\|\begin{pmatrix}3\\4\end{pmatrix}\right\| = \sqrt{25} = \boxed{5}$$

The Dot Product and Distance

Definition

The **distance** between two points x, y in $\mathbf{R}^n$ is

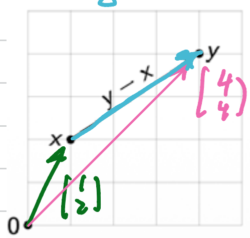
$$\mathrm{dist}(x,y) = \|y - x\|.$$

Example

Let $x = (1,2)$ and $y = (4,4)$. Then

$$\mathrm{dist}(x,y) = \left\|\begin{bmatrix}4\\4\end{bmatrix} - \begin{bmatrix}1\\2\end{bmatrix}\right\| = \boxed{\sqrt{13}}$$

$$\vec{y} - \vec{x} = \begin{bmatrix}4\\4\end{bmatrix} - \begin{bmatrix}1\\2\end{bmatrix} = \begin{bmatrix}3\\2\end{bmatrix}$$



$$\left\|\begin{bmatrix}3\\2\end{bmatrix}\right\| = \sqrt{9+4} = \sqrt{13}$$

Unit Vectors

Definition
A **unit vector** is a vector v with length $\|v\| = 1$.

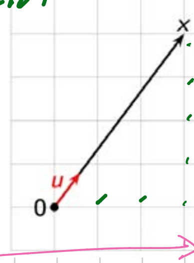
Example
The unit coordinate vectors are unit vectors:

$$\|e_1\| = \left\| \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\| = \sqrt{1^2 + 0^2 + 0^2} = 1$$

is $\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ both unit vectors

$\begin{pmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{pmatrix}$ also unit. $\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$, $\theta \in \mathbb{R}$.

Definition
Let x be a nonzero vector in $\mathbf{R}^n$. The **unit vector in the direction of** x is the vector $\frac{x}{\|x\|}$.



$\begin{bmatrix} 3 \\ 4 \end{bmatrix} = \vec{x}$ w/ length $\underline{5}$.

??

$\vec{u} = c\begin{bmatrix} 3 \\ 4 \end{bmatrix}$ $c = \frac{1}{5}$

$\boxed{\vec{u} = \frac{1}{5}\begin{bmatrix} 3 \\ 4 \end{bmatrix}}$ is a unit vector.

Example
What is the unit vector in the direction of $x = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$?

Check. $\|\vec{u}\| = \left\| \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix} \right\|$

$= \sqrt{(3/5)^2 + (4/5)^2} = \sqrt{\frac{9+16}{25}} = 1$ ✓

Orthogonality

Definition
Two vectors x, y are **orthogonal** or **perpendicular** if $x \cdot y = 0$.
Notation: $x \perp y$ means $x \cdot y = 0$.

$$\boxed{\text{Fact: } x \perp y \iff \|x - y\|^2 = \|x\|^2 + \|y\|^2}$$

You Try It!

Example: Which of the following pairs of vectors are orthogonal? Find all pairs.

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad v_4 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad v_5 = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix}.$$

$V_1 \cdot V_2 = 0$ ✓
$V_1 \cdot V_3 = 0$ ✓
$V_1 \cdot V_4 = 1$ X
$V_1 \cdot V_5 = -4$ X

$V_2 \cdot V_3 = 0$ ✓
$V_2 \cdot V_4 = 0$ ✓
$V_2 \cdot V_5 = 1$ X

$V_3 \cdot V_4 = 0$ ✓
$V_3 \cdot V_5 = 0$ ✓

$V_4 \cdot V_5 = -1$ X.

Orthogonality
Example

Problem: Find *all* vectors orthogonal to $v = \begin{pmatrix}1\\1\\-1\end{pmatrix}$.

Key insight:

$$0 = x \cdot v = \begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} \cdot \begin{pmatrix}1\\1\\-1\end{pmatrix} = x_1 + x_2 - x_3.$$

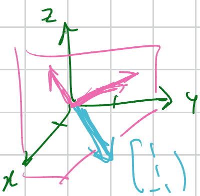
e.g. $w_1 = \begin{bmatrix}2\\0\\2\end{bmatrix}$ $\quad w_2 = \begin{bmatrix}0\\0\\0\end{bmatrix}$ $\quad w_3 = \begin{bmatrix}0\\1\\1\end{bmatrix}$

Solve for x_1, x_2, x_3

$$\boxed{\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} \cdot \begin{bmatrix}1\\1\\-1\end{bmatrix} = 0} \implies x_1 + x_2 - x_3 = 0$$

Solve!!

$$\begin{array}{cccc} & x_2 & x_3 & \\ \end{array}$$
$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & 0 \end{array}\right] \quad \text{RREF ✓}$$



$$X = \begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = \begin{bmatrix}-x_2+x_3\\x_2\\x_3\end{bmatrix} = x_2\begin{bmatrix}-1\\1\\0\end{bmatrix} + x_3\begin{pmatrix}1\\0\\1\end{pmatrix}.$$

X is orthogonal to $\begin{pmatrix}1\\1\\-1\end{pmatrix}$ iff x in $\operatorname{Span}\left\{\begin{bmatrix}-1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\0\\1\end{bmatrix}\right\}$

Orthogonality
Example

Problem: Find *all* vectors orthogonal to both $v = \begin{pmatrix}1\\1\\-1\end{pmatrix}$ and $w = \begin{pmatrix}1\\1\\1\end{pmatrix}$.

Key insight:

$$0 = x \cdot v = \begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} \cdot \begin{pmatrix}1\\1\\-1\end{pmatrix} = x_1 + x_2 - x_3$$
$$0 = x \cdot w = \begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} \cdot \begin{pmatrix}1\\1\\1\end{pmatrix} = x_1 + x_2 + x_3.$$

Now solve for $x = \begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$

$$\left\{\begin{aligned} &\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} \cdot \begin{bmatrix}1\\1\\-1\end{bmatrix} = 0 \\ &\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} \cdot \begin{bmatrix}1\\1\\1\end{bmatrix} = 0 \end{aligned}\right. \implies \left\{\begin{aligned} x_1 + x_2 - x_3 &= 0 \\ x_1 + x_2 + x_3 &= 0 \end{aligned}\right.$$

two homogeneous eqns in 3 variables x_1, x_2, x_3

both must be true simultaneously

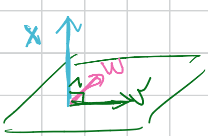
Coeff. matrix.

$$\Rightarrow A = \begin{bmatrix}1 & 1 & -1\\1 & 1 & 1\end{bmatrix} \sim \begin{bmatrix}1 & 1 & -1\\0 & 0 & -2\end{bmatrix} \sim \begin{array}{c} \begin{array}{ccc} & x_2 & \end{array} \\ \begin{bmatrix}1 & 1 & 0\\0 & 0 & 1\end{bmatrix} \end{array}$$

$$X = x_2\begin{bmatrix}-1\\1\\0\end{bmatrix}.$$

So $X = \begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$ is orthogonal to Both $v = \begin{pmatrix}1\\1\\-1\end{pmatrix}$ and $w = \begin{pmatrix}1\\1\\1\end{pmatrix}$

iff X is in $\operatorname{span}\left\{\begin{pmatrix}-1\\1\\0\end{pmatrix}\right\}$.



Orthogonality
General procedure

Problem: Find all vectors orthogonal to some number of vectors $v_1, v_2, \ldots, v_m$ in $\mathbf{R}^n$.

This is the same as finding all vectors x such that

$$0 = v_1^T x = v_2^T x = \cdots = v_m^T x.$$

Putting the *row* vectors $v_1^T, v_2^T, \ldots, v_m^T$ into a matrix, this is the same as finding all x such that

$$\begin{pmatrix} — v_1^T — \\ — v_2^T — \\ \vdots \\ — v_m^T — \end{pmatrix} x = \begin{pmatrix} v_1 \cdot x \\ v_2 \cdot x \\ \vdots \\ v_m \cdot x \end{pmatrix} = 0.$$

Important

The set of all vectors orthogonal to some vectors $v_1, v_2, \ldots, v_m$ in $\mathbf{R}^n$ is the *null space* of the $m \times n$ matrix you get by "turning them sideways and smooshing them together:"

$$\begin{pmatrix} — v_1^T — \\ — v_2^T — \\ \vdots \\ — v_m^T — \end{pmatrix}.$$

In particular, this set is a subspace!

If you want to find all vectors x such that

$x \cdot v_1 = 0$ AND $x \cdot v_2 = 0$ just solve $\begin{bmatrix} v_1^T \\ v_2^T \end{bmatrix} \vec{x} = \vec{0}$

just put v_1^T, v_2^T as ROWS of a matrix A

& compute the null space of A.

Summary

- The **dot product** of vectors x, y in $\mathbf{R}^n$ is the number $x^T y$.
- The **length** or $\boxed{\text{norm}}$ of a vector x in $\mathbf{R}^n$ is $\|x\| = \sqrt{x \cdot x}$.
- The **distance** between two vectors x, y in $\mathbf{R}^n$ is $\text{dist}(x, y) = \|y - x\|$.
- A **unit vector** is a vector v with length $\|v\| = 1$.
- The **unit vector in the direction of** x is $x/\|x\|$.
- Two vectors x, y are **orthogonal** if $x \cdot y = 0$.
- The set of all vectors orthogonal to some vectors $v_1, v_2, \ldots, v_m$ in $\mathbf{R}^n$ is the null space of the matrix

$$\begin{pmatrix} — v_1^T — \\ — v_2^T — \\ \vdots \\ — v_m^T — \end{pmatrix}.$$

Section 6.2

Orthogonal Complements

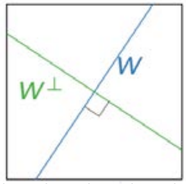
Orthogonal Complements

Definition
Let W be a subspace of $\mathbf{R}^n$. Its **orthogonal complement** is

$$W^\perp = \{v \text{ in } \mathbf{R}^n \mid v \cdot w = 0 \text{ for all } w \text{ in } W\} \qquad \text{read "}W\text{ perp".}$$

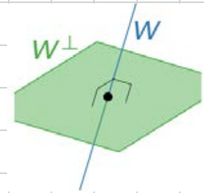
$W^\perp$ is orthogonal complement
A^T is transpose

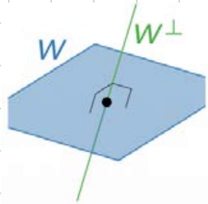
$W^\perp$ "W perp"
or The orthogonal complement of W.



W a line
$W^\perp$ the orthogonal line to W.
(both pass through $\begin{pmatrix}0\\0\end{pmatrix}$)

https://textbooks.math.gatech.edu/ila/demos/spans.html?v1=.3,0,1&captions=orthog&range=3





You Try It!

Poll

Let W be a 2-plane in $\mathbf{R}^4$. How would you describe $W^\perp$?

A. The zero space $\{0\}$.
B. A line in $\mathbf{R}^4$.
C. A plane in $\mathbf{R}^4$.
D. A 3-dimensional space in $\mathbf{R}^4$.
E. All of $\mathbf{R}^4$.

Hint: try an example! What is the set of vectors which are orthogonal to both e1 and e2 in R^4?

For example, if W is the xy-plane, then $W^\perp$ is the zw-plane:

$$\begin{pmatrix}x\\y\\0\\0\end{pmatrix} \cdot \begin{pmatrix}0\\0\\z\\w\end{pmatrix} = 0.$$

e.g. If $W = \text{Span}\left\{\begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\end{bmatrix}\right\}$

then $W^\perp = \text{Nul}\left(\begin{bmatrix}1&0&0&0\\0&1&0&0\end{bmatrix}\right)$

& $\dim W^\perp = 2$ (# free var of $\begin{array}{cccc} & & x_3 & x_4 \end{array}$ $\begin{bmatrix}1&0&0&0\\0&1&0&0\end{bmatrix}$).

Orthogonal Complements
Basic properties

Let W be a subspace of $\mathbf{R}^n$.

Facts:

1. $W^\perp$ is also a subspace of $\mathbf{R}^n$
2. $(W^\perp)^\perp = W$
3. $\dim W + \dim W^\perp = n$
4. If $A = \begin{pmatrix} v_1 & v_2 & \cdots & v_m \end{pmatrix}$ and $W = \mathrm{Col}\ A$, then $W^\perp = \mathrm{Nul}(A^T)$

Problem: if $W = \mathrm{Span}\left\{ \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$, compute $W^\perp$.

Hint: use (4.)

(4.) says:

$$\mathrm{Span}\{v_1, v_2, \ldots, v_m\}^\perp = \mathrm{Nul} \begin{pmatrix} — v_1^T — \\ — v_2^T — \\ \vdots \\ — v_m^T — \end{pmatrix}$$

Soln. To find basis for $W^\perp$ we will row reduce

$$A^T = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ -1 & 1 \end{bmatrix}^T = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix} \sim \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

x_2

$$X = x_2 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

So $\boxed{W^\perp = \mathrm{Span}\left\{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}\right\}}$ if $W = \mathrm{Col}\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ -1 & 1 \end{pmatrix}$

Orthogonal Complements
Row space, column space, null space

Definition
The **row space** of an $m \times n$ matrix A is the span of the *rows* of A. It is denoted Row A. Equivalently, it is the column space of A^T:

$$\text{Row } A = \text{Col } A^T.$$

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

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

Compute (1) $\text{Row } A = \text{span}\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\} = \mathbb{R}^2$

(2) $(\text{Row } A)^\perp = \text{Nul}\left(\begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix}^T \right) = \text{Nul}\left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ -1 & 1 \end{bmatrix} \right) = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$

Dimension of the row space

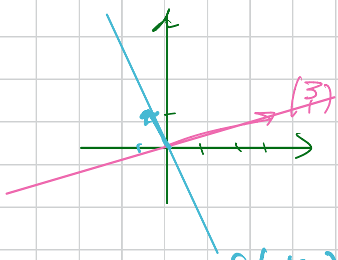
Even though Row(A) lives in $\mathbf{R}^n$ and Col(A) lives in $\mathbf{R}^m$ if A is an $m \times n$ matrix, both subspaces have the same dimension.

Theorem
If A is an $m \times n$ matrix, then dim(Row A) = dim(Col A).

Idea: the dimension of Row A is equal to the number of pivots of A, which is the dimension of the column space of A.

$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ -1 & 1 \end{bmatrix}$ e.g. Row A is a subspace of $\mathbb{R}^2$, Col A is a subspace of $\mathbb{R}^3$ } but both have dim equal to 2.

EX. In $\mathbb{R}^2$ $W = \text{span}\left\{ \begin{bmatrix} 3 \\ 1 \end{bmatrix} \right\}$ Find $W^\perp$?



Span $\left\{ \begin{bmatrix} -1/3 \\ 1 \end{bmatrix} \right\}$

Idea: $W^\perp = \text{Nul}\left(\begin{bmatrix} 3 \\ 1 \end{bmatrix}^T \right) = \text{Nul}\left(\begin{bmatrix} 3 & 1 \end{bmatrix} \right)$

$$\begin{bmatrix} 3 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1/3 \end{bmatrix} \quad \text{RREF ✓} \qquad x = x_2 \begin{bmatrix} -1/3 \\ 1 \end{bmatrix}$$

So $\boxed{W^\perp = \text{span}\left\{ \begin{bmatrix} -1/3 \\ 1 \end{bmatrix} \right\}}$

EX. In $\mathbb{R}^3$ $W = \text{span}\left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$ Find $W^\perp$?

$\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = 0$ ✓

$\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = 0$ ✓

$W^\perp = \text{Nul}\left(\begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 1 & 1 \end{bmatrix}^T \right) = \text{Nul}\left(\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix} \right)$

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix}$$

x_3 is free.

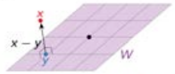
$x = x_3 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$ $\boxed{W^\perp = \text{span}\left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}}$

Section 6.3

Orthogonal Projections (will finish in next set of slides)

Best Approximation

Suppose you measure a data point x which you know for theoretical reasons must lie on a subspace W.



Due to measurement error, though, the measured x is not actually in W. Best approximation: y is the *closest* point to x on W.

How do you know that y is the closest point? The vector from y to x is orthogonal to W: it is in the *orthogonal complement* $W^\perp$.

Orthogonal Decomposition

Theorem

Every vector x in $\mathbf{R}^n$ can be written as

in W   in W⊥

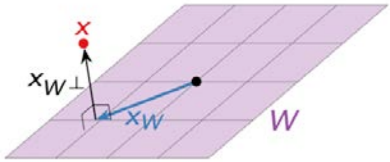
$$x = x_W + x_{W^\perp}$$

for unique vectors x_W in W and $x_{W^\perp}$ in $W^\perp$.

The equation $x = x_W + x_{W^\perp}$ is called the **orthogonal decomposition** of x (with respect to W).

The vector x_W is the **orthogonal projection** of x onto W.

https://textbooks.math.gatech.edu/ila/demos/projection.html?u1=1,0,0&u2=0,1.1,-.2&vec=-1.1,2,1.5&range=3&mode=decomp&closed



https://textbooks.math.gatech.edu/ila/demos/projection.html?u1=1,0,0&u2=0,1,0&vec=-1.1,2,1.5&range=3&mode=decomp&closed

Orthogonal Decomposition

Example

Let W be the xy-plane in $\mathbf{R}^3$. Then $W^\perp$ is the z-axis.

Find the orthogonal decomposition of x=[2;1;3]

in W   in W⊥

$$x = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}.$$

in xy-plane   in z-axis

Orthogonal Decomposition
Computation?

Problem: Given x and W, how do you compute the decomposition $x = x_W + x_{W^\perp}$?

Observation: It is enough to compute x_W, because $x_{W^\perp} = x - x_W$.

Theorem (The A^TA Trick)

Let W be a subspace of $\mathbf{R}^n$, let $v_1, v_2, \ldots, v_m$ be a spanning set for W (e.g., a basis), and let

$$A = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_m \\ | & | & & | \end{pmatrix}.$$

Then for any x in $\mathbf{R}^n$, the matrix equation

$$\boxed{A^TAv = A^Tx} \quad \text{(in the unknown vector } v\text{)}$$

is consistent, and $x_W = Av$ for any solution v.

Recipe for Computing $x = x_W + x_{W^\perp}$

- Write W as a column space of a matrix A.
- Find a solution v of $A^TAv = A^Tx$ (by row reducing).
- Then $x_W = Av$ and $x_{W^\perp} = x - x_W$.

$$\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = \boxed{\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}} + \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$$

in W (xy-plane) in $W^\perp$ (z-axis)

use $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$ w/ A^TA trick

The A^TA Trick
Example

Problem: Compute the orthogonal projection of a vector $x = (x_1, x_2, x_3)$ in $\mathbf{R}^3$ onto the xy-plane.

Write $\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ as $x_W + x_{W^\perp}$ $\quad W = \text{Span}\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right\} = \text{Col}\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$

$$\underset{2\times 3}{}\quad \underset{3\times 2}{}\quad \underset{2\times 2}{}$$

$$A^TA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\underset{2\times 3}{}\quad \underset{3\times 1}{}\quad \underset{2\times 1}{}$$

$$A^Tx = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ \cancel{z} \end{bmatrix}$$

Soln $\left[A^TA \mid A^Tx\right] = \left[\begin{array}{cc|c} 1 & 0 & x \\ 0 & 1 & y \end{array}\right]$

$$v = \begin{bmatrix} x \\ y \end{bmatrix}.$$

$$x_W = Av = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \boxed{\begin{bmatrix} x \\ y \\ 0 \end{bmatrix} = x_W}$$

The A^TA Trick
Another Example

Problem: Let

$$\boxed{x = \begin{pmatrix}1\\2\\3\end{pmatrix}} \qquad W = \left\{ \begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} \text{ in } \mathbf{R}^3 \mid x_1 - x_2 + x_3 = 0 \right\} = \text{Col } A$$

Recipe for Computing $x = x_W + x_{W^\perp}$

- Write W as a column space of a matrix A.
- Find a solution v of $A^TAv = A^Tx$ (by row reducing).
- Then $\boxed{x_W = Av}$ and $x_{W^\perp} = x - x_W$.

(1) compute the orthogonal decomposition of x for the given subspace W

$$\boxed{x = \begin{bmatrix}1\\2\\3\end{bmatrix} = \underbrace{x_W}_{\text{in } W} + \underbrace{x_{W^\perp}}_{\text{in } W^\perp}}$$

(2) find the distance from x to the subspace W

$W = \text{Span}\{v_1, v_2\}$

where v_1, v_2 are sols to

$$\begin{bmatrix}1 & -1 & 1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = 0$$

$$x = x_2\begin{pmatrix}1\\1\\0\end{pmatrix} + x_3\begin{pmatrix}-1\\0\\1\end{pmatrix}$$

$\uparrow v_1 \qquad \uparrow v_2$

$$A = \begin{bmatrix}v_1 & v_2\end{bmatrix} = \begin{bmatrix}1 & -1\\1 & 0\\0 & 1\end{bmatrix}$$

$\uparrow\ \uparrow$ need basis $\{v_1, v_2\}$ for W.

$$A^TA = \begin{bmatrix}1 & 1 & 0\\-1 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & -1\\1 & 0\\0 & 1\end{bmatrix} = \begin{bmatrix}2 & -1\\-1 & 2\end{bmatrix}$$

$$A^Tx = \begin{bmatrix}1 & 1 & 0\\-1 & 0 & 1\end{bmatrix}\boxed{\begin{bmatrix}1\\2\\3\end{bmatrix}} = \begin{bmatrix}3\\2\end{bmatrix}$$

next solve $A^TAv = A^Tx$

$$-2 + \frac{14}{3} = \frac{-6+14}{3} = 8/3$$

$$\left[\begin{array}{cc|c}2 & -1 & 3\\-1 & 2 & 2\end{array}\right] \sim \updownarrow \left[\begin{array}{cc|c}1 & -2 & -2\\2 & -1 & 3\end{array}\right] \sim \left[\begin{array}{cc|c}1 & -2 & -2\\0 & 3 & 7\end{array}\right]$$

$$\sim \left[\begin{array}{cc|c}1 & -2 & -2\\0 & 1 & 7/3\end{array}\right] \sim \left[\begin{array}{cc|c}1 & 0 & 8/3\\0 & 1 & 7/3\end{array}\right]$$

$$v = \begin{bmatrix}8/3\\7/3\end{bmatrix}$$

Check

$$\begin{pmatrix}2 & -1\\-1 & 2\end{pmatrix}\begin{pmatrix}8/3\\7/3\end{pmatrix} = \begin{pmatrix}(16-7)/3\\(-8+14)/3\end{pmatrix} = \begin{bmatrix}3\\2\end{bmatrix} \checkmark$$

$$x_W = \begin{bmatrix}1 & -1\\1 & 0\\0 & 1\end{bmatrix}\begin{bmatrix}8/3\\7/3\end{bmatrix} = \frac{8}{3}\begin{bmatrix}1\\1\\0\end{bmatrix} + \frac{7}{3}\begin{bmatrix}-1\\0\\1\end{bmatrix} = \begin{bmatrix}1/3\\8/3\\7/3\end{bmatrix} \text{ in } W$$

Recipe for Computing $x = x_W + x_{W^\perp}$

- Write W as a column space of a matrix A.
- Find a solution v of $A^TAv = A^Tx$ (by row reducing).
- Then $\boxed{x_W = Av}$ and $\boxed{x_{W^\perp} = x - x_W}$.

dist: $\boxed{2\sqrt{3}/3}$

$$x_{W^\perp} = x - x_W = \begin{bmatrix}1\\2\\3\end{bmatrix} - \begin{bmatrix}1/3\\8/3\\7/3\end{bmatrix} = \begin{bmatrix}2/3\\-2/3\\2/3\end{bmatrix} \text{ in } W^\perp.$$

$$\boxed{\begin{bmatrix}1\\2\\3\end{bmatrix} = \begin{bmatrix}1/3\\8/3\\7/3\end{bmatrix} + \begin{bmatrix}2/3\\-2/3\\2/3\end{bmatrix}}$$

(b) $\|x_{W^\perp}\| = \sqrt{\frac{4}{9} + \frac{4}{9} + \frac{4}{9}} = \sqrt{12}/3$

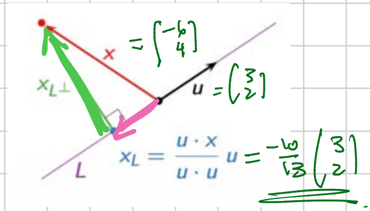
Orthogonal Projection onto a Line
Example

Problem: Compute the orthogonal projection of $x = \binom{-6}{4}$ onto the line L spanned by $u = \binom{3}{2}$, and find the distance from u to L.

Projection onto a Line

The projection of x onto a line $L = \mathrm{Span}\{u\}$ is

$$x_L = \frac{u \cdot x}{u \cdot u}\, u \qquad x_{L^\perp} = x - x_L.$$



X_L = projection of X onto line L

$$= \frac{u \cdot x}{u \cdot u} u$$

$$= \frac{\begin{bmatrix} 3 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} -6 \\ 4 \end{bmatrix}}{\begin{bmatrix} 3 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 2 \end{bmatrix}} \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

$$= \frac{-18+8}{9+4}\begin{bmatrix} 3 \\ 2 \end{bmatrix} = \frac{-10}{13}\begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} -30/13 \\ -20/13 \end{bmatrix}$$

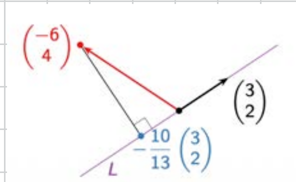
The closest vector in the line $L = \mathrm{Span}\{u\}$ to x.

Long way $\quad A = \begin{bmatrix} 3 \\ 2 \end{bmatrix} = u \quad L = \mathrm{Col}\ A$

https://textbooks.math.gatech.edu/ila/demos/projection.html?u1=3,2&vec=-6,4&labels=u&closed&mode=distance

$$A^TA = u^Tu = u \cdot u$$

$$A^Tx = u^Tx = u \cdot x$$



$$\boxed{\begin{aligned} &\text{Solve } A^TA v \overset{??}{=} A^Tx \quad ?? \\ &\Rightarrow u^Tu\, v = u^Tx \\ &\Rightarrow \end{aligned}}$$

Projection as a Transformation

Change in Perspective: let us consider orthogonal projection as a *transformation*.

Definition

Let W be a subspace of $\mathbf{R}^n$. Define a transformation

$$T\colon \mathbf{R}^n \longrightarrow \mathbf{R}^n \quad \text{by} \quad T(x) = x_W.$$

This transformation is also called **orthogonal projection** with respect to W.

Theorem

Let W be a subspace of $\mathbf{R}^n$ and let $T\colon \mathbf{R}^n \to \mathbf{R}^n$ be the orthogonal projection with respect to W. Then:

1. T is a *linear* transformation.
2. For every x in $\mathbf{R}^n$, $T(x)$ is the *closest* vector to x in W.
3. For every x in W, we have $T(x) = x$.
4. For every x in $W^\perp$, we have $T(x) = 0$.
5. $T \circ T = T$. ← $A^2 = A$
6. The range of T is W and the null space of T is $W^\perp$.

Projection Matrix

Example

Problem: Let $L = \mathrm{Span}\left\{\binom{3}{2}\right\}$ and let $T\colon \mathbf{R}^2 \to \mathbf{R}^2$ be the orthogonal projection onto L. Compute the matrix A for T.

How to find A such that $T(x) = Ax$? $A = [T(e_1) \;\; T(e_2)]$

(1) project e_1 to L to get first col of A

(2) project e_2 to L to get second col of A.

(1) $\dfrac{\begin{pmatrix}3\\2\end{pmatrix}\cdot\begin{pmatrix}1\\0\end{pmatrix}}{\begin{pmatrix}3\\2\end{pmatrix}\cdot\begin{pmatrix}3\\2\end{pmatrix}}\begin{pmatrix}3\\2\end{pmatrix} = \dfrac{3}{9}\begin{pmatrix}3\\2\end{pmatrix} = \begin{pmatrix}1\\2/3\end{pmatrix}$

(2) $\dfrac{\begin{pmatrix}3\\2\end{pmatrix}\cdot\begin{pmatrix}0\\1\end{pmatrix}}{\begin{pmatrix}3\\2\end{pmatrix}\cdot\begin{pmatrix}3\\2\end{pmatrix}}\begin{pmatrix}3\\2\end{pmatrix} = \dfrac{2}{9}\begin{pmatrix}3\\2\end{pmatrix} = \begin{pmatrix}2/3\\4/9\end{pmatrix}$

$$\boxed{A = \begin{pmatrix}1 & 2/3\\ 2/3 & 4/9\end{pmatrix}.}$$

You Try It!

Poll

Let A be the matrix for proj_W. What is/are the eigenvalue(s) of A?

A. 0 B. 1 C. -1 D. 0, 1 E. 1, -1 F. 0, -1 G. -1, 0, 1

Fact:

Theorem

Let W be an m-dimensional subspace of $\mathbf{R}^n$, let $T\colon \mathbf{R}^n \to W$ be the projection, and let A be the matrix for T. Then:

1. $\mathrm{Col}\,A = W$, which is the 1-eigenspace.
2. $\mathrm{Nul}\,A = W^\perp$, which is the 0-eigenspace.
3. $A^2 = A$.
4. A is similar to the diagonal matrix with m ones and $n - m$ zeros on the diagonal.

Projection Matrix
Another Example

Problem: Let

$$W = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \text{ in } \mathbf{R}^3 \;\middle|\; x_1 - x_2 + x_3 = 0 \right\}$$

and let $T: \mathbf{R}^3 \to \mathbf{R}^3$ be orthogonal projection onto W. Compute the matrix B for T.

Solution:

In the slides for the last lecture we computed $W = \operatorname{Col} A$ for

$$A = \begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

To compute $T(e_i)$ we have to solve the matrix equation $A^T A v = A^T e_i$. We have

$$A^T A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \qquad A^T e_i = \text{the } i\text{th column of } A^T = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}.$$

$$\left(\begin{array}{cc|c} 2 & -1 & 1 \\ -1 & 2 & -1 \end{array}\right) \xrightarrow{\text{RREF}} \left(\begin{array}{cc|c} 1 & 0 & 1/3 \\ 0 & 1 & -1/3 \end{array}\right) \implies T(e_1) = \frac{1}{3} A \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$$

$$\left(\begin{array}{cc|c} 2 & -1 & 1 \\ -1 & 2 & 0 \end{array}\right) \xrightarrow{\text{RREF}} \left(\begin{array}{cc|c} 1 & 0 & 2/3 \\ 0 & 1 & 1/3 \end{array}\right) \implies T(e_2) = \frac{1}{3} A \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

$$\left(\begin{array}{cc|c} 2 & -1 & 0 \\ -1 & 2 & 1 \end{array}\right) \xrightarrow{\text{RREF}} \left(\begin{array}{cc|c} 1 & 0 & 1/3 \\ 0 & 1 & 2/3 \end{array}\right) \implies T(e_2) = \frac{1}{3} A \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}$$

$$\implies B = \frac{1}{3} \begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix}.$$

Fact:

Theorem
Let $\{v_1, v_2, \ldots, v_m\}$ be a *linearly independent* set in $\mathbf{R}^n$, and let

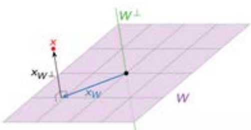
$$A = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_m \\ | & | & & | \end{pmatrix}.$$

Then the $m \times m$ matrix $A^T A$ is invertible.

Summary

Recall: Let W be a subspace of $\mathbf{R}^n$.

- The **orthogonal complement** $W^\perp$ is the set of vectors orthogonal to everything in W.
- The **orthogonal decomposition** of a vector x with respect to W is the unique way of writing $x = x_W + x_{W^\perp}$ for x_W in W and $x_{W^\perp}$ in $W^\perp$.
- The vector x_W is the **orthogonal projection** of x onto W. It is the closest vector to x in W.
- To compute x_W, write W as $\operatorname{Col} A$ and solve $A^T A v = A^T x$; then $x_W = Av$.



- If $W = L = \operatorname{Span}\{u\}$ is a line then $x_L = \frac{u \cdot x}{u \cdot u} u$.
- Orthogonal projection is a linear transformation.
- The matrix for projection onto W has eigenvalues 1 and 0 with eigenspaces W and $W^\perp$.
- A projection matrix is diagonalizable.

Section 6.5

The Method of Least Squares

Least Squares Solutions

We now are in a position to solve the motivating problem of this third part of the course:

> **Problem**
> Suppose that $Ax = b$ does not have a solution. What is the best possible approximate solution?

want instead Ax as close to b as possible

$\|Ax - b\|$ as small as possible.

Let A be an $m \times n$ matrix.

Definition
A **least squares solution** of $Ax = b$ is a vector $\widehat{x}$ in $\mathbf{R}^n$ such that

$$\|b - A\widehat{x}\| \leq \|b - Ax\|$$

Theorem
The least squares solutions of $Ax = b$ are the solutions of

$$\boxed{(A^T A)\widehat{x} = A^T b.}$$

(normal eqns)

Note we compute $\widehat{x}$ directly, without computing $\widehat{b}$ first.

Main idea:

> To say $Ax = b$ does not have a solution means that b is not in Col A.
> The closest possible $\widehat{b}$ for which $Ax = \widehat{b}$ does have a solution is $\widehat{b} = b_{\text{Col } A}$.
> Then $A\widehat{x} = \widehat{b}$ is a consistent equation.
> A solution $\widehat{x}$ to $A\widehat{x} = \widehat{b}$ is a **least squares solution.**

How to compute?

Conclusion: $\widehat{x}$ is just a solution of $A^T A v = A^T b$

Least Squares Solutions
Example

Find the least squares solutions of $Ax = b$ where:

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix} \qquad b = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}.$$

idea $\left[\begin{array}{cc|c} 0 & 1 & 6 \\ 1 & 1 & 0 \\ 2 & 1 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right]$ inconsistent system $Ax = b$.

Solve $A^T A x = A^T b$ solve for x to get a least squares solution.

$$A^T A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 3 \\ 3 & 3 \end{bmatrix}$$

$$A^T b = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 6 \end{bmatrix}$$

row reduce $[A^T A \mid A^T b] = \left[\begin{array}{cc|c} 5 & 3 & 0 \\ 3 & 3 & 6 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 5 & 3 & 0 \end{array}\right] \underset{-5R_1+R_2}{\sim} \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & -2 & -10 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & -3 \\ 0 & 1 & 5 \end{array}\right]$

So least squares solution is $\boxed{\widehat{x} = \begin{bmatrix} -3 \\ 5 \end{bmatrix}}$

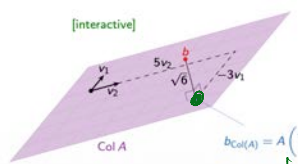
And the closest vector in Col A to $b = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$ is $A\widehat{x} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix}\begin{pmatrix} -3 \\ 5 \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \\ -1 \end{pmatrix} = b_W$

Least Squares Solutions
Example, continued

How close did we get?

$$\widehat{b} = A\widehat{x} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} -3 \\ 5 \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \\ -1 \end{pmatrix}$$

$$= -3\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} + 5\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = b_W$$



$= \begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix}$

Note that $\boxed{\begin{pmatrix} -3 \\ 5 \end{pmatrix}}$ records the coefficients of v_1 and v_2 in $\widehat{b}$.

Least Squares Solutions
Second example

Find the least squares solutions of $Ax = b$ where:

$$A = \begin{pmatrix} 2 & 0 \\ -1 & 1 \\ 0 & 2 \end{pmatrix} \qquad b = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

META. ① compute A^TA and A^Tb.

② row reduce $[A^TA \mid A^Tb]$ to get soln $\hat{x}$.

$A^TA x = A^Tb$.

$$① \; A^TA = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ -1 & 1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 5 & -1 \\ -1 & 5 \end{bmatrix}$$

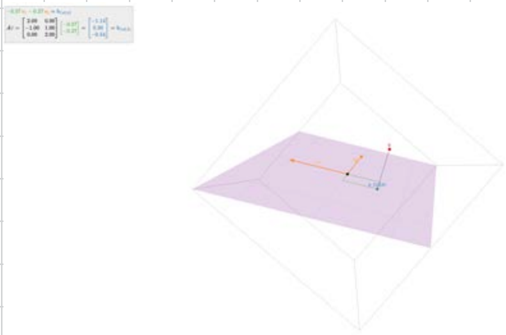
$$A^Tb = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$$

$-\frac{5}{3} + 2 =$

$$② \; \left[\begin{array}{cc|c} 5 & -1 & 2 \\ -1 & 5 & -2 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & -5 & 2 \\ 5 & -1 & 2 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & -5 & 2 \\ 0 & 24 & -8 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & -5 & 2 \\ 0 & 1 & -1/3 \end{array}\right]$$

$$\sim \left[\begin{array}{cc|c} 1 & 0 & 1/3 \\ 0 & 1 & -1/3 \end{array}\right] \qquad \boxed{\hat{x} = \begin{bmatrix} 1/3 \\ -1/3 \end{bmatrix}}$$ least squares soln

https://textbooks.math.gatech.edu/ila/demos/leastsquares.html?v1=2,-1,0&v2=0,1,2&vec=1,0,-1&range=3



Least Squares Solutions
Uniqueness

When does $Ax = b$ have a *unique* least squares solution $\hat{x}$?

Theorem

Let A be an $m \times n$ matrix. The following are equivalent:

1. $Ax = b$ has a *unique* least squares solution for all b in $\mathbf{R}^m$.
2. The columns of A are linearly independent.
3. A^TA is invertible.

In this case, the least squares solution is $(A^TA)^{-1}(A^Tb)$.

Why? If the columns of A are linearly *dependent*, then $A\hat{x} = \hat{b}$ has many solutions:

Note: A^TA is always a square matrix, but it need not be invertible.

Application
Data modeling: best fit line

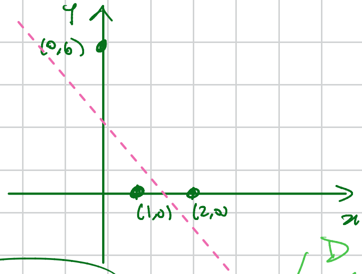
Find the best fit line through $(0,6)$, $(1,0)$, and $(2,0)$.

The general equation of a line is

$$y = C + Dx.$$

slope

y-intercept



Key insight:

So we want to solve:

$$6 = C + D \cdot 0$$
$$0 = C + D \cdot 1$$
$$0 = C + D \cdot 2.$$

Note to self: this is the first example from the lecture.

Plug in $(0,6)$, $(1,0)$, $(2,0)$ into line equation

$$\begin{cases} 6 = C + D(0) \\ 0 = C + D(1) \\ 0 = C + D(2) \end{cases}$$

$$\boxed{\begin{array}{l} C = 6 \\ \begin{cases} D + C = 0 \\ 2D + C = 0 \end{cases} \end{array}}$$

D coeff

C coeff

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} \quad b = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$$

least squares soln $\hat{X} = \begin{bmatrix} -3 \\ 5 \end{bmatrix}$

← best D-value

← best C-value

best line $\boxed{y = 5 - 3x}$

Summary

- A **least squares solution** of $Ax = b$ is a vector $\widehat{x}$ such that $\widehat{b} = A\widehat{x}$ is as *close to b as possible*.
- This means that $\widehat{b} = b_{\text{Col} A}$.
- One way to compute a least squares solution is by solving the system of equations
$$(A^T A)\widehat{x} = A^T b.$$
Note that $A^T A$ is a (symmetric) square matrix.
- Least-squares solutions are unique when the columns of A are linearly independent.
- You can use least-squares to find best-fit lines, parabolas, ellipses, planes, etc.

Story time ...

Application
Best fit parabola

What least squares problem $Ax = b$ finds the best parabola through the points $(-1, 0.5)$, $(1, -1)$, $(2, -0.5)$, $(3, 2)$?

The general equation for a parabola is

$$\boxed{y = Ax^2 + Bx + C.}$$

Key plug in points into model

(pt 1) $0.5 = A(-1)^2 + B(-1) + C$

(pt 2) $-1 = A(1)^2 + B(1) + C$

(pt 3) $-.5 = A(2)^2 + B(2) + C$

(pt 4) $2 = A(3)^2 + B(3) + C$

So we want to solve:

$$\begin{aligned} 0.5 &= A(-1)^2 + B(-1) + C \\ -1 &= A(1)^2 + B(1) + C \\ -0.5 &= A(2)^2 + B(2) + C \\ 2 &= A(3)^2 + B(3) + C \end{aligned}$$

In matrix form:

$$\begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \end{pmatrix} \begin{pmatrix} A \\ B \\ C \end{pmatrix} = \begin{pmatrix} 0.5 \\ -1 \\ -0.5 \\ 2 \end{pmatrix}.$$

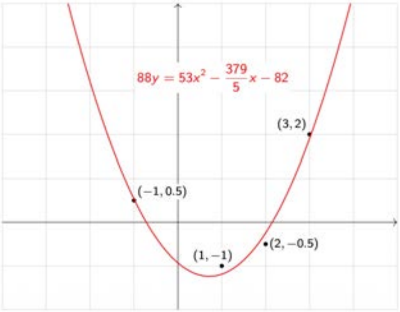
M

$$M^T M = \begin{pmatrix} 1 & 1 & 4 & 9 \\ -1 & 1 & 2 & 3 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 99 & 35 & 15 \\ 35 & 15 & 5 \\ 15 & 5 & 4 \end{pmatrix}$$

$$M^T b = \begin{pmatrix} 1 & 1 & 4 & 9 \\ -1 & 1 & 2 & 3 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0.5 \\ -1 \\ -0.5 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{31}{2} \\ \frac{7}{2} \\ 1 \end{pmatrix}$$

https://textbooks.math.gatech.edu/ila/demos/bestfit.html?func=A*x%5E2+B*x+C&v1=-1,.5&v2=1,-1&v3=2,-.5&v4=3,2&range=5

$$M^T M = \begin{pmatrix} 99 & 35 & 15 \\ 35 & 15 & 5 \\ 15 & 5 & 4 \end{pmatrix} \quad M^T b = \begin{pmatrix} \frac{31}{2} \\ \frac{7}{2} \\ 1 \end{pmatrix}$$



$$\left[\begin{array}{ccc|c} 99 & 35 & 15 & \frac{31}{2} \\ 35 & 15 & 5 & \frac{7}{2} \\ 15 & 5 & 4 & 1 \end{array}\right] \rightsquigarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{341}{6292} \\ 0 & 1 & 0 & -\frac{4479}{5720} \\ 0 & 0 & 1 & -\frac{41}{44} \end{array}\right]$$

$$\boxed{A = \frac{341}{6292} \approx 0.0542, \qquad B = -\frac{4479}{5720} \approx -0.78299, \qquad C = -\frac{41}{44} \approx -0.93182.}$$

Application

Best-fit Trigonometric Function

For fun: what is the best-fit function of the form

$$y = A + B\cos(x) + C\sin(x) + D\cos(2x) + E\sin(2x) + F\cos(3x) + G\sin(3x)$$

passing through the points

$$\begin{pmatrix}-4\\-1\end{pmatrix}, \begin{pmatrix}-3\\0\end{pmatrix}, \begin{pmatrix}-2\\-1.5\end{pmatrix}, \begin{pmatrix}-1\\.5\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix}, \begin{pmatrix}2\\-.5\end{pmatrix}, \begin{pmatrix}3\\2\end{pmatrix}, \begin{pmatrix}4\\-1\end{pmatrix}?$$

$$M = \begin{pmatrix}
1 & \cos(-4) & \sin(-4) & \cos(-8) & \sin(-8) & \cos(-12) & \sin(-12) \\
1 & \cos(-3) & \sin(-3) & \cos(-6) & \sin(-6) & \cos(-9) & \sin(-9) \\
1 & \cos(-2) & \sin(-2) & \cos(-4) & \sin(-4) & \cos(-6) & \sin(-6) \\
1 & \cos(-1) & \sin(-1) & \cos(-2) & \sin(-2) & \cos(-3) & \sin(-3) \\
1 & 1 & 0 & 1 & 0 & 1 & 0 \\
1 & \cos(1) & \sin(1) & \cos(2) & \sin(2) & \cos(3) & \sin(3) \\
1 & \cos(2) & \sin(2) & \cos(4) & \sin(4) & \cos(6) & \sin(6) \\
1 & \cos(3) & \sin(3) & \cos(6) & \sin(6) & \cos(9) & \sin(9) \\
1 & \cos(4) & \sin(4) & \cos(8) & \sin(8) & \cos(12) & \sin(12)
\end{pmatrix}, \qquad b = \begin{pmatrix}-1\\0\\-1.5\\0.5\\1\\-1\\-0.5\\2\\-1\end{pmatrix}$$

$$M \approx \begin{pmatrix}
1 & -0.65 & 0.76 & -0.15 & -0.99 & 0.84 & 0.54 \\
1 & -0.99 & 0.14 & 0.96 & 0.28 & -0.91 & -0.41 \\
1 & -0.42 & -0.91 & -0.65 & 0.76 & 0.96 & 0.28 \\
1 & 0.54 & -0.84 & -0.42 & -0.91 & -0.99 & -0.14 \\
1 & 1.00 & 0.00 & 1.00 & 0.00 & 1.00 & 0.00 \\
1 & 0.54 & 0.84 & -0.42 & 0.91 & -0.99 & 0.14 \\
1 & -0.42 & 0.91 & -0.65 & -0.76 & 0.96 & -0.28 \\
1 & -0.99 & -0.14 & 0.96 & -0.28 & -0.91 & 0.41 \\
1 & -0.65 & -0.76 & -0.15 & 0.99 & 0.84 & -0.54
\end{pmatrix}.$$

$$M^TM \approx \begin{pmatrix}
9.00 & -2.04 & 0.00 & 0.33 & 0.00 & -0.20 & 0.00 \\
-2.04 & 4.50 & 0.00 & -1.12 & 0.00 & 0.47 & 0.00 \\
0.00 & 0.00 & 4.50 & 0.00 & -0.16 & 0.00 & 0.09 \\
0.33 & -1.12 & 0.00 & 4.50 & 0.00 & -0.55 & 0.00 \\
0.00 & 0.00 & -0.16 & 0.00 & 4.50 & 0.00 & -0.31 \\
-0.20 & 0.47 & 0.00 & -0.55 & 0.00 & 4.50 & 0.00 \\
0.00 & 0.00 & 0.09 & 0.00 & -0.31 & 0.00 & 4.50
\end{pmatrix}, \qquad M^Tb \approx \begin{pmatrix}-1.50\\1.69\\-1.55\\0.90\\0.76\\-2.29\\2.02\end{pmatrix}$$

$$\left[\begin{array}{ccccccc|c}
9.00 & -2.04 & 0 & 0.33 & 0 & -0.20 & 0 & -1.50 \\
-2.04 & 4.50 & 0 & -1.12 & 0 & 0.47 & 0 & 1.69 \\
0 & 0 & 4.50 & 0 & -0.16 & 0 & 0.09 & -1.55 \\
0.33 & -1.12 & 0 & 4.50 & 0 & -0.55 & 0 & 0.90 \\
0 & 0 & -0.16 & 0 & 4.50 & 0 & -0.31 & 0.76 \\
-0.20 & 0.47 & 0 & -0.55 & 0 & 4.50 & 0 & -2.29 \\
0 & 0 & 0.09 & 0 & -0.31 & 0 & 4.50 & 2.02
\end{array}\right] \qquad
\left[\begin{array}{ccccccc|c}
1 & 0 & 0 & 0 & 0 & 0 & 0 & -0.07 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0.37 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & -0.34 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0.29 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0.16 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & -0.47 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0.46
\end{array}\right]$$

$A \approx -0.07,$
$B \approx 0.37,$
$C \approx -0.34,$
$D \approx 0.29,$
$E \approx 0.16,$
$F \approx -0.47,$
$G \approx 0.46.$

https://textbooks.math.gatech.edu/ila/demos/bestfit.html?func=A+B*cos(x)+C*sin(x)+D*cos(2*x)+EE*sin(2*x)+F*cos(3*x)+G*sin(3*x)&v1=-4,-1&v2=-3,0&v3=-2,-1.5&v4=-1,.5&v5=0,1&v6=1,-1&v7=2,-.5&v8=3,2&v9=4,-1&range=5