MA1553

Chapter 6

Orthogonality

Section 6.1

Dot Products and Orthogonality

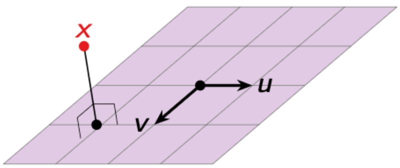
Orientation

Recall: This course is about learning to:

- Solve the matrix equation $Ax = b$
- Solve the matrix equation $Ax = \lambda x$
- Almost solve the equation $Ax = b$

We are now aiming at the last topic.

Idea: In the real world, data is imperfect. Suppose you measure a data point x 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

$$x \cdot y = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \stackrel{\text{def}}{=} x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

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

Example

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} =$$

Example

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} =$$

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

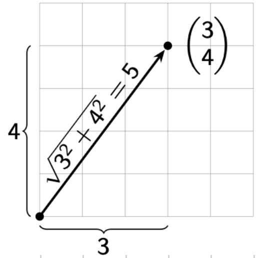
The **length** or **norm** of a vector x in $\mathbf{R}^n$ is

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

Fact

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

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



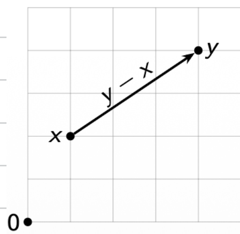
$$\left\|\begin{pmatrix}3\\4\end{pmatrix}\right\| =$$

The Dot Product and Distance

Definition

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

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



Example

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

$$\text{dist}(x, y) =$$

Unit Vectors

Definition

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

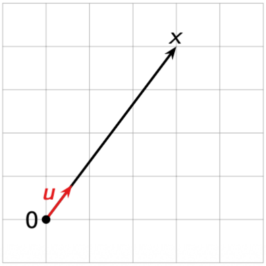
Example

The unit coordinate vectors are unit vectors:

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

Definition

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



Example

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

Orthogonality

Definition

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

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.

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

Orthogonality
Example

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

Key insight:

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

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:

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

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

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!

Summary

- The **dot product** of vectors x, y in $\mathbf{R}^n$ is the number $x^T y$.
- The **length** or **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 $\mathrm{dist}(x, y) = \|y - x\|$.
- 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

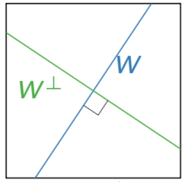
Orthogonal Complements

Orthogonal Complements

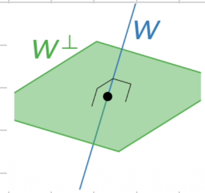
Definition
Let W be a subspace of $\mathbf{R}^n$. Its **orthogonal complement** is

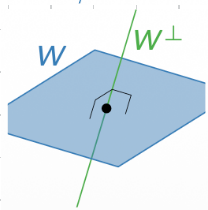
$$W^{\perp} = \{v \text{ in } \mathbf{R}^n \mid v \cdot w = 0 \text{ for all } w \text{ in } W\} \qquad \text{read “}W \text{ perp”}.$$

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



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?

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

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

Hint: use (4.)

(4.) says:

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

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

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

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 $\dim(\text{Row}\ A) = \dim(\text{Col}\, A)$.

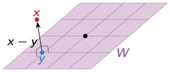
Idea: the dimension of Row A is equal to the number of pivots of A, which is the dimension of the column space of A.

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

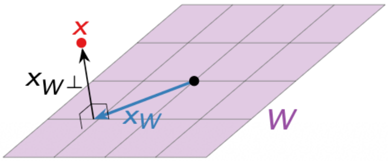
$$x = x_W + x_{W^\perp}$$

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]

$$x = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}$$

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 $A^T A$ 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

$$A^T A v = A^T x \qquad \text{(in the unknown vector } v\text{)}$$

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 $A^T A v = A^T x$ (by row reducing).
- Then $x_W = Av$ and $x_{W^\perp} = x - x_W$.

The $A^T A$ 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.

The A^TA Trick

Another Example

Problem: Let

$$x = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \qquad W = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \text{ in } \mathbf{R}^3 \;\middle|\; x_1 - x_2 + x_3 = 0 \right\}.$$

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 $x_{W^\perp} = x - x_W$.

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

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

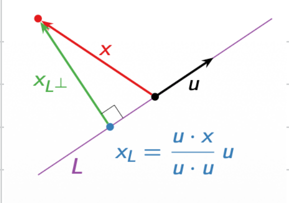
Orthogonal Projection onto a Line
Example

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.

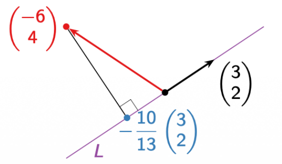
Projection onto a Line

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

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



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



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

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

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. $T \circ T = T$.
6. The range of T is W and the null space of T is $W^\perp$.

Projection Matrix
Example

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

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. $\text{Col}\, A = W$, which is the 1-eigenspace.
2. $\text{Nul}\, A = W^\perp$, 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 $T\colon \mathbf{R}^3 \to \mathbf{R}^3$ 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 $A^TAv = A^Te_i$. We have

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

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

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

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

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

Fact:

Theorem

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

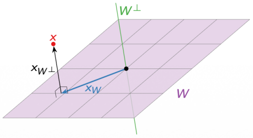
$$A = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_m \\ | & | & & | \end{pmatrix}.$$

Then the $m \times m$ matrix A^TA 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 $A^TAv = A^Tx$; 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?

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

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

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

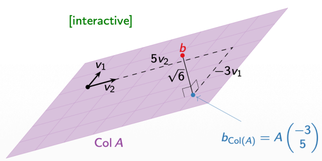
Find the least squares solutions of $Ax = b$ where:

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

Least Squares Solutions
Example, continued

How close did we get?

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



Note that

$$\begin{pmatrix} -3 \\ 5 \end{pmatrix}$$

records the coefficients of v_1 and v_2 in $\widehat{b}$.

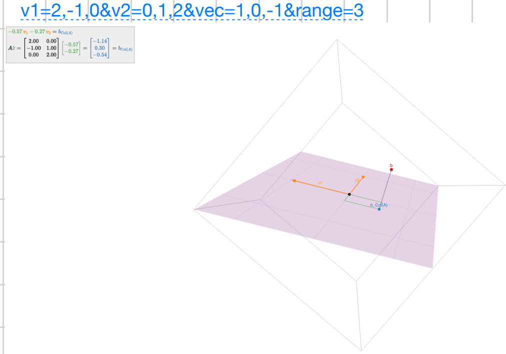
Least Squares Solutions

Second example

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

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

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 $\widehat{x}$?

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. $A^T A$ is invertible.

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

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

Note: $A^T A$ 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

$$y = C + Dx.$$

Key insight:

So we want to solve:

$$\begin{aligned} 6 &= C + D \cdot 0 \\ 0 &= C + D \cdot 1 \\ 0 &= C + D \cdot 2. \end{aligned}$$

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

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 $\widehat{b} = b_{\mathrm{Col}\,A}$.
- 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.