Class 10: 10/9

Reminder: Midterm I is one week from today! Details TBA.

<u>Outline</u>

I. Projection onto spaces
II. Least squares approximation.
III. Orthogonal matrices + Gram Schmidt.

I. See yesterday's notes.

II. Say we can't solve $A\vec{x} = \vec{b}$ ($\vec{b}$ is not in the column space of A)

What do we do instead?

* Instead, solve $A^TA\hat{x} = A^T\vec{b}$.

Why? We want to find $\hat{x}$ so that $A\hat{x}$ is as close to $\vec{b}$ as possible.

That is, if $\vec{e} = A\hat{x} - \vec{b}$, we want to minimize $\|\vec{e}\|$, and find the vector in the column space of A closest to $\vec{b}$.

* This is the projection of $\vec{b}$ onto the column space of A.

The solution $\hat{x}$ that minimizes $\|\vec{e}\|$ is called the <u>least</u> <u>squares</u> solution.

<u>Applications</u>

Fitting data:

Example: Find the best fit line for the points $(0,8)$, $(1,2)$ and $(2,0)$.

We want to solve

$$\begin{aligned} C + D\cdot 0 &= 8 \\ C + D\cdot 1 &= 2 \\ C + D\cdot 2 &= 0 \end{aligned} \quad \rightsquigarrow \quad \underbrace{\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}}_{A} \begin{bmatrix} C \\ D \end{bmatrix} = \underbrace{\begin{bmatrix} 8 \\ 2 \\ 0 \end{bmatrix}}_{\vec{b}}.$$

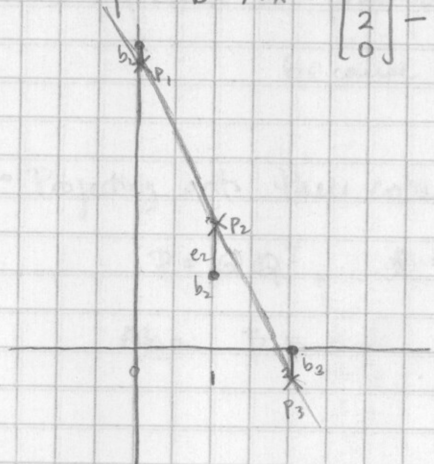
But there's no solution.

Solve $A^TA\hat{x} = A^T\vec{b}$ instead: $\begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 10 \\ 2 \end{bmatrix}$

$$\rightsquigarrow \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 22/3 \\ -4 \end{bmatrix}$$

So $\vec{p} = A\hat{x} = \begin{bmatrix} 22/3 \\ 10/3 \\ -2/3 \end{bmatrix}$

$\vec{e} = \vec{b} - \vec{p} = \vec{b} - A\hat{x} = \begin{bmatrix} 8 \\ 2 \\ 0 \end{bmatrix} - \begin{bmatrix} 22/3 \\ 10/3 \\ -2/3 \end{bmatrix} = \begin{bmatrix} 2/3 \\ -4/3 \\ 2/3 \end{bmatrix}$



Hypothetical question: Fitting a parabola?

III. Orthogonal matrices.

A projection matrix has the formula $P = A(A^TA)^{-1}A^T$.

Complicated! To simplify, would be nice if $A^TA = I$.

<u>Definition</u>: A set of vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ is <u>orthonormal</u> if
1) $\vec{v}_i \cdot \vec{v}_j = 0$ if $i \neq j$
2) $\vec{v}_i \cdot \vec{v}_i = 1$

(vectors are unit vectors and orthogonal to each other).

A matrix Q with orthonormal columns satisfies $Q^TQ = I$.

If Q is square, this means $Q^T = Q^{-1}$, and we call such a matrix an <u>orthogonal matrix</u>.

<u>Example</u>: $Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ (From prev. HW).

or $Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ (permutation matrix).

Matrices with orthonormal columns are awesome:

- They preserve lengths: $\|Q\vec{x}\| = \|\vec{x}\|$ (like rotation matrix, permutation matrix)
- They preserve dot products: $(Q\vec{x})\cdot(Q\vec{y}) = (\vec{x}\cdot\vec{y})$

  because $(Q\vec{x})\cdot(Q\vec{y}) = (Q\vec{x})^T Q y = \vec{x}^T Q^T Q \vec{y} = \vec{x}^T\vec{y}$.

- Projecting onto their column space is easy

$$P = QQ^T, \quad \hat{x} = Q^T\vec{b}, \quad \vec{p} = Q\hat{x}$$

<u>Also</u>: $\vec{p} = \begin{bmatrix} | & & | \\ q_1 & \cdots & q_n \\ | & & | \end{bmatrix} \begin{bmatrix} \vec{q}_1^T\vec{b} \\ \vdots \\ q_n^T\vec{b} \end{bmatrix} = \vec{q}_1\vec{q}_1^T\vec{b} + \vec{q}_2\vec{q}_2^T\vec{b} + \dots + \vec{q}_n\vec{q}_n^T\vec{b}$

<u>Note</u>: $\vec{p}$ = sum of projections of $\vec{b}$ onto each line spanned by q_i's

<u>If</u> you wanted to, you could project separately onto each column, then sum

Example: Project $\begin{bmatrix}1\\2\\2\end{bmatrix}$ onto the column space of $Q = \begin{bmatrix} 1/\sqrt{3} & 0 \\ 1/\sqrt{3} & 1/\sqrt{2} \\ -1/\sqrt{3} & 1/\sqrt{2} \end{bmatrix}$

Method 1: $QQ^T = \begin{bmatrix} 1/\sqrt{3} & 0 \\ 1/\sqrt{3} & 1/\sqrt{2} \\ -1/\sqrt{3} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & -1/\sqrt{3} \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1/3 & 1/3 & -1/3 \\ 1/3 & 5/6 & 1/6 \\ -1/3 & 1/6 & 5/6 \end{bmatrix}$

So $\vec{p} = QQ^T\vec{b} = \begin{bmatrix} 1/3 \\ 7/3 \\ 5/3 \end{bmatrix}$

Method 2: $\vec{q}_1\vec{q}_1^T = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ -1/\sqrt{3} \end{bmatrix} \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & -1/\sqrt{3} \end{bmatrix} = \begin{bmatrix} 1/3 & 1/3 & -1/3 \\ 1/3 & 1/3 & -1/3 \\ -1/3 & -1/3 & 1/3 \end{bmatrix}$

So $\vec{q}_1\vec{q}_1^T\vec{b} = \begin{bmatrix} 1/3 \\ 1/3 \\ -1/3 \end{bmatrix}$

$\vec{q}_2\vec{q}_2^T\vec{b} = \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 0 & 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}\vec{b} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}$

Q: How do we make matrices with orthonormal columns?

A: Gram-Schmidt.

Explanation by example:

Say you start with 3 independent vectors, $\vec{a}, \vec{b}, \vec{c}$, and you want 3 orthonormal vectors that span the same space.

1) Choose $\vec{A} = \vec{a}$. To make $\vec{B} \perp \vec{A}$, subtract off the projection of $\vec{b}$ onto $\vec{A}$

$$\vec{B} = \vec{b} - \frac{\vec{A}^T \vec{b}}{\vec{A}^T \vec{A}} \vec{A} = \vec{b} - \frac{\vec{a}^T \vec{b}}{\vec{a}^T \vec{a}} \vec{a}$$

2) How do we choose C? It needs to be perpendicular to the space spanned by A and B. Subtract off the projection of C onto this space, which is the sum of the projections onto A and B.

$$\vec{C} = \vec{c} - \frac{\vec{A}^T \vec{c}}{\vec{A}^T \vec{A}} \vec{A} - \frac{\vec{B}^T \vec{c}}{\vec{B}^T \vec{B}} \vec{B}$$

3) If there were more vectors, continue

4) Normalize.

$$\vec{q}_1 = \frac{\vec{A}}{\|\vec{A}\|}, \quad \vec{q}_2 = \frac{\vec{B}}{\|\vec{B}\|}, \quad \vec{q}_3 = \frac{\vec{C}}{\|\vec{C}\|}.$$

Fact: This gives a factorization of A

$$A = QR = (\text{orthonormal columns}) \cdot (\text{triangular})$$

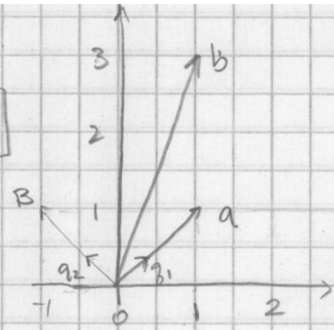
$$\begin{bmatrix} | & | & | \\ a & b & c \\ | & | & | \end{bmatrix} = \begin{bmatrix} | & | & | \\ q_1 & q_2 & q_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} q_1^T a & q_1^T b & q_1^T c \\ 0 & q_2^T b & q_2^T c \\ 0 & 0 & q_3^T c \end{bmatrix}$$

$R = Q^T A$ since $Q^T A = Q^T Q R = R$.

(So $A = Q Q^T A$. But $Q Q^T \neq I$. What's going on?)

Example 2×2 example

$\vec{a} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$



$$A = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 3 \end{bmatrix} - \frac{a^Tb}{a^Ta} a = \begin{bmatrix} 1 \\ 3 \end{bmatrix} - \frac{4}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} - \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Normalize: $\quad q_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \qquad q_2 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$