Example from last Friday.

$$B = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 4 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} \right\} \qquad O = \left\{ \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Exercise 3a Find the $A = Q\,R$ factorization based on the data above, for

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 4 & -2 \\ 0 & 0 & 3 \end{bmatrix}$$

$A = Q\,R$

shortcut for R

$$Q^T A = \underbrace{Q^T Q}_{I} R = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 4 & -2 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} \sqrt{2} & 2\sqrt{2} & -\frac{1}{\sqrt{2}} \\ 0 & 2\sqrt{2} & -\frac{3}{\sqrt{2}} \\ 0 & 0 & 3 \end{bmatrix}$$

✓

$$\text{solution } A = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \underbrace{\begin{bmatrix} \sqrt{2} & 2\sqrt{2} & -\frac{1}{\sqrt{2}} \\ 0 & 2\sqrt{2} & -\frac{3}{\sqrt{2}} \\ 0 & 0 & 3 \end{bmatrix}}_{R}$$

Exercise 3b Further factor R into a diagonal matrix times a volume-preserving shear and interpret the transformation $T(\underline{\mathbf{x}}) = A\,\underline{\mathbf{x}}$ as a composition of (1) a volume preserving shear, followed by (2) a coordinate scaling that increases volume by a factor of 12, followed by a rotation about the x_3 axis in $\mathbb{R}^3$, which preserves volume. The generalization of this example explains why the determinant of A (or its absolute value in general) is the volume expansion factor for the transformation $T(\underline{\mathbf{x}}) = A\,\underline{\mathbf{x}}$.

$$\frac{-3}{\sqrt{2}\cdot 2\sqrt{2}}$$

$$R = \underbrace{\begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 2\sqrt{2} & 0 \\ 0 & 0 & 3 \end{bmatrix}}_{\text{vol} \times 12} \underbrace{\begin{bmatrix} 1 & 2 & -\frac{1}{2} \\ 0 & 1 & -3/4 \\ 0 & 0 & 1 \end{bmatrix}}_{\substack{\text{shear} \\ \text{vol pres.}}}$$

Q

Definition A square $n \times n$ matrix Q is called *orthogonal* if its columns are ortho-normal. (You can read more about orthogonal matrices at e.g. Wikipedia.)

Theorem. Let Q be an orthogonal matrix. Then

a) $Q^{-1} = Q^T$.

$$\text{entry}_{ij}\ Q^T Q = \left[\text{Row}_i\ Q^T\right]\left[\text{Col}_j\ Q\right] = \text{Col}_i Q \cdot \text{col}_j Q = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$$

$$\Rightarrow Q^T Q = I$$

$$\Rightarrow Q Q^T = I \text{ also (prev. Chapter).}$$

i.e. $Q^{-1} = Q^T$

b) The rows of Q are also ortho-normal.

$$Q Q^T = I$$

$$ij \text{ entry: } \text{row}_i Q \cdot \underbrace{\text{col}_j Q^T}_{\text{row}_j(Q)} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$$

c) the transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ given by

$$T(\underline{x}) = Q\,\underline{x}$$

preserves dot products and magnitudes, (so also volumes, since cubes generated by perpendicular vectors will be transformed into equal-volume cubes). In other words, for all $\underline{x}, \underline{y} \in \mathbb{R}^n$,

$$T(\underline{x}) \cdot T(\underline{y}) = \underline{x} \cdot \underline{y}$$

$$\|T(\underline{x})\| = \|\underline{x}\|.$$

d) The only matrix transformations $T : \mathbb{R}^n \to \mathbb{R}^n$ that preserve dot products are orthogonal transformations. (These transformations are often referred to as *isometries*.)

Tues Apr 10

- 6.5 Least squares "solutions" to $A\vec{x} = \vec{b}$, and projection revisited.

Announcements:

- graded HW + quizzes
- $\S$6.5 first, then go back to rest of Monday notes.

'til 12:56

Warm-up Exercise: What is the condition on $\vec{y}$ so that the system

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

is consistent?

low tech (Chptrs 1-2)

$$\left[\begin{array}{cc|c} 1 & 2 & y_1 \\ 0 & 1 & y_2 \\ 1 & 0 & y_3 \end{array}\right]$$

$R_3 \to R_1$, $R_1 \to R_3$

$$\left[\begin{array}{cc|c} 1 & 0 & y_3 \\ 0 & 1 & y_2 \\ 1 & 2 & y_1 \end{array}\right]$$

$-R_1 + R_3 \to R_3$

$$\left[\begin{array}{cc|c} 1 & 0 & y_3 \\ 0 & 1 & y_2 \\ 0 & 2 & y_1 - y_3 \end{array}\right]$$

$-2R_2 + R_3 \to R_3$

$$\left[\begin{array}{cc|c} 1 & 0 & y_3 \\ 0 & 1 & y_2 \\ 0 & 0 & y_1 - 2y_2 - y_3 \end{array}\right]$$

$\vec{y}$ is on plane with implicit eqtn

$$\boxed{y_1 - 2y_2 - y_3 = 0}$$

high tech Chapter 6. $(W^\perp)^\perp = W$.

$W^\perp$ for span$\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \right\}$

$W^\perp$: $\begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 2 & 1 & 0 & 0 \end{array}\right]$$

$-2R_1 + R_2 \to R_2$

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & -2 & 0 \end{array}\right]$$

$W^\perp$: span$\left\{ \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \right\}$

$(W^\perp)^\perp = W$

$\parallel$

$$\begin{bmatrix} -1 & 2 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = 0$$

$$\boxed{-y_1 + 2y_2 + y_3 = 0}$$

equivalent

(better in high dims)

Least squares solutions, section 6.5

In trying to fit experimental data to a linear model you must often find a "solution" to

$$A\,\underline{\mathbf{x}} = \underline{\boldsymbol{b}}$$

where no exact solution actually exists. Mathematically speaking, the issue is that $\underline{\boldsymbol{b}}$ is not in the range of the transformation

$$T(\underline{\mathbf{x}}) = A\,\underline{\mathbf{x}},$$

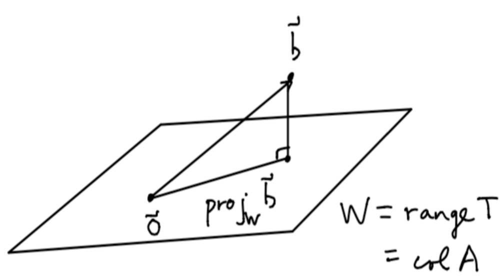
i.e.

$$\underline{\mathbf{x}} \notin Range\ T = Col\ A.$$

In such a case, the *least squares solution(s)* $\hat{\mathbf{x}}$ solve(s)

$$\boxed{A\,\hat{\mathbf{x}} = proj_{Col\ A}\,\underline{\boldsymbol{b}}\,.}$$

Thus, for the least squares solution(s), $A\,\hat{\mathbf{x}}$ is as close to $\underline{\mathbf{b}}$ as possible. Note that there will be a unique least squares solution $\hat{\mathbf{x}}$ if and only if $Nul\ A = \{\underline{\mathbf{0}}\}$, i.e. if and only if the columns of A are linearly independent. (Recall, any two solutions to the same nonhomogeneous matrix equation differ by a solution to the homogeneous equation.)



warmup

Exercise 1 Find the least squares solution to

$$\begin{array}{cc} \vec{w}_1 & \vec{w}_2 \end{array}$$

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}.$$

Note that the implicit equation of the plane spanned by the two columns of A is

- $$-y_1 + 2y_2 + y_3 = 0. \qquad -3+6+3 \neq 0$$

You know two ways to find that implicit equation (!) .....at least it's easy to check that the the two column vectors satisfy it. Since $[3\ 3\ 3]^T$ does not satisfy the implicit equation, there is no exact solution to this problem. If you wish, it could be instructive review the two ways.

You may use the Gram-Schmidt ortho-normal basis for *Col A*, namely

$$\begin{array}{cc} \vec{u}_1 & \vec{u}_2 \end{array}$$

$$\mathsf{O} = \left\{ \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \right\}. \leftarrow$$

Step 1 find $\operatorname{proj}_{Col A} \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix} = \hat{b}$

Step 2 Solve $A\hat{x} = \hat{b}$. $\hat{x}$ = "least squares soln"

$$\operatorname{proj}_W \vec{b} = (\vec{b}\cdot\vec{u}_1)\vec{u}_1 + (\vec{b}\cdot\vec{u}_2)\vec{u}_2$$

$$= \left(\begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}\cdot\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right)\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \left(\begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}\cdot\frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}\right)\frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

$$= \frac{1}{2}(6)\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \frac{1}{3}(3)\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

$$= 3\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix}$$

$$\left[\begin{array}{cc|c} 1 & 2 & 4 \\ 0 & 1 & 1 \\ 1 & 0 & 2 \end{array}\right]$$

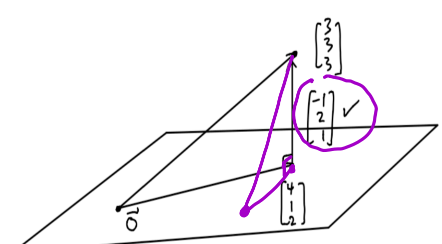
$$\begin{array}{l} R_3 \to R_1 \\ \\ R_1 \to R_3 \end{array} \left[\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 1 & 2 & 4 \end{array}\right]$$

$$\begin{array}{l} \\ \\ -R_1 - 2R_2 + R_3 \to R_3 \end{array} \left[\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{array}\right]$$

$x_1 = 2$

$x_2 = 1$

Solution:



$$\begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

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

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix}$$

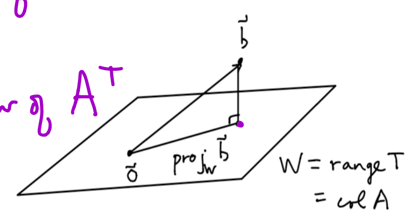
There's actually a smart way to find the least squares solutions that doesn't require an orthonormal basis for *Col A*. And as a result, it turns out that one can also compute projections onto a subspace without first constructing an orthonormal basis for the subspace !!! Consider the following chain of equivalent conditions on $\underline{\boldsymbol{x}}$:

$$\boxed{A\,\underline{\mathbf{x}} = proj_{Col\,A}\,\underline{\boldsymbol{b}}}$$

$$\vec{z} = b - A\,\underline{\mathbf{x}} \in (Col\,A)^{\perp}$$

$b - A\vec{x} \perp$ each col of A

$b - A\vec{x} \perp$ each row of A^T



$$A^T(\underline{\boldsymbol{b}} - A\,\underline{\mathbf{x}}) = \underline{\mathbf{0}}$$

$$A^T\underline{\boldsymbol{b}} - A^TA\,\underline{\mathbf{x}} = \underline{\mathbf{0}}$$

$$\boxed{A^TA\,\underline{\mathbf{x}} = A^T\underline{\boldsymbol{b}}\,.}$$

This last equation will always be consistent because projections exist. And if the columns of A are linearly independent the solutions to the top equation, and hence the final equation, will be unique. So the matrix A^TA will be invertible in that case. The final matrix equation is called the *normal equation* for least squares solutions.

Exercise 2 Re-do Exercise 1 using the normal equation, i.e find the least squares solution $\hat{\boldsymbol{x}}$ to

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}.$$

And then note that $A\,\hat{\boldsymbol{x}}$ is $\underbrace{proj_{Col\,A}\underline{\boldsymbol{b}}}_{\hat b}$, i.e. you found the projection of $[3\ 3\ 3]^T$ without ever finding and using an ortho-normal basis!!!

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

$$A^TA\,\hat{x} = A^T\vec{b}$$

after find $\hat{x}$,

$$\begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix}\begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}$$

$$A\hat{x} = proj_{Col\,A}\vec{b}$$

$$\begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 9 \end{bmatrix}$$

Same as in Exercise 1 !!

$$A\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix}$$

no Gram-Schmidt $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 5 & -2 \\ -2 & 2 \end{bmatrix}\begin{bmatrix} 6 \\ 9 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 12 \\ 6 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

Exercise 3 In the case that $A^T A$ is invertible we may take the normal equation for finding the least squares solution to $A\,x = b$ and find $A\,\hat{x} = proj_{Col\,A} b$ directly:

$$\boxed{A^T A\,\underline{\mathbf{x}} = A^T\,\underline{\mathbf{b}}}$$

$$\underline{\mathbf{x}} = \left(A^T A\right)^{-1} A^T\,\underline{\mathbf{b}}$$

$$\underline{proj_{Col\,A}\underline{\mathbf{b}}} = A\,\underline{\mathbf{x}} = \underline{A\left(A^T A\right)^{-1} A^T\underline{\mathbf{b}}}.$$

Verify for the third time that for $W = span\left\{\underbrace{\begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}2\\1\\0\end{bmatrix}}\right\}$, $proj_W\begin{bmatrix}3\\3\\3\end{bmatrix} = \begin{bmatrix}4\\1\\2\end{bmatrix}$ by "plug and chug".

$$proj_W \begin{bmatrix}3\\3\\3\end{bmatrix} = \begin{bmatrix}1&2\\0&1\\1&0\end{bmatrix}\begin{bmatrix}2&2\\2&5\end{bmatrix}^{-1}\begin{bmatrix}1&0&1\\2&1&0\end{bmatrix}\begin{bmatrix}3\\3\\3\end{bmatrix}$$

$$= \begin{bmatrix}1&2\\0&1\\1&0\end{bmatrix}\frac{1}{6}\begin{bmatrix}5&-2\\-2&2\end{bmatrix}\begin{bmatrix}6\\9\end{bmatrix}$$

See previous page

$$= \begin{bmatrix}4\\1\\2\end{bmatrix}$$