Solution Key HW # 7

Section 3.2 # 1, 3, 12, 18
Section 3.3 # 1, 2, 3, 4, 8, 14

Section 3.2

• #1.

a) $\vec{b} = (\sqrt{x}, \sqrt{y}) \qquad \vec{a} = (\sqrt{y}, \sqrt{x})$

$$\vec{a}^T\vec{b} = \sqrt{xy} + \sqrt{xy} = 2\sqrt{xy}$$

$$\|a\| = \sqrt{x+y}$$

$$\|b\| = \sqrt{x+y}$$

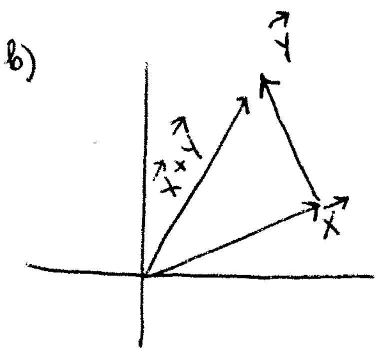
Cauchy - Schwarz states that

$\vec{a}^T\vec{b} \leq \|a\| \, \|b\|$ so

$$2\sqrt{xy} \leq x+y \Rightarrow \sqrt{xy} \leq \frac{x+y}{2}$$

The geometric mean is always less than or equal to the arithmetic mean!

b)



$$\|\vec{x}+\vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\| \quad \text{(Triangle inequality)}$$

we square both sides

$$\|\vec{x}+\vec{y}\|^2 \leq (\|\vec{x}\| + \|\vec{y}\|)^2$$

$$(\vec{x}+\vec{y})^T(\vec{x}+\vec{y}) \leq \|\vec{x}\|^2 + \|\vec{y}\|^2 + 2\|\vec{x}\|\,\|\vec{y}\|$$

$$\cancel{\vec{x}^T\vec{x}} + \cancel{2}\vec{y}^T\vec{x} + \cancel{\vec{y}^T\vec{y}} \leq \cancel{\vec{x}^T\vec{x}} + \cancel{\vec{y}^T\vec{y}} + \cancel{2}\|\vec{x}\|\,\|\vec{y}\|$$

$$\vec{y}^T\vec{x} \leq \|\vec{x}\|\,\|\vec{y}\| \quad \text{(Cauchy-Schwarz)}$$

For vectors in $\mathbb{R}^n$, inequalities are equivalent!

#3. First we project $\vec{b}$ onto $\vec{a}$

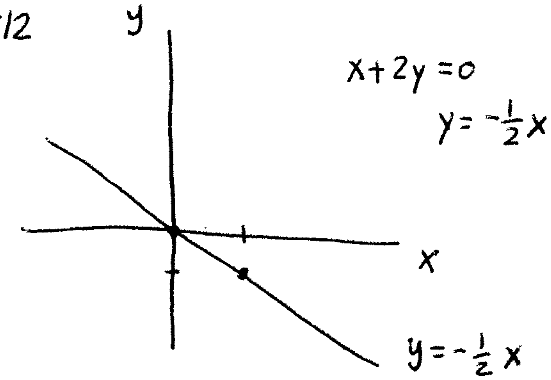
$$\vec{a} = (1,1,1) \quad \vec{b} = (2,4,4)$$

$$\frac{\vec{b}^T\vec{a}}{\vec{a}^T\vec{a}}\vec{a} = \frac{10}{3}\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}10/3\\10/3\\10/3\end{bmatrix}$$

Then we project $\vec{a}$ onto $\vec{b}$

$$\frac{\vec{a}^T\vec{b}}{\vec{b}^T\vec{b}}\vec{b} = \frac{10}{36}\begin{bmatrix}2\\4\\4\end{bmatrix} = \begin{bmatrix}5/9\\10/9\\10/9\end{bmatrix}$$

• #12



We need a matrix that projects a point (a,b) onto the line containing $(x, -\frac{1}{2}x)$ for any x.

We pick $x = 1$ to simplify but it doesn't matter what x we pick!

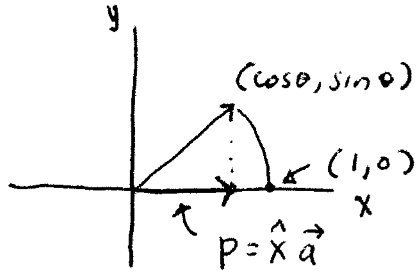
$$P = \frac{\vec{a}\vec{a}^T}{\vec{a}^T\vec{a}} = \frac{1}{\frac{5}{4}x^2}\begin{bmatrix} x^2 & -\frac{1}{2}x^2 \\ -\frac{1}{2}x^2 & \frac{1}{4}x^2 \end{bmatrix}$$

$$= \begin{bmatrix} 4/5 & -\frac{2}{5} \\ -2/5 & 1/5 \end{bmatrix}$$

#18 Draw the projection of $\vec{b}$ onto $\vec{a}$ and compute $p = \hat{x}\vec{a}$:

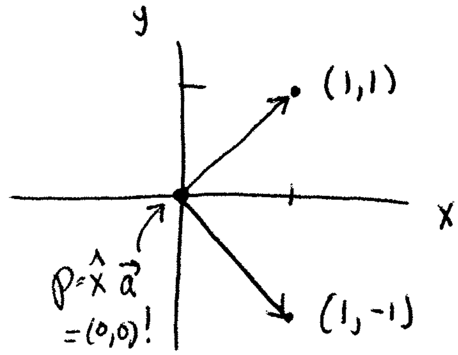
a) $\vec{b} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$ $\vec{a} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

$$p = \frac{\vec{b}^T\vec{a}}{\vec{a}^T\vec{a}}\vec{a} = \cos\theta\begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$



b) $\vec{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\vec{a} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

$$p = \frac{\vec{b}^T\vec{a}}{\vec{a}^T\vec{a}}\vec{a} = \frac{0}{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$



They are orthogonal!

Section 3.3

#1 Clearly there is no solution to $3x=10$ and $4x=5$. Least squares is the best we can do. We first write the two equations in vector form:

$$\begin{bmatrix}3\\4\end{bmatrix} x = \begin{bmatrix}10\\5\end{bmatrix} \qquad \hat{x} = (A^T A)^{-1} A^T \vec{b}$$

$$A \quad x = \vec{b} \qquad A^T A = 25$$

$$(A^T A)^{-1} = \frac{1}{25} \qquad A^T \vec{b} = 30 + 20 = 50$$

$$\hat{x} = \frac{50}{25} = 2.$$

The error that is minimized is $(3x-10)^2 + (4x-5)^2 = E^2$.
Note for $x=2$, $E^2 = (-4)^2 + (+3)^2 = 25$.

But imagine you didn't know about least-squares and solved each equation separately:

$$x = \frac{10}{3} \quad \text{and} \quad x = 5/4$$

For $x = \frac{10}{3}$, $\|\vec{E}\|^2 = 0^2 + \left(\frac{40}{3} - \frac{15}{3}\right)^2 = \left(\frac{25}{3}\right)^2$ which is more than 25.

For $x = 5/4$, $\|\vec{E}\|^2 = \left(\frac{15}{4} - \frac{40}{4}\right)^2 + 0^2 = \left(\frac{25}{4}\right)^2$ which is more than 25.

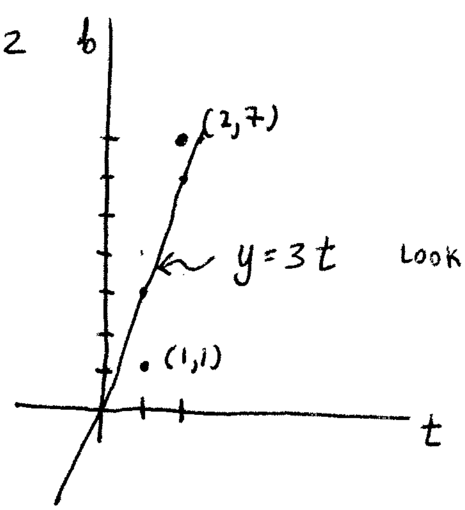
or $\bar{x} = \frac{1}{2}\left(\frac{10}{3} + 5/4\right) \approx 2.3$, $\|\vec{E}\|^2 \approx 27.1$ which is close but still more than 25.

We didn't have to do all this because we know that $\hat{x}$ minimizes E^2 but it is nice to see that the computation matches our intuition.

We check that $\vec{E}$ is perpendicular to the column space of A:

$\vec{E} = (-4, +3)$ $\qquad \vec{E}^T \begin{bmatrix}3\\4\end{bmatrix} = -12 + 12 = 0$. Yep, they are perpendicular.

• #2



$$\begin{bmatrix}1\\2\end{bmatrix} D = \begin{bmatrix}1\\7\end{bmatrix} \qquad \hat{D} = \frac{1}{5}(1+14) = 3$$

Looks reasonable.

#3.

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \qquad \hat{x} = \left(A^T A\right)^{-1} A^T \vec{b}$$

$$A^T A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \qquad \left(A^T A\right)^{-1} = \frac{1}{3} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

$$A^T \vec{b} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \text{so } \hat{x} = \frac{1}{3} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

The error $\vec{b} - A\hat{x} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2/3 \\ 2/3 \\ -2/3 \end{bmatrix}$

is perpendicular to the columns of A

$$\begin{bmatrix} 1 & 0 & 1 \end{bmatrix} \frac{1}{3} \begin{bmatrix} 2 \\ 2 \\ -2 \end{bmatrix} = 0 \quad \text{and} \quad \begin{bmatrix} 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2/3 \\ 2/3 \\ -2/3 \end{bmatrix} = 0.$$

#4 $\quad E^2 = (u-1)^2 + (v-3)^2 + (u+v-4)^2$

$E^2(u,v)$ is a function of two variables, to find its minimum we set its partial derivatives to zero:

$$\frac{\partial}{\partial u} E^2 = 2(u-1) + 0 + 2(u+v-4) = 0 \Rightarrow 2u - 2 + 2u + 2v - 8 = 0$$

$$4u + 2v = 10$$

$$2u + v = 5$$

and

$$\frac{\partial}{\partial v} E^2 = 0 + 2(v-3) + 2(u+v-4) = 0 \Rightarrow 2v - 6 + 2u + 2v - 8 = 0$$

$$2u + 4v = 14$$

$$u + 2v = 7$$

If we write these equations in matrix form:

$$\begin{matrix} 2u + v = 5 \\ u + 2v = 7 \end{matrix} \Rightarrow \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$$

$A^T A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \quad A^T \vec{b} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$ Hence $A^T A \begin{bmatrix} u \\ v \end{bmatrix} = A^T \vec{b}$ is $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$

the same as with Calculus!

#4 continued

$$\hat{x} = \frac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 5 \\ 7 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

$$\vec{p} = A\hat{x} = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix} \; !$$

$\vec{p} = \vec{b}$ because $\vec{b}$ is in the column space of A and in that case, the least squares solution is the exact solution!

#8 First let's think of two simple examples

$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ is a projection matrix onto the x-axis.

The column space is the x-axis and the rank is one.

$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ is a projection matrix onto the x-y plane.

The column space is the x-y plane and the rank is two.

O.K., now we answer the question:

If P is the projection matrix onto a k-dimensional subspace S, then the column space is S and the rank is k like in the previous two examples.

• #14 let $\vec{a}_1 = (1,1,0)$ and $\vec{a}_2 = (1,1,1)$ span a plane in $\mathbb{R}^3$. The simplest way to find a projection matrix into that plane is to make them columns of a matrix and use our formula for the projection into the column space of that matrix:

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \qquad P = A(A^TA)^{-1}A^T$$

$$A^TA = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 3 \end{bmatrix} \qquad (A^TA)^{-1} = \frac{1}{2}\begin{bmatrix} 3 & -2 \\ -2 & 2 \end{bmatrix}$$

$$P = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 3/2 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

To find a vector that gets projected to zero, we just need to find an element of the nullspace:

$$P\vec{x} = \vec{0} \quad \text{such as} \quad \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}.$$