18.06 Problem Set 5

Due Wednesday, Oct. 18, 2006 at **4:00 p.m.** in 2-106

Problem 1 *Wednesday 10/11*

For each of these, find a matrix satisfying the conditions given or explain why none can exist.

(a) Column space contains $\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$, and nullspace contains $\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$

(b) Row space contains $\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$, and nullspace contains $\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$

(c) $Ax = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ is solvable; $A^{\top}\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$ is zero.

(d) A *nonzero* matrix where every row is perpendicular to every column

(e) Rows sum to a row of zeros, and columns sum to $\begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix}$

Solution 1

(a) $\begin{bmatrix} -1 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -5 \end{bmatrix}$

(b) Anything with $\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$ in $\mathbf{N}(A)$ (and no other basis vectors!) automatically gives us the correct row space: $\begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix}$, for example.

(c) Can't do this: the column space and the left nullspace have to be orthogonal, but the vectors we're given from each have dot product $\begin{bmatrix} 1 & 0 & -1 \end{bmatrix}\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} = 3 + 0 - 1 = 2 \neq 0$.

(d) I'll give you two examples: $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$.

(e) Not possible (for any row length): $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ is in the left nullspace, so the column sum (which is in the column space) would have to be orthogonal to this.

Problem 2 *Wednesday 10/11*

Do Problem #12 from section 4.1 in your book.

Solution 2

See figure.

Problem 3 *Wednesday 10/11*

Do Problem #26 from section 4.1 in your book.

Solution 3

I used $A = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$; then $A^{\top}A$ is $S = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{bmatrix}$. All the off-diagonal entries $s_{ij} = A_i^{\top} A_j$ are zero, because the columns A_i and A_j are perpendicular.

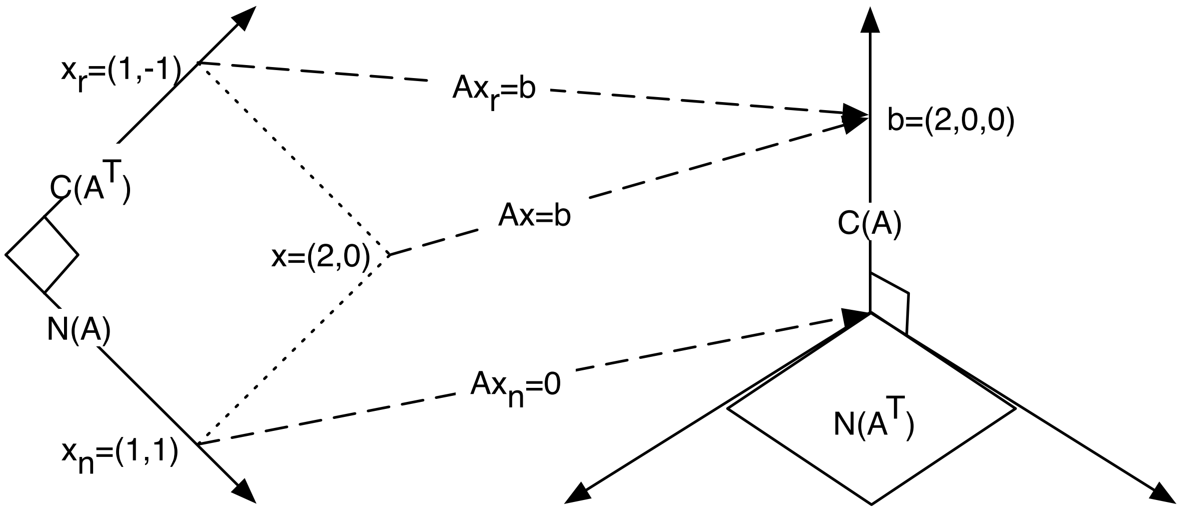
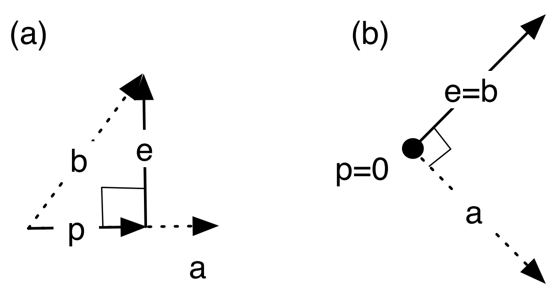


Figure 1: Solution to problem #2.

Problem 4 *Friday 10/13*

Do Problem #2 from section 4.2 in your book. What is the permutation matrix P? What is the error $e = b - p$?

Solution 4



For (a), $p = \begin{bmatrix} \cos\theta \\ 0 \end{bmatrix}$, $P = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, and $e = \begin{bmatrix} 0 \\ \sin\theta \end{bmatrix}$.
For (b), $p = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $P = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$, and $e = b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Problem 5 *Friday 10/13*

Do Problem #13 from section 4.2 in your book. Do this two different ways:
(a) geometrically, tell what subspace we're projecting b orthogonally onto
(b) algebraically, calculate $P = A(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}$

Solution 5

(a) Our subspace is just the hyperplane[1] $x_4 = 0$, so we project b straight down onto it to get

[1] A 1-dimensional subspace is a "line", a 2-dimensional subspace is a "plane", and an $(n-1)$-dimensional subspace is a "hyperplane". I think an $(n-2)$-dimensional subspace is a "hyperline".

$p = Pb = (1, 2, 3, 0)$.

(b) Here $A^\top A$ is the 3-by-3 identity matrix, so $P = AA^\top = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 0 \end{bmatrix}$ and $p = Pb = (1, 2, 3, 0)$.

Problem 6 *Friday 10/13*

A subspace **S** has basis $\left\{ a = \begin{bmatrix} 2 \\ -1 \\ 0 \\ 1 \end{bmatrix}, b = \begin{bmatrix} 5 \\ 0 \\ -1 \\ 2 \end{bmatrix}, c = \begin{bmatrix} 0 \\ 5 \\ 4 \\ -1 \end{bmatrix} \right\}$.

(a) What are the dot products $a^\top b$, $a^\top c$, $b^\top c$? Are the basis vectors orthogonal?
Now let's compute a new basis $\left\{ \hat{a}, \hat{b}, \hat{c} \right\}$ for the same subspace. Start by letting $\hat{a} = a$.
(b) Compute the projection Pb of b onto the line described by a. What is the error $(b - Pb)$? Call this error vector $\hat{b}$.
(c) Compute the projection $P_1 c$ of c onto the plane described by a and b. What is the error $(c - P_1 c)$? Call this error vector $\hat{c}$. Does $\hat{c}$ change if we project onto the plane with basis $\hat{a}$ and $\hat{b}$ instead? Why or why not?
(d) What are the dot products $\hat{a}^\top \hat{b}$, $\hat{a}^\top \hat{c}$, $\hat{b}^\top \hat{c}$? Are the new basis vectors orthogonal?
(This process for finding an orthogonal basis is called the "Gram-Schmidt Process" — the full version also scales each vector to "normalize" it to unit length.)
(e) Find the matrix R relating the old basis and the new basis: $\begin{bmatrix} a & b & c \end{bmatrix} = \begin{bmatrix} \hat{a} & \hat{b} & \hat{c} \end{bmatrix} \begin{bmatrix} ? & ? & ? \\ ? & ? & ? \\ ? & ? & ? \end{bmatrix}$
(f) Explain how you know $\left\{ \hat{a}, \hat{b}, \hat{c} \right\}$ is a basis for **S**. (Don't forget to show it both spans the subspace, and is linearly independent!)

Solution 6

(a) No, they're not orthogonal: $a^\top b = 12$, $a^\top c = -6$, $b^\top c = -6$.
(b) $Pb = \begin{bmatrix} 4 \\ -2 \\ 0 \\ 2 \end{bmatrix}$, so $\hat{b} = \begin{bmatrix} 1 \\ 2 \\ -1 \\ 0 \end{bmatrix}$.
(c) $P_1 c = \begin{bmatrix} -1 \\ 3 \\ -1 \\ -1 \end{bmatrix}$, so $\hat{c} = \begin{bmatrix} 1 \\ 2 \\ 5 \\ 0 \end{bmatrix}$.

The subspace we're projecting onto is the same, whether we use $\{a, b\}$ or $\left\{ \hat{a}, \hat{b} \right\}$ as the basis for it, so $\hat{c}$ is the same either way.
(d) All these dot products are zero, so the new basis vectors are orthogonal.
(e) $\begin{bmatrix} 2 & 5 & 0 \\ -1 & 0 & 5 \\ 0 & -1 & 4 \\ 1 & 2 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 \\ -1 & 2 & 2 \\ 0 & -1 & 5 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$. (Notice R is upper-triangular!)
(f) All the old basis vectors are linear combinations of the new ones, $(A = \hat{A}R)$ and all the new basis vectors are linear combinations of the old ones ($\hat{A} = AR^{-1}$, since R is invertible!) So they span the same subspace, and they have the same dimension.

Problem 7 *Monday 10/16*

Do Problem #17 from section 4.3 in your book.

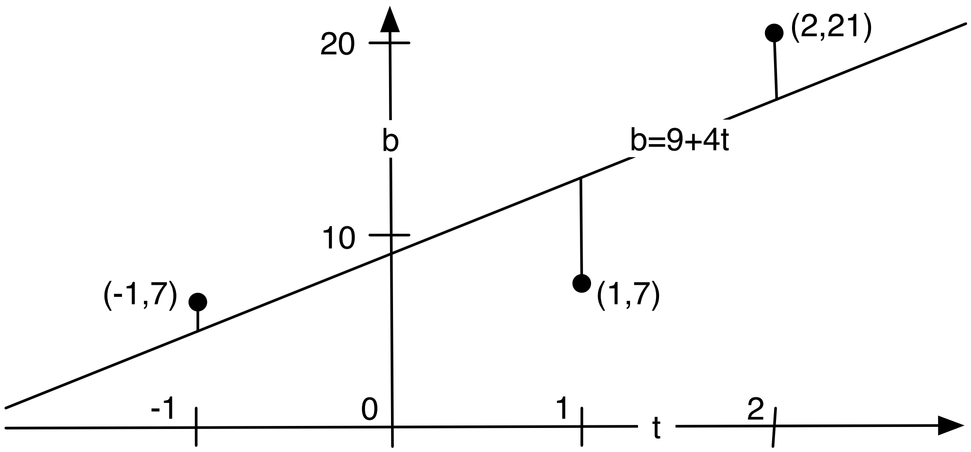


Figure 2: Graph of line for problem #7. The three vertical segments are the components of our error vector e, whose length we minimize with least-squares.

Solution 7

Plugging in $(t,b) = (-1,7),(1,7),(2,21)$ gives us the three equations $\begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 7 \\ 7 \\ 21 \end{bmatrix}$.

There are lots of solutions to this system $A\hat{x} = b$; for the *least squares* solution, I multiply by A^{T}:
$A^{\mathsf{T}}A\hat{x} = A^{\mathsf{T}}b \qquad \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix}\begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 35 \\ 42 \end{bmatrix}$

The unique solution to this system is $(C,D) = (9,4)$ and our line is $b = 9 + 4t$.

(We can write the solution to $A^{\mathsf{T}}A\hat{x} = A^{\mathsf{T}}b$ (the least-squares solution to $A\hat{x} = b$) as $\hat{x} = [(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}]b$. That thing in brackets is sometimes called the "pseudoinverse" A^+. It works almost like a regular inverse:

When A is invertible, the solution to $Ax = b$ is $x = A^{-1}b$.

When A isn't invertible,[2] a least-squares *solution to $A\hat{x} = b$ is $\hat{x} = A^+b$.)*

Problem 8 *Monday 10/16*

Do Problem #27 from section 4.3 in your book.

Solution 8

The equation $Ax = b$ we'd like to "solve" approximately, by least-squares: $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}\begin{bmatrix} C \\ D \\ E \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 3 \\ 4 \end{bmatrix}$

To "solve" it, multiply by A^{T} to get the 'normal equation' $A^{\mathsf{T}}A\hat{x} = A^{\mathsf{T}}b$: $\begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}\begin{bmatrix} C \\ D \\ E \end{bmatrix} = \begin{bmatrix} 8 \\ -3 \\ -3 \end{bmatrix}$

(Notice how $A^{\mathsf{T}}A$ is diagonal — that's because A's columns were orthogonal. Orthogonality makes life easier!)

This has the solution $C = 2$, $D = E = -3/2$, so the equation of the plane of 'best fit' is $2 - \frac{3}{2}x - \frac{3}{2}y = b$. At $(x,y) = (0,0)$ this is just $C = 2$, the average of the b-values 0,1,3,4.

[2] If $(A^{\mathsf{T}}A)^{-1}$ doesn't exist, the formula above won't work and we have to define A^+ a different way.