INTRODUCTION

TO

LINEAR

ALGEBRA

Fifth Edition

MANUAL FOR INSTRUCTORS

Gilbert Strang

Massachusetts Institute of Technology

math.mit.edu/linearalgebra

web.mit.edu/18.06

video lectures: ocw.mit.edu

math.mit.edu/∼gs

www.wellesleycambridge.com

email: linearalgebrabook@gmail.com

Wellesley - Cambridge Press

Box 812060

Wellesley, Massachusetts 02482

Problem Set 4.1, page 202

1 Both nullspace vectors will be orthogonal to the row space vector in $\mathbf{R}^3$. The column space of A and the nullspace of A^{T} are perpendicular lines in $\mathbf{R}^2$ because rank $= 1$.

2 The nullspace of a 3 by 2 matrix with rank 2 is $\mathbf{Z}$ (only the zero vector because the 2 columns are independent). So $\boldsymbol{x}_n = \mathbf{0}$, and row space $= \mathbf{R}^2$. Column space $=$ plane perpendicular to left nullspace $=$ line in $\mathbf{R}^3$ (because the rank is 2).

3 (a) One way is to use these two columns directly : $A = \begin{bmatrix} 1 & 2 & -3 \\ 2 & -3 & 1 \\ -3 & 5 & -2 \end{bmatrix}$

(b) Impossible because $\boldsymbol{N}(A)$ and $\boldsymbol{C}(A^{\mathrm{T}})$ are orthogonal subspaces : $\begin{bmatrix} 2 \\ -3 \\ 5 \end{bmatrix}$ is not orthogonal to $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

(c) $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ in $\boldsymbol{C}(A)$ and $\boldsymbol{N}(A^{\mathrm{T}})$ is impossible: not perpendicular

(d) Rows orthogonal to columns makes A times $A =$ zero matrix ρ. An example is $A = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$

(e) $(1, 1, 1)$ in the nullspace (columns add to the zero vector) and also $(1, 1, 1)$ is in the row space: no such matrix.

4 If $AB = 0$, the columns of B are in the *nullspace* of A and the rows of A are in the *left nullspace* of B. If rank $= 2$, all those four subspaces have dimension at least 2 which is impossible for 3 by 3.

5 (a) If $A\boldsymbol{x} = \boldsymbol{b}$ has a solution and $A^{\mathrm{T}}\boldsymbol{y} = \mathbf{0}$, then $\boldsymbol{y}$ is perpendicular to $\boldsymbol{b}$. $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{y} = (A\boldsymbol{x})^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{x}^{\mathrm{T}}(A^{\mathrm{T}}\boldsymbol{y}) = 0$. This says again that $\boldsymbol{C}(A)$ is orthogonal to $\boldsymbol{N}(A^{\mathrm{T}})$.
(b) If $A^{\mathrm{T}}\boldsymbol{y} = (1, 1, 1)$ has a solution, $(1, 1, 1)$ is a combination of the rows of A. It is in the **row space** and is orthogonal to every $\boldsymbol{x}$ in the **nullspace**.

6 Multiply the equations by $y_1, y_2, y_3 = 1, 1, -1$. Now the equations add to $0 = 1$ so there is no solution. In subspace language, $\boldsymbol{y} = (1, 1, -1)$ is in the left nullspace. $A\boldsymbol{x} = \boldsymbol{b}$ would need $0 = (\boldsymbol{y}^{\mathrm{T}}A)\boldsymbol{x} = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{b} = 1$ but here $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{b} = 1$.

7 Multiply the 3 equations by $\boldsymbol{y} = (1, 1, -1)$. Then $x_1 - x_2 = 1$ plus $x_2 - x_3 = 1$ *minus* $x_1 - x_3 = 1$ is $0 = 1$. Key point: This $\boldsymbol{y}$ in $\boldsymbol{N}(A^{\mathrm{T}})$ is not orthogonal to $\boldsymbol{b} = (1, 1, 1)$ so $\boldsymbol{b}$ is not in the column space and $A\boldsymbol{x} = \boldsymbol{b}$ has *no solution*.

8 Figure 4.3 has $\boldsymbol{x} = \boldsymbol{x}_r + \boldsymbol{x}_n$, where $\boldsymbol{x}_r$ is in the row space and $\boldsymbol{x}_n$ is in the nullspace. Then $A\boldsymbol{x}_n = \mathbf{0}$ and $A\boldsymbol{x} = A\boldsymbol{x}_r + A\boldsymbol{x}_n = A\boldsymbol{x}_r$. The example has $\boldsymbol{x} = (1, 0)$ and row space = line through $(1, 1)$ so the splitting is $\boldsymbol{x} = \boldsymbol{x}_r + \boldsymbol{x}_n = \left(\frac{1}{2}, \frac{1}{2}\right) + \left(\frac{1}{2}, -\frac{1}{2}\right)$. All $A\boldsymbol{x}$ are in $\boldsymbol{C}(A)$.

9 $A\boldsymbol{x}$ is always in the *column space* of A. If $A^{\mathrm{T}}A\boldsymbol{x} = \mathbf{0}$ then $A\boldsymbol{x}$ is also in the *nullspace* of A^{T}. Those subspaces are perpendicular. So $A\boldsymbol{x}$ is perpendicular to itself. Conclusion: $A\boldsymbol{x} = \mathbf{0}$ if $A^{\mathrm{T}}A\boldsymbol{x} = \mathbf{0}$.

10 (a) With $A^{\mathrm{T}} = A$, the column and row spaces are the *same*. The nullspace is always perpendicular to the row space. (b) $\boldsymbol{x}$ is in the nullspace and $\boldsymbol{z}$ is in the column space = row space: so these "eigenvectors" $\boldsymbol{x}$ and $\boldsymbol{z}$ have $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{z} = 0$.

11 ***For A***: The nullspace is spanned by $(-2, 1)$, the row space is spanned by $(1, 2)$. The column space is the line through $(1, 3)$ and $\boldsymbol{N}(A^{\mathrm{T}})$ is the perpendicular line through $(3, -1)$. ***For B***: The nullspace of B is spanned by $(0, 1)$, the row space is spanned by $(1, 0)$. The column space and left nullspace are the same as for A.

12 $\boldsymbol{x} = (2, 0)$ splits into $\boldsymbol{x}_r + \boldsymbol{x}_n = (1, -1) + (1, 1)$. Notice $\boldsymbol{N}(A^{\mathrm{T}})$ is the $y - z$ plane.

13 $V^{\mathrm{T}}W =$ zero matrix makes each column of V orthogonal to each column of W. This means: each basis vector for $\boldsymbol{V}$ is orthogonal to each basis vector for $\boldsymbol{W}$. Then *every* $\boldsymbol{v}$ in $\boldsymbol{V}$ (combinations of the basis vectors) is orthogonal to *every* $\boldsymbol{w}$ in $\boldsymbol{W}$.

14 $A\boldsymbol{x} = B\widehat{\boldsymbol{x}}$ means that $[\,A \;\; B\,]\begin{bmatrix} \boldsymbol{x} \\ -\widehat{\boldsymbol{x}} \end{bmatrix} = \mathbf{0}$. Three homogeneous equations (zero right hand sides) in four unknowns always have a nonzero solution. Here $\boldsymbol{x} = (3, 1)$ and

$\widehat{\boldsymbol{x}} = (1, 0)$ and $A\boldsymbol{x} = B\widehat{\boldsymbol{x}} = (5, 6, 5)$ is in both column spaces. Two planes in $\mathbf{R}^3$ must share a line.

15 A p-dimensional and a q-dimensional subspace of $\mathbf{R}^n$ share at least a line if $\boldsymbol{p} + \boldsymbol{q} > \boldsymbol{n}$. (The $p + q$ basis vectors of $\boldsymbol{V}$ and $\boldsymbol{W}$ cannot be independent, so same combination of the basis vectors of $\boldsymbol{V}$ is also a combination of the basis vectors of $\boldsymbol{W}$.)

16 $A^{\mathrm{T}}\boldsymbol{y} = \mathbf{0}$ leads to $(A\boldsymbol{x})^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}\boldsymbol{y} = 0$. Then $\boldsymbol{y} \perp A\boldsymbol{x}$ and $\boldsymbol{N}(A^{\mathrm{T}}) \perp \boldsymbol{C}(A)$.

17 If $\boldsymbol{S}$ is the subspace of $\mathbf{R}^3$ containing only the zero vector, then $\boldsymbol{S}^\perp$ is all of $\mathbf{R}^3$. If $\boldsymbol{S}$ is spanned by $(1, 1, 1)$, then $\boldsymbol{S}^\perp$ is the plane spanned by $(1, -1, 0)$ and $(1, 0, -1)$. If $\boldsymbol{S}$ is spanned by $(1, 1, 1)$ and $(1, 1, -1)$, then $\boldsymbol{S}^\perp$ is the line spanned by $(1, -1, 0)$.

18 $\boldsymbol{S}^\perp$ contains all vectors perpendicular to those two given vectors. So $\boldsymbol{S}^\perp$ is the nullspace of $A = \begin{bmatrix} 1 & 5 & 1 \\ 2 & 2 & 2 \end{bmatrix}$. Therefore $\boldsymbol{S}^\perp$ is a *subspace* even if $\boldsymbol{S}$ is not.

19 $\boldsymbol{L}^\perp$ is the 2-*dimensional subspace* (*a plane*) in $\mathbf{R}^3$ perpendicular to $\boldsymbol{L}$. Then $(\boldsymbol{L}^\perp)^\perp$ is a 1-*dimensional subspace* (*a line*) perpendicular to $\boldsymbol{L}^\perp$. In fact $(\boldsymbol{L}^\perp)^\perp$ is $\boldsymbol{L}$.

20 If $\boldsymbol{V}$ is the whole space $\mathbf{R}^4$, then $\boldsymbol{V}^\perp$ contains only the *zero vector*. Then $(\boldsymbol{V}^\perp)^\perp =$ all vectors perpendicular to the zero vector $= \mathbf{R}^4 = \boldsymbol{V}$.

21 For example $(-5, 0, 1, 1)$ and $(0, 1, -1, 0)$ span $\boldsymbol{S}^\perp =$nullspace of $A = \begin{bmatrix} 1 & 2 & 2 & 3 \\ 1 & 3 & 3 & 2 \end{bmatrix}$.

22 $(1, 1, 1, 1)$ is a basis for the line $\boldsymbol{P}^\perp$ orthogonal to P. $A = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$ has $\boldsymbol{P}$ as its nullspace and $\boldsymbol{P}^\perp$ as its row space.

23 $\boldsymbol{x}$ in $\boldsymbol{V}^\perp$ is perpendicular to every vector in $\boldsymbol{V}$. Since $\boldsymbol{V}$ contains all the vectors in $\boldsymbol{S}$, $\boldsymbol{x}$ is perpendicular to every vector in $\boldsymbol{S}$. So every $\boldsymbol{x}$ in $\boldsymbol{V}^\perp$ is also in $\boldsymbol{S}^\perp$.

24 $AA^{-1} = I$: Column 1 of A^{-1} is orthogonal to rows $2, 3, \ldots, n$ and therefore to the space spanned by those rows.

25 If the columns of A are unit vectors, all mutually perpendicular, then $A^{\mathrm{T}}A = I$. Simple but important! We write Q for such a matrix.

26 $A = \begin{bmatrix} 2 & 2 & -1 \\ -1 & 2 & 2 \\ 2 & -1 & 2 \end{bmatrix}$, This example shows a matrix with perpendicular columns. $A^{\mathrm{T}}A = 9I$ is *diagonal*: $(A^{\mathrm{T}}A)_{ij} =$ (column i of A) $\cdot$ (column j of A). When the columns are *unit vectors*, then $A^{\mathrm{T}}A = I$.

27 The lines $3x + y = b_1$ and $6x + 2y = b_2$ are **parallel**. They are the same line if $b_2 = 2b_1$. In that case (b_1, b_2) is perpendicular to $\mathbf{(-2, 1)}$. The nullspace of the 2 by 2 matrix is the line $3x + y = \mathbf{0}$. One particular vector in the nullspace is $\mathbf{(-1, 3)}$.

28 (a) $(1, -1, 0)$ is in both planes. Normal vectors are perpendicular, but planes still intersect! Two planes in $\mathbf{R}^3$ can't be orthogonal. (b) Need *three* orthogonal vectors to span the whole orthogonal complement in $\mathbf{R}^5$. (c) Lines in $\mathbf{R}^3$ can meet at the zero vector without being orthogonal.

29 $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 1 & -1 \\ 2 & -1 & 0 \\ 3 & 0 & -1 \end{bmatrix}$; A has $\boldsymbol{v} = (1, 2, 3)$ in row and column spaces; B has $\boldsymbol{v}$ in its column space and nullspace. $\boldsymbol{v}$ **can not** be in the nullspace and row space, or in the left nullspace and column space. These spaces are orthogonal and $\boldsymbol{v}^{\mathrm{T}}\boldsymbol{v} \neq 0$.

30 When $AB = 0$, every column of B is multiplied by A to give zero. So the column space of B is contained in the nullspace of A. Therefore the dimension of $\boldsymbol{C}(B) \leq$ dimension of $\boldsymbol{N}(A)$. This means $\text{rank}(B) \leq 4 -$ $\text{rank}(A)$.

31 null(N') produces a basis for the *row space* of A (perpendicular to $\boldsymbol{N}(A)$).

32 We need $\boldsymbol{r}^{\mathrm{T}}\boldsymbol{n} = 0$ and $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{\ell} = 0$. All possible examples have the form $a\boldsymbol{c}\boldsymbol{r}^{\mathrm{T}}$ with $a \neq 0$.

33 Both $\boldsymbol{r}$'s must be orthogonal to both $\boldsymbol{n}$'s, both $\boldsymbol{c}$'s must be orthogonal to both $\boldsymbol{\ell}$'s, each pair ($\boldsymbol{r}$'s, $\boldsymbol{n}$'s, $\boldsymbol{c}$'s, and $\boldsymbol{\ell}$'s) must be independent. Fact: All A's with these subspaces have the form $[\boldsymbol{c}_1\ \boldsymbol{c}_2]M[\boldsymbol{r}_1\ \boldsymbol{r}_2]^{\mathrm{T}}$ for a 2 by 2 invertible M.

You must take $[\boldsymbol{c}_1, \boldsymbol{c}_2]$ times $[\boldsymbol{r}_1, \boldsymbol{r}_2]^{\mathrm{T}}$.

Problem Set 4.2, page 214

1 (a) $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}/\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a} = 5/3$; $\boldsymbol{p} = 5\boldsymbol{a}/3 = (5/3, 5/3, 5/3)$; $\boldsymbol{e} = (-2, 1, 1)/3$

(b) $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}/\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}=-1$; $\boldsymbol{p}=-\boldsymbol{a}$; $\boldsymbol{e}=\mathbf{0}$.

2 (a) The projection of $\boldsymbol{b} = (\cos\theta, \sin\theta)$ onto $\boldsymbol{a} = (1,0)$ is $\boldsymbol{p} = (\cos\theta, 0)$

(b) The projection of $\boldsymbol{b} = (1,1)$ onto $\boldsymbol{a} = (1,-1)$ is $\boldsymbol{p} = (0,0)$ since $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b} = 0$.

The picture for part (a) has the vector $\boldsymbol{b}$ at an angle θ with the horizontal $\boldsymbol{a}$. The picture for part (b) has vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ at a 90° angle.

3 $P_1 = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ and $P_1\boldsymbol{b} = \frac{1}{3}\begin{bmatrix} 5 \\ 5 \\ 5 \end{bmatrix}$. $P_2 = \frac{1}{11}\begin{bmatrix} 1 & 3 & 1 \\ 3 & 9 & 3 \\ 1 & 3 & 1 \end{bmatrix}$ and $P_2\boldsymbol{b} = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$.

4 $P_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $P_2 = \dfrac{\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}} = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$. P_1 projects onto $(1,0)$, P_2 projects onto $(1,-1)$. $P_1P_2 \neq 0$ and $P_1 + P_2$ is not a projection matrix. $(P_1 + P_2)^2$ is different from $P_1 + P_2$.

5 $P_1 = \frac{1}{9}\begin{bmatrix} 1 & -2 & -2 \\ -2 & 4 & 4 \\ -2 & 4 & 4 \end{bmatrix}$ and $P_2 = \frac{1}{9}\begin{bmatrix} 4 & 4 & -2 \\ 4 & 4 & -2 \\ -2 & -2 & 1 \end{bmatrix}$.

P_1 and P_2 are the projection matrices onto the lines through $\boldsymbol{a}_1 = (-1,2,2)$ and $\boldsymbol{a}_2 = (2,2,-1)$. $P_1P_2 =$ *zero matrix because* $\boldsymbol{a}_1 \perp \boldsymbol{a}_2$.

6 $\boldsymbol{p}_1 = (\frac{1}{9}, -\frac{2}{9}, -\frac{2}{9})$ and $\boldsymbol{p}_2 = (\frac{4}{9}, \frac{4}{9}, -\frac{2}{9})$ and $\boldsymbol{p}_3 = (\frac{4}{9}, -\frac{2}{9}, \frac{4}{9})$. So $\boldsymbol{p}_1 + \boldsymbol{p}_2 + \boldsymbol{p}_3 = \boldsymbol{b}$.

7 $$P_1 + P_2 + P_3 = \frac{1}{9}\begin{bmatrix} 1 & -2 & -2 \\ -2 & 4 & 4 \\ -2 & 4 & 4 \end{bmatrix} + \frac{1}{9}\begin{bmatrix} 4 & 4 & -2 \\ 4 & 4 & -2 \\ -2 & -2 & 1 \end{bmatrix} + \frac{1}{9}\begin{bmatrix} 4 & -2 & 4 \\ -2 & 1 & -2 \\ 4 & -2 & 4 \end{bmatrix} = I.$$

We *can* add projections onto *orthogonal vectors* to get the projection matrix onto the larger space. This is important.

8 The projections of $(1,1)$ onto $(1,0)$ and $(1,2)$ are $\boldsymbol{p}_1 = (1,0)$ and $\boldsymbol{p}_2 = \frac{3}{5}(1,2)$. Then $\boldsymbol{p}_1 + \boldsymbol{p}_2 \neq \boldsymbol{b}$. The sum of projections is not a projection onto the space spanned by $(1,0)$ and $(1,2)$ because those vectors are *not orthogonal*.

9 Since A is invertible, $P = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$ separates into $AA^{-1}(A^{\mathrm{T}})^{-1}A^{\mathrm{T}} = I$. And I is the projection matrix onto all of $\mathbf{R}^2$.

10 $P_2 = \dfrac{\boldsymbol{a}_2\boldsymbol{a}_2^{\mathrm{T}}}{\boldsymbol{a}_2^{\mathrm{T}}\boldsymbol{a}_2} = \begin{bmatrix} 0.2 & 0.4 \\ 0.4 & 0.8 \end{bmatrix}$, $P_2\boldsymbol{a}_1 = \begin{bmatrix} 0.2 \\ 0.4 \end{bmatrix}$, $P_1 = \dfrac{\boldsymbol{a}_1\boldsymbol{a}_1^{\mathrm{T}}}{\boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{a}_1} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $P_1P_2\boldsymbol{a}_1 = \begin{bmatrix} 0.2 \\ 0 \end{bmatrix}$. This is not $\boldsymbol{a}_1 = (1, 0)$

No, $\boldsymbol{P_1P_2} \neq (P_1P_2)^2$.

11 (a) $\boldsymbol{p} = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{b} = (2, 3, 0)$, $\boldsymbol{e} = (0, 0, 4)$, $A^{\mathrm{T}}\boldsymbol{e} = \boldsymbol{0}$

(b) $\boldsymbol{p} = (4, 4, 6)$ and $\boldsymbol{e} = \boldsymbol{0}$ because $\boldsymbol{b}$ is in the column space of A.

12 $P_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ = projection matrix onto the column space of A (the xy plane)

$P_2 = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.5 & 0.5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ = Projection matrix $A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$ onto the second column space. Certainly $(P_2)^2 = P_2$. A true projection matrix.

13 $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, P = square matrix = $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, $\boldsymbol{p} = P\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 0 \end{bmatrix}$.

14 The projection of this $\boldsymbol{b}$ onto the column space of A is $\boldsymbol{b}$ itself because $\boldsymbol{b}$ is in that column space. But P is not necessarily I. Here $\boldsymbol{b} = 2(\text{column } 1 \text{ of } A)$:

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 2 \\ 2 & 0 \end{bmatrix} \text{ gives } P = \frac{1}{21}\begin{bmatrix} 5 & 8 & -4 \\ 8 & 17 & 2 \\ -4 & 2 & 20 \end{bmatrix} \text{ and } \boldsymbol{b} = P\boldsymbol{b} = \boldsymbol{p} = \begin{bmatrix} 0 \\ 2 \\ 4 \end{bmatrix}.$$

15 $2A$ has the same column space as A. Then P is the same for A and $2A$, but $\widehat{\boldsymbol{x}}$ for $2A$ is *half* of $\widehat{\boldsymbol{x}}$ for A.

16 $\frac{1}{2}(1, 2, -1) + \frac{3}{2}(1, 0, 1) = (2, 1, 1)$. So $\boldsymbol{b}$ is in the plane. Projection shows $P\boldsymbol{b} = \boldsymbol{b}$.

17 If $P^2 = P$ then $\boldsymbol{(I - P)^2} = (I - P)(I - P) = I - PI - IP + P^2 = \boldsymbol{I - P}$. When P projects onto the column space, $I - P$ projects onto the *left nullspace*.

18 (a) $I - P$ is the projection matrix onto $(1, -1)$ in the perpendicular direction to $(1, 1)$

(b) $I - P$ projects onto the plane $x + y + z = 0$ perpendicular to $(1, 1, 1)$.

19 For any basis vectors in the plane $x - y - 2z = 0$, say $(1, 1, 0)$ and $(2, 0, 1)$, the matrix $P = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$ is $\begin{bmatrix} 5/6 & 1/6 & 1/3 \\ 1/6 & 5/6 & -1/3 \\ 1/3 & -1/3 & 1/3 \end{bmatrix}$.

20 $\boldsymbol{e} = \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}$, $Q = \frac{\boldsymbol{e}\boldsymbol{e}^{\mathrm{T}}}{\boldsymbol{e}^{\mathrm{T}}\boldsymbol{e}} = \begin{bmatrix} 1/6 & -1/6 & -1/3 \\ -1/6 & 1/6 & 1/3 \\ -1/3 & 1/3 & 2/3 \end{bmatrix}$, $I - Q = \begin{bmatrix} 5/6 & 1/6 & 1/3 \\ 1/6 & 5/6 & -1/3 \\ 1/3 & -1/3 & 1/3 \end{bmatrix}$.

21 $\left(A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\right)^2 = A(A^{\mathrm{T}}A)^{-1}(A^{\mathrm{T}}A)(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}} = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$. So $P^2 = P$. $P\boldsymbol{b}$ is in the column space (where P projects). Then its projection $P(P\boldsymbol{b})$ is also $P\boldsymbol{b}$.

22 $P^{\mathrm{T}} = (A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}})^{\mathrm{T}} = A((A^{\mathrm{T}}A)^{-1})^{\mathrm{T}}A^{\mathrm{T}} = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}} = P$. ($A^{\mathrm{T}}A$ is symmetric!)

23 If A is invertible then its column space is all of $\mathbf{R}^n$. So $P = I$ and $\boldsymbol{e} = \mathbf{0}$.

24 The nullspace of A^{T} is *orthogonal* to the column space $\boldsymbol{C}(A)$. So if $A^{\mathrm{T}}\boldsymbol{b} = \mathbf{0}$, the projection of $\boldsymbol{b}$ onto $\boldsymbol{C}(A)$ should be $\boldsymbol{p} = \mathbf{0}$. Check $P\boldsymbol{b} = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{b} = A(A^{\mathrm{T}}A)^{-1}\mathbf{0}$.

25 **The column space of P is the space that P projects onto**. The column space of A always contains all outputs $A\boldsymbol{x}$ and here the outputs $P\boldsymbol{x}$ fill the subspace S. Then rank of P = dimension of $S = n$.

26 A^{-1} exists since the rank is $r = m$. Multiply $A^2 = A$ by A^{-1} to get $A = I$.

27 If $A^{\mathrm{T}}A\boldsymbol{x} = \mathbf{0}$ then $A\boldsymbol{x}$ is in the **nullspace of A^{T}**. But $A\boldsymbol{x}$ is always in the **column space of A**. To be in both of those perpendicular spaces, $A\boldsymbol{x}$ must be zero. So A and $A^{\mathrm{T}}A$ have the *same nullspace*: $A^{\mathrm{T}}A\boldsymbol{x} = \mathbf{0}$ exactly when $A\boldsymbol{x} = \mathbf{0}$.

28 $P^2 = P = P^{\mathrm{T}}$ give $P^{\mathrm{T}}P = P$. Then the $(2, 2)$ entry of P equals the $(2, 2)$ entry of $P^{\mathrm{T}}P$. But the $(2, 2)$ entry of $P^{\mathrm{T}}P$ is the length squared of column 2.

29 $A = B^{\mathrm{T}}$ has independent columns, so $A^{\mathrm{T}}A$ (which is BB^{T}) must be invertible.

30 (a) The column space is the line through $\boldsymbol{a} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ so $P_C = \dfrac{\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}} = \dfrac{1}{25}\begin{bmatrix} 9 & 12 \\ 12 & 16 \end{bmatrix}$.

The formula $P = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$ needs independent columns—this A has dependent columns. The update formula is correct.

(b) The row space is the line through $\boldsymbol{v} = (1,2,2)$ and $P_R = \boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}/\boldsymbol{v}^{\mathrm{T}}\boldsymbol{v}$. Always $P_C A = A$ (columns of A project to themselves) and $AP_R = A$. Then $\boldsymbol{P_C A P_R = A}$.

31 *Test*: The error $\boldsymbol{e} = \boldsymbol{b} - \boldsymbol{p}$ must be perpendicular to all the $\boldsymbol{a}$'s.

32 Since $P_1\boldsymbol{b}$ is in $\boldsymbol{C}(A)$ and P_2 projects onto that column space, $P_2(P_1\boldsymbol{b})$ equals $P_1\boldsymbol{b}$. So $P_2P_1 = P_1 = \boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}/\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}$ where $\boldsymbol{a} = (1,2,0)$.

33 Each $\boldsymbol{b}_1$ to $\boldsymbol{b}_{99}$ is multiplied by $\frac{1}{999} - \frac{1}{1000}\left(\frac{1}{999}\right) = \frac{999}{1000}\frac{1}{999} = \frac{1}{1000}$. The last pages of the book discuss least squares and the Kalman filter.

Problem Set 4.3, page 229

1 $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}$ and $\boldsymbol{b} = \begin{bmatrix} 0 \\ 8 \\ 8 \\ 20 \end{bmatrix}$ give $A^{\mathrm{T}}A = \begin{bmatrix} 4 & 8 \\ 8 & 26 \end{bmatrix}$ and $A^{\mathrm{T}}\boldsymbol{b} = \begin{bmatrix} 36 \\ 112 \end{bmatrix}$.

$A^{\mathrm{T}}A\widehat{\boldsymbol{x}} = A^{\mathrm{T}}\boldsymbol{b}$ gives $\widehat{\boldsymbol{x}} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$ and $\boldsymbol{p} = A\widehat{\boldsymbol{x}} = \begin{bmatrix} 1 \\ 5 \\ 13 \\ 17 \end{bmatrix}$ and $\boldsymbol{e} = \boldsymbol{b} - \boldsymbol{p} = \begin{bmatrix} -1 \\ 3 \\ -5 \\ 3 \end{bmatrix}$

$E = \|\boldsymbol{e}\|^2 = \mathbf{44}$

2 $\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ 8 \\ 20 \end{bmatrix}$. This $A\boldsymbol{x} = \boldsymbol{b}$ is unsolvable. Project $\boldsymbol{b}$ to $\boldsymbol{p} = P\boldsymbol{b} = \begin{bmatrix} 1 \\ 5 \\ 13 \\ 17 \end{bmatrix}$; When $\boldsymbol{p}$ replaces $\boldsymbol{b}$,

$\widehat{\boldsymbol{x}} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$ exactly solves $A\widehat{\boldsymbol{x}} = \boldsymbol{p}$.

3 In Problem 2, $\boldsymbol{p} = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{b} = (1,5,13,17)$ and $\boldsymbol{e} = \boldsymbol{b} - \boldsymbol{p} = (-1,3,-5,3)$. This $\boldsymbol{e}$ is perpendicular to both columns of A. This shortest distance $\|\boldsymbol{e}\|$ is $\sqrt{44}$.

4 $E = (C + \mathbf{0}D)^2 + (C + \mathbf{1}D - 8)^2 + (C + \mathbf{3}D - 8)^2 + (C + \mathbf{4}D - 20)^2$. Then $\partial E/\partial C = 2C + 2(C + D - 8) + 2(C + 3D - 8) + 2(C + 4D - 20) = 0$ and $\partial E/\partial D = 1 \cdot 2(C + D - 8) + 3 \cdot 2(C + 3D - 8) + 4 \cdot 2(C + 4D - 20) = 0$.
These two normal equations are again $\begin{bmatrix} 4 & 8 \\ 8 & 26 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 36 \\ 112 \end{bmatrix}$.

5 $E = (C-0)^2 + (C-8)^2 + (C-8)^2 + (C-20)^2$. $A^{\mathrm{T}} = [\,1 \;\; 1 \;\; 1 \;\; 1\,]$ and $A^{\mathrm{T}}A = [\,4\,]$. $A^{\mathrm{T}}\boldsymbol{b} = [\,36\,]$ and $(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{b} = \mathbf{9}$ = best height C for the horizontal line. Errors $\boldsymbol{e} = \boldsymbol{b} - \boldsymbol{p} = (-9, -1, -1, 11)$ still add to zero.

6 $\boldsymbol{a} = (1, 1, 1, 1)$ and $\boldsymbol{b} = (0, 8, 8, 20)$ give $\widehat{x} = \boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}/\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a} = 9$ and the projection is $\widehat{x}\boldsymbol{a} = \boldsymbol{p} = (9, 9, 9, 9)$. Then $\boldsymbol{e}^{\mathrm{T}}\boldsymbol{a} = (-9, -1, -1, 11)^{\mathrm{T}}(1, 1, 1, 1) = 0$ and the shortest distance from $\boldsymbol{b}$ to the line through $\boldsymbol{a}$ is $\|\boldsymbol{e}\| = \sqrt{204}$.

7 Now the 4 by 1 matrix in $A\boldsymbol{x} = \boldsymbol{b}$ is $A = [\,0 \;\; 1 \;\; 3 \;\; 4\,]^{\mathrm{T}}$. Then $A^{\mathrm{T}}A = [\,26\,]$ and $A^{\mathrm{T}}\boldsymbol{b} = [\,112\,]$. Best $D = 112/26 = 56/13$.

8 $\widehat{\boldsymbol{x}} = \boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}/\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a} = 56/13$ and $\boldsymbol{p} = (56/13)(0, 1, 3, 4)$. $(C, D) = (9, 56/13)$ don't match $(C, D) = (1, 4)$ from Problems 1-4. Columns of A were not perpendicular so we can't project separately to find C and D.

9 Parabola Project $\boldsymbol{b}$ 4D to 3D $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{bmatrix} \begin{bmatrix} C \\ D \\ E \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ 8 \\ 20 \end{bmatrix}$. $A^{\mathrm{T}}A\widehat{\boldsymbol{x}} = \begin{bmatrix} 4 & 8 & 26 \\ 8 & 26 & 92 \\ 26 & 92 & 338 \end{bmatrix} \begin{bmatrix} C \\ D \\ E \end{bmatrix} = \begin{bmatrix} 36 \\ 112 \\ 400 \end{bmatrix}$.

Figure 4.9 (a) is fitting 4 points and 4.9 (b) is a projection in $\mathbf{R}^4$: same problem !

10 $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 3 & 9 & 27 \\ 1 & 4 & 16 & 64 \end{bmatrix} \begin{bmatrix} C \\ D \\ E \\ F \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ 8 \\ 20 \end{bmatrix}$. Then $\begin{bmatrix} C \\ D \\ E \\ F \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 0 \\ 47 \\ -28 \\ 5 \end{bmatrix}$. **Exact cubic so $\boldsymbol{p} = \boldsymbol{b}$, $\boldsymbol{e} = \mathbf{0}$.** This Vandermonde matrix gives exact interpolation by a cubic at $0, 1, 3, 4$

11 (a) The best line $x = 1 + 4t$ gives the center point $\widehat{\boldsymbol{b}} = 9$ at center time, $\widehat{t} = 2$.
(b) The first equation $Cm + D\sum t_i = \sum b_i$ divided by m gives $C + D\widehat{t} = \widehat{\boldsymbol{b}}$. This shows : The best line goes through $\widehat{\boldsymbol{b}}$ at time $\widehat{t}$.

12 (a) $\boldsymbol{a} = (1,\ldots,1)$ has $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a} = m$, $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b} = b_1 + \cdots + b_m$. Therefore $\widehat{x} = \boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}/m$ is the **mean** of the b's (their average value)

(b) $\boldsymbol{e} = \boldsymbol{b} - \widehat{x}\boldsymbol{a}$ and $\|\boldsymbol{e}\|^2 = (b_1 - \text{ mean })^2 + \cdots + (b_m - \text{ mean })^2 =$ **variance** (denoted by $\boldsymbol{\sigma^2}$).

(c) $\boldsymbol{p} = (3,3,3)$ and $\boldsymbol{e} = (-2,-1,3)$ $\boldsymbol{p}^{\mathrm{T}}\boldsymbol{e} = 0$. Projection matrix $P = \dfrac{1}{3}\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix}$.

13 $(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}(\boldsymbol{b} - A\boldsymbol{x}) = \widehat{\boldsymbol{x}} - \boldsymbol{x}$. This tells us: When the components of $A\boldsymbol{x} - \boldsymbol{b}$ add to zero, so do the components of $\widehat{\boldsymbol{x}} - \boldsymbol{x}$: Unbiased.

14 The matrix $(\widehat{\boldsymbol{x}} - \boldsymbol{x})(\widehat{\boldsymbol{x}} - \boldsymbol{x})^{\mathrm{T}}$ is $(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}(\boldsymbol{b} - A\boldsymbol{x})(\boldsymbol{b} - A\boldsymbol{x})^{\mathrm{T}}A(A^{\mathrm{T}}A)^{-1}$. When the average of $(\boldsymbol{b} - A\boldsymbol{x})(\boldsymbol{b} - A\boldsymbol{x})^{\mathrm{T}}$ is $\sigma^2 I$, the average of $(\widehat{\boldsymbol{x}} - \boldsymbol{x})(\widehat{\boldsymbol{x}} - \boldsymbol{x})^{\mathrm{T}}$ will be the *output covariance matrix* $(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\sigma^2 A(A^{\mathrm{T}}A)^{-1}$ which simplifies to $\sigma^2(A^{\mathrm{T}}A)^{-1}$. That gives the average of the squared output errors $\widehat{\boldsymbol{x}} - \boldsymbol{x}$.

15 When A has 1 column of 4 ones, Problem 14 gives the expected error $(\widehat{x} - x)^2$ as $\sigma^2(A^{\mathrm{T}}A)^{-1} = \sigma^2/4$. By taking m measurements, the variance drops from σ^2 to σ^2/m. This leads to the **Monte Carlo method** in Section 12.1.

16 $\dfrac{1}{10}b_{10} + \dfrac{\mathbf{9}}{\mathbf{10}}\widehat{x}_9 = \dfrac{1}{10}(b_1 + \cdots + b_{10})$. Knowing $\widehat{x}_9$ avoids adding all ten b's.

17 $\begin{bmatrix}1&-1\\1&1\\1&2\end{bmatrix}\begin{bmatrix}C\\D\end{bmatrix} = \begin{bmatrix}7\\7\\21\end{bmatrix}$. The solution $\widehat{\boldsymbol{x}} = \begin{bmatrix}\mathbf{9}\\\mathbf{4}\end{bmatrix}$ comes from $\begin{bmatrix}3&2\\2&6\end{bmatrix}\begin{bmatrix}C\\D\end{bmatrix} = \begin{bmatrix}35\\42\end{bmatrix}$.

18 $\boldsymbol{p} = A\widehat{\boldsymbol{x}} = (5,13,17)$ gives the heights of the closest line. The vertical errors are $\boldsymbol{b} - \boldsymbol{p} = (2,-6,4)$. This error $\boldsymbol{e}$ has $P\boldsymbol{e} = P\boldsymbol{b} - P\boldsymbol{p} = \boldsymbol{p} - \boldsymbol{p} = \mathbf{0}$.

19 If $\boldsymbol{b} =$ error $\boldsymbol{e}$ then $\boldsymbol{b}$ is perpendicular to the column space of A. Projection $\boldsymbol{p} = \mathbf{0}$.

20 The matrix A has columns $1,1,1$ and $-1,1,2$. If $\boldsymbol{b} = A\widehat{\boldsymbol{x}} = (5,13,17)$ then $\widehat{\boldsymbol{x}} = (9,4)$ and $\boldsymbol{e} = \mathbf{0}$ since $\boldsymbol{b} = 9\,(\text{column } 1) + 4\,(\text{column } 2)$ is *in the column space of* A.

21 $\boldsymbol{e}$ is in $\boldsymbol{N}(A^{\mathrm{T}})$; $\boldsymbol{p}$ is in $\boldsymbol{C}(A)$; $\widehat{\boldsymbol{x}}$ is in $\boldsymbol{C}(A^{\mathrm{T}})$; $\boldsymbol{N}(A) = \{\mathbf{0}\}$ = zero vector only.

22 The least squares equation is $\begin{bmatrix} 5 & \mathbf{0} \\ \mathbf{0} & 10 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 5 \\ -10 \end{bmatrix}$. Solution: $C = 1$, $D = -1$. The best line is $b = 1 - t$. Symmetric t's $\Rightarrow$ diagonal $A^{\mathrm{T}}A \Rightarrow$ easy solution.

23 $\boldsymbol{e}$ is orthogonal to $\boldsymbol{p}$ in $\mathbf{R}^m$; then $\|\boldsymbol{e}\|^2 = e^{\mathrm{T}}(\boldsymbol{b} - \boldsymbol{p}) = \boldsymbol{e}^{\mathrm{T}}\boldsymbol{b} = \boldsymbol{b}^{\mathrm{T}}\boldsymbol{b} - \boldsymbol{b}^{\mathrm{T}}\boldsymbol{p}$.

24 The derivatives of $\|A\boldsymbol{x} - \boldsymbol{b}\|^2 = \boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}A\boldsymbol{x} - 2\boldsymbol{b}^{\mathrm{T}}A\boldsymbol{x} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{b}$ (this last term is constant) are zero when $2A^{\mathrm{T}}A\boldsymbol{x} = 2A^{\mathrm{T}}\boldsymbol{b}$, or $\boldsymbol{x} = (A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{b}$.

25 3 points on a linewill give **equal slopes** $(b_2 - b_1)/(t_2 - t_1) = (b_3 - b_2)/(t_3 - t_2)$. Linear algebra: Orthogonal to the columns $(1, 1, 1)$ and (t_1, t_2, t_3) is $\boldsymbol{y} = (t_2 - t_3, t_3 - t_1, t_1 - t_2)$ in the left nullspace of A. $\boldsymbol{b}$ is in the column space ! Then $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{b} = 0$ is the same equal slopes condition written as $(b_2 - b_1)(t_3 - t_2) = (b_3 - b_2)(t_2 - t_1)$.

26 The unsolvable equations for $C + Dx + Ey = (0, 1, 3, 4)$ at the 4 corners are $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} C \\ D \\ E \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 3 \\ 4 \end{bmatrix}$. Then $A^{\mathrm{T}}A = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ and $A^{\mathrm{T}}\boldsymbol{b} = \begin{bmatrix} 8 \\ -2 \\ -3 \end{bmatrix}$ and $\begin{bmatrix} C \\ D \\ E \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ -3/2 \end{bmatrix}$. At $x, y = 0, 0$ the best plane $2 - x - \frac{3}{2}y$ has height $C = \mathbf{2}$ = average of $0, 1, 3, 4$.

27 The shortest link connecting two lines in space is *perpendicular to those lines*.

28 If A has dependent columns, then $A^{\mathrm{T}}A$ is not invertable and the usual formula $P = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$ will fail. Replace A in that formula by the matrix B that keeps *only the pivot columns of* A.

29 Only 1 plane contains $\mathbf{0}, \boldsymbol{a}_1, \boldsymbol{a}_2$ unless $\boldsymbol{a}_1, \boldsymbol{a}_2$ are *dependent*. Same test for $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{n-1}$. If they are dependent, there is a vector $\boldsymbol{v}$ perpendicular to all the $\boldsymbol{a}$'s. Then they all lie on the plane $\boldsymbol{v}^{\mathrm{T}}\boldsymbol{x} = 0$ going through $\boldsymbol{x} = (0, 0, \ldots, 0)$.

30 When A has orthogonal columns $(1,\ldots,1)$ and $(T_1,\ldots,T_m)$, the matrix $A^{\mathrm{T}}A$ is **diagonal** with entries m and $T_1^2+\cdots+T_m^2$. Also $A^{\mathrm{T}}b$ has entries $b_1+\cdots+b_m$ and $T_1b_1+\cdots+T_mb_m$. The solution with that diagonal $A^{\mathrm{T}}A$ is just the given $\widehat{\boldsymbol{x}}=(C,D)$.

Problem Set 4.4, page 242

1 (a) *Independent* (b) *Independent* and *orthogonal* (c) *Independent* and *orthonormal*. For orthonormal vectors, (a) becomes $(1,0)$, $(0,1)$ and (b) is $(.6,.8)$, $(.8,-.6)$.

2 Divide by length 3 to get $\boldsymbol{q}_1=(\frac{2}{3},\frac{2}{3},-\frac{1}{3})$. $\boldsymbol{q}_2=(-\frac{1}{3},\frac{2}{3},\frac{2}{3})$. $Q^{\mathrm{T}}Q=\begin{bmatrix}1&0\\0&1\end{bmatrix}$ but $QQ^{\mathrm{T}}=\begin{bmatrix}5/9&2/9&-4/9\\2/9&8/9&2/9\\-4/9&2/9&5/9\end{bmatrix}$.

3 (a) $A^{\mathrm{T}}A$ will be $16I$ (b) $A^{\mathrm{T}}A$ will be diagonal with entries $1^2,2^2,3^2=1,4,9$.

4 (a) $Q=\begin{bmatrix}1&0\\0&1\\0&0\end{bmatrix}$, $QQ^{\mathrm{T}}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}\neq I$. Any Q with $n<m$ has $QQ^{\mathrm{T}}\neq I$.

(b) $(1,0)$ and $(0,0)$ are *orthogonal*, not *independent*. Nonzero orthogonal vectors *are* independent. (c) From $\boldsymbol{q}_1=(1,1,1)/\sqrt{3}$ my favorite is $\boldsymbol{q}_2=(1,-1,0)/\sqrt{2}$ and $\boldsymbol{q}_3=(1,1,-2)/\sqrt{6}$.

5 *Orthogonal* vectors are $(1,-1,0)$ and $(1,1,-1)$. *Orthonormal* after dividing by their lengths: $\left(\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}},0\right)$ and $\left(\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},-\frac{1}{\sqrt{3}}\right)$.

6 Q_1Q_2 is orthogonal because $(Q_1Q_2)^{\mathrm{T}}Q_1Q_2=Q_2^{\mathrm{T}}Q_1^{\mathrm{T}}Q_1Q_2=Q_2^{\mathrm{T}}Q_2=I$.

7 When Gram-Schmidt gives Q with orthonormal columns, $Q^{\mathrm{T}}Q\widehat{\boldsymbol{x}}=Q^{\mathrm{T}}\boldsymbol{b}$ becomes $\widehat{\boldsymbol{x}}=Q^{\mathrm{T}}\boldsymbol{b}$. No cost to solve the normal equations!

8 If $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$ are *orthonormal* vectors in $\mathbf{R}^5$ then $\boldsymbol{p}=(\boldsymbol{q}_1^{\mathrm{T}}\boldsymbol{b})\boldsymbol{q}_1+(\boldsymbol{q}_2^{\mathrm{T}}\boldsymbol{b})\boldsymbol{q}_2$ is closest to $\boldsymbol{b}$. The error $\boldsymbol{e}=\boldsymbol{b}-\boldsymbol{p}$ is orthogonal to $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$.

9 (a) $Q=\begin{bmatrix}.8&-.6\\.6&.8\\0&0\end{bmatrix}$ has $P=QQ^{\mathrm{T}}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}$ = projection on the xy plane.

(b) $(QQ^{\mathrm{T}})(QQ^{\mathrm{T}}) = Q(Q^{\mathrm{T}}Q)Q^{\mathrm{T}} = QQ^{\mathrm{T}}$.

10 (a) If $\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3$ are *orthonormal* then the dot product of $\boldsymbol{q}_1$ with $c_1\boldsymbol{q}_1 + c_2\boldsymbol{q}_2 + c_3\boldsymbol{q}_3 = \boldsymbol{0}$ gives $c_1 = 0$. Similarly $c_2 = c_3 = 0$. This proves : *Independent* q's

(b) $Q\boldsymbol{x} = \boldsymbol{0}$ leads to $Q^{\mathrm{T}}Q\boldsymbol{x} = \boldsymbol{0}$ which says $\boldsymbol{x} = \boldsymbol{0}$.

11 (a) Two *orthonormal* vectors are $\boldsymbol{q}_1 = \frac{1}{10}(1,3,4,5,7)$ and $\boldsymbol{q}_2 = \frac{1}{10}(-7,3,4,-5,1)$

(b) Closest projection in the plane = *projection* $QQ^{\mathrm{T}}(1,0,0,0,0) = (0.5, -0.18, -0.24, 0.4, 0)$.

12 (a) Orthonormal $\boldsymbol{a}$'s: $\boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{b} = \boldsymbol{a}_1^{\mathrm{T}}(x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + x_3\boldsymbol{a}_3) = x_1(\boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{a}_1) = x_1$

(b) Orthogonal $\boldsymbol{a}$'s: $\boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{b} = \boldsymbol{a}_1^{\mathrm{T}}(x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + x_3\boldsymbol{a}_3) = x_1(\boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{a}_1)$. Therefore $x_1 = \boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{b}/\boldsymbol{a}_1^{\mathrm{T}}\boldsymbol{a}_1$

(c) x_1 is the first component of A^{-1} times $\boldsymbol{b}$ (A is 3 by 3 and invertible).

13 The multiple to subtract is $\frac{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}}$. Then $\boldsymbol{B} = \boldsymbol{b} - \frac{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}}\boldsymbol{a} = \begin{bmatrix} 4 \\ 0 \end{bmatrix} - 2\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$.

14 $$\begin{bmatrix} 1 & 4 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \boldsymbol{q}_1 & \boldsymbol{q}_2 \end{bmatrix} \begin{bmatrix} \|\boldsymbol{a}\| & \boldsymbol{q}_1^{\mathrm{T}}\boldsymbol{b} \\ 0 & \|\boldsymbol{B}\| \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{2} & 2\sqrt{2} \\ 0 & 2\sqrt{2} \end{bmatrix} = QR.$$

15 (a) Gram-Schmidt chooses $\boldsymbol{q}_1 = \boldsymbol{a}/||\boldsymbol{a}|| = \boldsymbol{\frac{1}{3}}(1,2,-2)$ and $\boldsymbol{q}_2 = \frac{1}{3}(2,1,2)$. Then $\boldsymbol{q}_3 = \frac{1}{3}(2,-2,-1)$.

(b) The nullspace of A^{T} contains $\boldsymbol{q}_3$

(c) $\widehat{\boldsymbol{x}} = (A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}(1,2,7) = (1,2)$.

16 $\boldsymbol{p} = (\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}/\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a})\boldsymbol{a} = 14\boldsymbol{a}/49 = 2\boldsymbol{a}/7$ is the projection of $\boldsymbol{b}$ onto $\boldsymbol{a}$. $\boldsymbol{q}_1 = \boldsymbol{a}/\|\boldsymbol{a}\| = \boldsymbol{a}/7$ is $(4,5,2,2)/7$. $\boldsymbol{B} = \boldsymbol{b} - \boldsymbol{p} = (-1,4,-4,-4)/7$ has $\|\boldsymbol{B}\| = 1$ so $\boldsymbol{q}_2 = \boldsymbol{B}$.

17 $\boldsymbol{p} = (\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}/\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a})\boldsymbol{a} = (3,3,3)$ and $\boldsymbol{e} = (-2,0,2)$. Then Gram-Schmidt will choose $\boldsymbol{q}_1 = (1,1,1)/\sqrt{3}$ and $\boldsymbol{q}_2 = (-1,0,1)/\sqrt{2}$.

18 $\boldsymbol{A} = \boldsymbol{a} = (1,-1,0,0)$; $\boldsymbol{B} = \boldsymbol{b} - \boldsymbol{p} = (\frac{1}{2}, \frac{1}{2}, -1, 0)$; $\boldsymbol{C} = \boldsymbol{c} - \boldsymbol{p}_A - \boldsymbol{p}_B = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, -1)$. Notice the pattern in those orthogonal $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$. In $\mathbf{R}^5$, $\boldsymbol{D}$ would be $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, -1)$.

Gram-Schmidt would go on to normalize $\boldsymbol{q}_1 = \boldsymbol{A}/||\boldsymbol{A}||$, $\boldsymbol{q}_2 = \boldsymbol{B}/||\boldsymbol{B}||$, $\boldsymbol{q}_3 = \boldsymbol{C}/||\boldsymbol{C}||$.

19 If $A = QR$ then $A^{\mathrm{T}}A = R^{\mathrm{T}}Q^{\mathrm{T}}QR = R^{\mathrm{T}}R =$ *lower* triangular times *upper* triangular (this Cholesky factorization of $A^{\mathrm{T}}A$ uses the same R as Gram-Schmidt!). The example has $A = \begin{bmatrix} -1 & 1 \\ 2 & 1 \\ 2 & 4 \end{bmatrix} = \dfrac{1}{3}\begin{bmatrix} -1 & 2 \\ 2 & -1 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 3 & 3 \\ 0 & 3 \end{bmatrix} = QR$ and the same R appears in

$A^{\mathrm{T}}A = \begin{bmatrix} 9 & 9 \\ 9 & 18 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 3 & 3 \end{bmatrix}\begin{bmatrix} 3 & 3 \\ 0 & 3 \end{bmatrix} = R^{\mathrm{T}}R.$

20 (a) *True* because $Q^{\mathrm{T}}Q = I$ leads to $(Q^{-1})\,(Q^{-1}) = I$.

(b) *True*. $Q\boldsymbol{x} = x_1\boldsymbol{q}_1 + x_2\boldsymbol{q}_2$. $\|Q\boldsymbol{x}\|^2 = x_1^2 + x_2^2$ because $\boldsymbol{q}_1 \cdot \boldsymbol{q}_2 = 0$. Also $\|Q\boldsymbol{x}\|^2 = \boldsymbol{x}^{\mathrm{T}}Q^{\mathrm{T}}Q\boldsymbol{x} = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}$.

21 The orthonormal vectors are $\boldsymbol{q}_1 = (1,1,1,1)/2$ and $\boldsymbol{q}_2 = (-5,-1,1,5)/\sqrt{52}$. Then $\boldsymbol{b} = (-4,-3,3,0)$ projects to $\boldsymbol{p} = (\boldsymbol{q}_1^{\mathrm{T}}\boldsymbol{b})\boldsymbol{q}_1 + (\boldsymbol{q}_2^{\mathrm{T}}\boldsymbol{b})\boldsymbol{q}_2 = (-7,-3,-1,3)/2$. And $\boldsymbol{b} - \boldsymbol{p} = (-1,-3,7,-3)/2$ is orthogonal to both $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$.

22 $A = (1,1,2),\ B = (1,-1,0),\ C = (-1,-1,1)$. These are not yet unit vectors. As in Problem 18, Gram-Schmidt will divide by $\|\boldsymbol{A}\|$ and $\|\boldsymbol{B}\|$ and $\|\boldsymbol{C}\|$.

23 You can see why $\boldsymbol{q}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{q}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \boldsymbol{q}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$. $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 & 4 \\ 0 & 3 & 6 \\ 0 & 0 & 5 \end{bmatrix} =$ QR. This Q is just a permutation matrix—certainly orthogonal.

24 (a) One basis for the subspace $\boldsymbol{S}$ of solutions to $x_1 + x_2 + x_3 - x_4 = 0$ is the 3 special solutions $\boldsymbol{v}_1 = (-1,1,0,0)$, $\boldsymbol{v}_2 = (-1,0,1,0)$, $\boldsymbol{v}_3 = (1,0,0,1)$

(b) Since $\boldsymbol{S}$ contains solutions to $(1,1,1,-1)^{\mathrm{T}}\boldsymbol{x} = 0$, a basis for $\boldsymbol{S}^{\perp}$ is $(1,1,1,-1)$

(c) Split $(1,1,1,1)$ into $\boldsymbol{b}_1 + \boldsymbol{b}_2$ by projection on $\boldsymbol{S}^{\perp}$ and $\boldsymbol{S}$: $\boldsymbol{b}_2 = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2})$ and $\boldsymbol{b}_1 = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{3}{2})$.

25 This question shows 2 by 2 formulas for QR; breakdown $R_{22} = 0$ for singular A. Nonsingular example $\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = \dfrac{1}{\sqrt{5}}\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} \cdot \dfrac{1}{\sqrt{5}}\begin{bmatrix} 5 & 3 \\ 0 & 1 \end{bmatrix}$.

Singular example $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \cdot \frac{1}{\sqrt{2}} \begin{bmatrix} 2 & 2 \\ 0 & \mathbf{0} \end{bmatrix}$.
The Gram-Schmidt process breaks down when $ad - bc = 0$.

26 $(\boldsymbol{q}_2^{\mathrm{T}} \boldsymbol{C}^*) \boldsymbol{q}_2 = \dfrac{\boldsymbol{B}^{\mathrm{T}} \boldsymbol{c}}{\boldsymbol{B}^{\mathrm{T}} \boldsymbol{B}} \boldsymbol{B}$ because $\boldsymbol{q}_2 = \dfrac{\boldsymbol{B}}{\|\boldsymbol{B}\|}$ and the extra $\boldsymbol{q}_1$ in $\boldsymbol{C}^*$ is orthogonal to $\boldsymbol{q}_2$.

27 When $\boldsymbol{a}$ and $\boldsymbol{b}$ are not orthogonal, the projections onto these lines *do not add* to the projection onto the plane of $\boldsymbol{a}$ and $\boldsymbol{b}$. We must use the orthogonal $\boldsymbol{A}$ and $\boldsymbol{B}$ (or orthonormal $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$) to be allowed to add projections on those lines.

28 There are $\frac{1}{2}m^2 n$ multiplications to find the numbers r_{kj} and the same for v_{ij}.

29 $\boldsymbol{q}_1 = \frac{1}{3}(2, 2, -1)$, $\boldsymbol{q}_2 = \frac{1}{3}(2, -1, 2)$, $\boldsymbol{q}_3 = \frac{1}{3}(1, -2, -2)$.

30 The columns of the wavelet matrix W are *orthonormal*. Then $W^{-1} = W^{\mathrm{T}}$. This is a useful orthonormal basis with many zeros.

31 (a) $c = \frac{1}{2}$ normalizes all the orthogonal columns to have unit length (b) The projection $(\boldsymbol{a}^{\mathrm{T}} \boldsymbol{b} / \boldsymbol{a}^{\mathrm{T}} \boldsymbol{a}) \boldsymbol{a}$ of $\boldsymbol{b} = (1, 1, 1, 1)$ onto the first column is $\boldsymbol{p}_1 = \frac{1}{2}(-1, 1, 1, 1)$. (Check $\boldsymbol{e} = \mathbf{0}$.) To project onto the plane, add $\boldsymbol{p}_2 = \frac{1}{2}(1, -1, 1, 1)$ to get $(0, 0, 1, 1)$.

32 $Q_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ reflects across x axis, $Q_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix}$ across plane $y + z = 0$.

33 Orthogonal and lower triangular $\Rightarrow \pm 1$ on the main diagonal and zeros elsewhere.

34 (a) $Q\boldsymbol{u} = (I - 2\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}})\boldsymbol{u} = \boldsymbol{u} - 2\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u}$. This is $-\boldsymbol{u}$, provided that $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u}$ equals 1 (b) $Q\boldsymbol{v} = (I - 2\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}})\boldsymbol{v} = \boldsymbol{u} - 2\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v} = \boldsymbol{u}$, provided that $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v} = 0$.

35 Starting from $\boldsymbol{A} = (1, -1, 0, 0)$, the orthogonal (not orthonormal) vectors $\boldsymbol{B} = (1, 1, -2, 0)$ and $\boldsymbol{C} = (1, 1, 1, -3)$ and $\boldsymbol{D} = (1, 1, 1, 1)$ are in the directions of $\boldsymbol{q}_2, \boldsymbol{q}_3, \boldsymbol{q}_4$. The 4 by 4 and 5 by 5 matrices with *integer orthogonal columns* (not orthogonal rows, since not orthonormal Q!) are

$$\begin{bmatrix} \\ \boldsymbol{A} & \boldsymbol{B} & \boldsymbol{C} & \boldsymbol{D} \\ \\ \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 \\ 0 & -2 & 1 & 1 \\ 0 & 0 & -3 & 1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 & 1 \\ 0 & -2 & 1 & 1 & 1 \\ 0 & 0 & -3 & 1 & 1 \\ 0 & 0 & 0 & -4 & 1 \end{bmatrix}$$

36 $[Q, R] = \boldsymbol{qr}(A)$ produces from A (m by n of rank n) a "*full-size*" square $Q = [\,Q_1 \;\; Q_2\,]$ and $\begin{bmatrix} R \\ 0 \end{bmatrix}$. The columns of Q_1 are the orthonormal basis from Gram-Schmidt of the *column space* of A. The $m - n$ columns of Q_2 are an orthonormal basis for the *left nullspace* of A. Together the columns of $Q = [\,Q_1 \;\; Q_2\,]$ are an orthonormal basis for $\mathbf{R}^m$.

37 This question describes the next $\boldsymbol{q}_{n+1}$ in Gram-Schmidt using the matrix Q with the columns $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n$ (instead of using those $\boldsymbol{q}$'s separately). Start from $\boldsymbol{a}$, *subtract its projection* $\boldsymbol{p} = QQ^{\mathrm{T}}\boldsymbol{a}$ onto the earlier $\boldsymbol{q}$'s, *divide by the length* of $\boldsymbol{e} = \boldsymbol{a} - QQ^{\mathrm{T}}\boldsymbol{a}$ to get the next $\boldsymbol{q}_{n+1} = \boldsymbol{e}/\|\boldsymbol{e}\|$.