

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 $\text{rank} = 1$.

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 $\mathbf{x}_n = \mathbf{0}$, 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 $\mathcal{N}(A)$ and $\mathcal{C}(A^T)$ 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 $\mathcal{C}(A)$ and $\mathcal{N}(A^T)$ 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 $\text{rank} = 2$, all those four subspaces have dimension at least 2 which is impossible for 3 by 3.

5 (a) If $A\mathbf{x} = \mathbf{b}$ has a solution and $A^T\mathbf{y} = \mathbf{0}$, then \mathbf{y} is perpendicular to \mathbf{b} . $\mathbf{b}^T\mathbf{y} = (A\mathbf{x})^T\mathbf{y} = \mathbf{x}^T(A^T\mathbf{y}) = 0$. This says again that $\mathcal{C}(A)$ is orthogonal to $\mathcal{N}(A^T)$.

(b) If $A^T\mathbf{y} = (1, 1, 1)$ 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 \mathbf{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, $\mathbf{y} = (1, 1, -1)$ is in the left nullspace. $A\mathbf{x} = \mathbf{b}$ would need $0 = (\mathbf{y}^T A)\mathbf{x} = \mathbf{y}^T \mathbf{b} = 1$ but here $\mathbf{y}^T \mathbf{b} = 1$.
- 7** Multiply the 3 equations by $\mathbf{y} = (1, 1, -1)$. Then $x_1 - x_2 = 1$ plus $x_2 - x_3 = 1$ minus $x_1 - x_3 = 1$ is $0 = 1$. Key point: This \mathbf{y} in $\mathcal{N}(A^T)$ is not orthogonal to $\mathbf{b} = (1, 1, 1)$ so \mathbf{b} is not in the column space and $A\mathbf{x} = \mathbf{b}$ has *no solution*.
- 8** Figure 4.3 has $\mathbf{x} = \mathbf{x}_r + \mathbf{x}_n$, where \mathbf{x}_r is in the row space and \mathbf{x}_n is in the nullspace. Then $A\mathbf{x}_n = \mathbf{0}$ and $A\mathbf{x} = A\mathbf{x}_r + A\mathbf{x}_n = A\mathbf{x}_r$. The example has $\mathbf{x} = (1, 0)$ and row space = line through $(1, 1)$ so the splitting is $\mathbf{x} = \mathbf{x}_r + \mathbf{x}_n = (\frac{1}{2}, \frac{1}{2}) + (\frac{1}{2}, -\frac{1}{2})$. All $A\mathbf{x}$ are in $\mathcal{C}(A)$.
- 9** $A\mathbf{x}$ is always in the *column space* of A . If $A^T A\mathbf{x} = \mathbf{0}$ then $A\mathbf{x}$ is also in the *nullspace* of A^T . Those subspaces are perpendicular. So $A\mathbf{x}$ is perpendicular to itself. Conclusion: $A\mathbf{x} = \mathbf{0}$ if $A^T A\mathbf{x} = \mathbf{0}$.
- 10** (a) With $A^T = A$, the column and row spaces are the *same*. The nullspace is always perpendicular to the row space. (b) \mathbf{x} is in the nullspace and \mathbf{z} is in the column space = row space: so these “eigenvectors” \mathbf{x} and \mathbf{z} have $\mathbf{x}^T \mathbf{z} = 0$.
- 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 $\mathcal{N}(A^T)$ 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** $\mathbf{x} = (2, 0)$ splits into $\mathbf{x}_r + \mathbf{x}_n = (1, -1) + (1, 1)$. Notice $\mathcal{N}(A^T)$ is the $y - z$ plane.
- 13** $V^T W = \text{zero matrix}$ makes each column of V orthogonal to each column of W . This means: each basis vector for \mathbf{V} is orthogonal to each basis vector for \mathbf{W} . Then *every* \mathbf{v} in \mathbf{V} (combinations of the basis vectors) is orthogonal to *every* \mathbf{w} in \mathbf{W} .
- 14** $A\mathbf{x} = B\hat{\mathbf{x}}$ means that $[A \ B] \begin{bmatrix} \mathbf{x} \\ -\hat{\mathbf{x}} \end{bmatrix} = \mathbf{0}$. Three homogeneous equations (zero right hand sides) in four unknowns always have a nonzero solution. Here $\mathbf{x} = (3, 1)$ and

$\hat{x} = (1, 0)$ and $Ax = B\hat{x} = (5, 6, 5)$ 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 $p + q > n$. (The $p + q$ basis vectors of \mathbf{V} and \mathbf{W} cannot be independent, so some combination of the basis vectors of \mathbf{V} is also a combination of the basis vectors of \mathbf{W} .)
- 16** $A^T \mathbf{y} = \mathbf{0}$ leads to $(Ax)^T \mathbf{y} = \mathbf{x}^T A^T \mathbf{y} = 0$. Then $\mathbf{y} \perp Ax$ and $\mathbf{N}(A^T) \perp \mathbf{C}(A)$.
- 17** If \mathbf{S} is the subspace of \mathbf{R}^3 containing only the zero vector, then \mathbf{S}^\perp is all of \mathbf{R}^3 . If \mathbf{S} is spanned by $(1, 1, 1)$, then \mathbf{S}^\perp is the plane spanned by $(1, -1, 0)$ and $(1, 0, -1)$. If \mathbf{S} is spanned by $(2, 0, 0)$ and $(0, 0, 3)$, then \mathbf{S}^\perp is the line spanned by $(0, 1, 0)$.
- 18** \mathbf{S}^\perp contains all vectors perpendicular to those two given vectors. So \mathbf{S}^\perp is the nullspace of $A = \begin{bmatrix} 1 & 5 & 1 \\ 2 & 2 & 2 \end{bmatrix}$. Therefore \mathbf{S}^\perp is a *subspace* even if \mathbf{S} is not.
- 19** L^\perp is the 2-dimensional subspace (a plane) in \mathbf{R}^3 perpendicular to L . Then $(L^\perp)^\perp$ is a 1-dimensional subspace (a line) perpendicular to L^\perp . In fact $(L^\perp)^\perp = L$.
- 20** If \mathbf{V} is the whole space \mathbf{R}^4 , then \mathbf{V}^\perp contains only the zero vector. Then $(\mathbf{V}^\perp)^\perp =$ all vectors perpendicular to the zero vector $= \mathbf{R}^4 = \mathbf{V}$.
- 21** For example $(-5, 0, 1, 1)$ and $(0, 1, -1, 0)$ span $\mathbf{S}^\perp =$ 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 \mathbf{P}^\perp orthogonal to \mathbf{P} . $A = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$ has \mathbf{P} as its nullspace and \mathbf{P}^\perp as its row space.
- 23** \mathbf{x} in \mathbf{V}^\perp is perpendicular to every vector in \mathbf{V} . Since \mathbf{V} contains all the vectors in \mathbf{S} , \mathbf{x} is perpendicular to every vector in \mathbf{S} . So every \mathbf{x} in \mathbf{V}^\perp is also in \mathbf{S}^\perp .
- 24** $AA^{-1} = I$: Column 1 of A^{-1} is orthogonal to rows 2, 3, ..., n and therefore to the space spanned by those rows.
- 25** If the columns of A are unit vectors, all mutually perpendicular, then $A^T A = I$. 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. $A^T A = 9I$ is *diagonal*: $(A^T A)_{ij} = (\text{column } i \text{ of } A) \cdot (\text{column } j \text{ of } A)$. When the columns are *unit vectors*, then $A^T A = I$.

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 $(-2, 1)$. The nullspace of the 2 by 2 matrix is the line $3x + y = 0$. One particular vector in the nullspace is $(-1, 3)$.

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 $\mathbf{v} = (1, 2, 3)$ in row and column spaces. B has \mathbf{v} in its column space and nullspace. \mathbf{v} **can not** be in the nullspace and row space, or in the left nullspace and column space. These spaces are orthogonal and $\mathbf{v}^T \mathbf{v} \neq 0$.

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 $C(B) \leq$ dimension of $N(A)$. This means $\text{rank}(B) \leq 4 - \text{rank}(A)$.

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

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

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

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

Problem Set 4.2, page 214

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

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

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

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

The picture for part (a) has the vector \mathbf{b} at an angle θ with the horizontal \mathbf{a} . The picture for part (b) has vectors \mathbf{a} and \mathbf{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 $P_1 \mathbf{b} = \frac{1}{3} \begin{bmatrix} 5 \\ 5 \\ 5 \end{bmatrix}$. $P_2 = \frac{1}{11} \begin{bmatrix} 1 & 3 & 1 \\ 3 & 9 & 3 \\ 1 & 3 & 1 \end{bmatrix}$ and $P_2 \mathbf{b} = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$.

4 $P_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $P_2 = \frac{\mathbf{a}\mathbf{a}^T}{\mathbf{a}^T \mathbf{a}} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$. P_1 projects onto $(1, 0)$, P_2 projects onto $(1, -1)$. $P_1 P_2 \neq 0$ 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 $\mathbf{a}_1 = (-1, 2, 2)$ and $\mathbf{a}_2 = (2, 2, -1)$. $P_1 P_2 = \text{zero matrix}$ because $\mathbf{a}_1 \perp \mathbf{a}_2$.

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

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 $\mathbf{p}_1 = (1, 0)$ and $\mathbf{p}_2 = \frac{3}{5}(1, 2)$. Then $\mathbf{p}_1 + \mathbf{p}_2 \neq \mathbf{b}$. The sum of projections is not a projection because $(1, 0)$ and $(2, 1)$ are *not orthogonal*.

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

$$10 \quad P_2 = \frac{\mathbf{a}_2 \mathbf{a}_2^T}{\mathbf{a}_2^T \mathbf{a}_2} = \begin{bmatrix} 0.2 & 0.4 \\ 0.4 & 0.8 \end{bmatrix}, P_2 \mathbf{a}_1 = \begin{bmatrix} 0.2 \\ 0.4 \end{bmatrix}, P_1 = \frac{\mathbf{a}_1 \mathbf{a}_1^T}{\mathbf{a}_1^T \mathbf{a}_1} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, P_1 P_2 \mathbf{a}_1 = \begin{bmatrix} 0.2 \\ 0 \end{bmatrix}.$$

This is not $\mathbf{a}_1 = (1, 0)$.
No, $\mathbf{P}_1 \mathbf{P}_2 \neq (\mathbf{P}_1 \mathbf{P}_2)^2$.

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

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

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

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

Certainly $(P_2)^2 = P_2$. A true projection matrix.

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

14 The projection of this $\mathbf{b} =$ column 1 of A onto the column space of A is \mathbf{b} itself because \mathbf{b} is in that column space. But P is not necessarily I .

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

15 $2A$ has the same column space as A . Then P is the same for A and $2A$ but $\hat{\mathbf{x}}$ for $2A$ is half of $\hat{\mathbf{x}}$ for A .

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

17 If $P^2 = P$ then $(I - P)^2 = (I - P)(I - P) = I - PI - IP + P^2 = I - P$. 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 $P = A(A^T A)^{-1} A^T$ is
$$\begin{bmatrix} 5/6 & 1/6 & 1/3 \\ 1/6 & 5/6 & -1/3 \\ 1/3 & -1/3 & 1/3 \end{bmatrix}.$$

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

21 $(A(A^T A)^{-1} A^T)^2 = A(A^T A)^{-1} (A^T A) (A^T A)^{-1} A^T = A(A^T A)^{-1} A^T$. So $P^2 = P$.

Pb is in the column space (where P projects). Then its projection $P(Pb)$ is also Pb .

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Problem Set 4.3, page 229

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

$$A^T A \hat{\mathbf{x}} = A^T \mathbf{b} \quad \text{gives} \quad \hat{\mathbf{x}} = \begin{bmatrix} 1 \\ 4 \end{bmatrix} \quad \text{and} \quad \mathbf{p} = A \hat{\mathbf{x}} = \begin{bmatrix} 1 \\ 5 \\ 13 \\ 17 \end{bmatrix} \quad \text{and} \quad \mathbf{e} = \mathbf{b} - \mathbf{p} = \begin{bmatrix} -1 \\ 3 \\ -5 \\ 3 \end{bmatrix} \\ E = \|\mathbf{e}\|^2 = 44$$

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

$$\hat{\mathbf{x}} = \begin{bmatrix} 1 \\ 4 \end{bmatrix} \quad \text{exactly solves } A\hat{\mathbf{x}} = \mathbf{p}.$$

3 In Problem 2, $\mathbf{p} = A(A^T A)^{-1} A^T \mathbf{b} = (1, 5, 13, 17)$ and $\mathbf{e} = \mathbf{b} - \mathbf{p} = (-1, 3, -5, 3)$.

This \mathbf{e} is perpendicular to both columns of A . This shortest distance $\|\mathbf{e}\|$ is $\sqrt{44}$.

- 4 $E = (C + 0D)^2 + (C + 1D - 8)^2 + (C + 3D - 8)^2 + (C + 4D - 20)^2$. 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 $E = (C - 0)^2 + (C - 8)^2 + (C - 8)^2 + (C - 20)^2$. $A^T = [1 \ 1 \ 1 \ 1]$ and $A^T A = [4]$. $A^T \mathbf{b} = [36]$ and $(A^T A)^{-1} A^T \mathbf{b} = \mathbf{9} =$ best height C for the horizontal line. Errors $\mathbf{e} = \mathbf{b} - \mathbf{p} = (-9, -1, -1, 11)$ still add to zero.

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

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

- 8 $\hat{x} = \mathbf{a}^T \mathbf{b} / \mathbf{a}^T \mathbf{a} = 56/13$ and $\mathbf{p} = (56/13)(0, 1, 3, 4)$. $(C, D) = (9, 56/13)$ 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
$$\begin{array}{l} \text{Parabola} \\ \text{Project } \mathbf{b} \\ \text{4D to 3D} \end{array} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{bmatrix} \begin{bmatrix} C \\ D \\ E \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ 8 \\ 20 \end{bmatrix}. \quad A^T A \hat{\mathbf{x}} = \begin{bmatrix} 4 & 8 & 26 \\ 8 & 26 & 92 \\ 26 & 92 & 338 \end{bmatrix} \begin{bmatrix} C \\ D \\ E \end{bmatrix} = \begin{bmatrix} 36 \\ 112 \\ 400 \end{bmatrix}.$$

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

10
$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 3 & 9 & 27 \\ 1 & 4 & 16 & 64 \end{bmatrix} \begin{bmatrix} C \\ D \\ E \\ F \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ 8 \\ 20 \end{bmatrix}. \quad \text{Then } \begin{bmatrix} C \\ D \\ E \\ F \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 0 \\ 47 \\ -28 \\ 5 \end{bmatrix}.$$
 Exact cubic so $\mathbf{p} = \mathbf{b}$, $\mathbf{e} = \mathbf{0}$. 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 $\hat{\mathbf{b}} = 9$ at center time, $\hat{t} = 2$.
 (b) The first equation $Cm + D \sum t_i = \sum b_i$ divided by m gives $C + D\hat{t} = \hat{\mathbf{b}}$. This shows: The best line goes through $\hat{\mathbf{b}}$ at time \hat{t} .

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

(b) $\mathbf{e} = \mathbf{b} - \hat{\mathbf{x}} \mathbf{a}$ and $\|\mathbf{e}\|^2 = (b_1 - \text{mean})^2 + \dots + (b_m - \text{mean})^2 = \mathbf{variance}$ (denoted by σ^2).

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

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

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

15 When A has 1 column of 4 ones, Problem 14 gives the expected error $(\hat{\mathbf{x}} - \mathbf{x})^2$ as $\sigma^2 (A^T A)^{-1} = \sigma^2 / 4$. 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 $\frac{1}{10} b_{10} + \frac{9}{10} \hat{x}_9 = \frac{1}{10} (b_1 + \dots + b_{10})$. Knowing \hat{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 $\hat{\mathbf{x}} = \begin{bmatrix} 9 \\ 4 \end{bmatrix}$ comes from $\begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 35 \\ 42 \end{bmatrix}$.

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

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

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

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

22 The least squares equation is $\begin{bmatrix} 5 & 0 \\ 0 & 10 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 5 \\ -10 \end{bmatrix}$. Solution: $C = 1, D = -1$.

The best line is $b = 1 - t$. Symmetric t 's \Rightarrow diagonal $A^T A \Rightarrow$ easy solution.

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

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

25 3 points on a line will 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 $y = (t_2 - t_3, t_3 - t_1, t_1 - t_2)$ in the left nullspace of A . b is in the column space! Then $y^T b = 0$ 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

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} C \\ D \\ E \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 3 \\ 4 \end{bmatrix}. \text{ Then } A^T A = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

and $A^T b = \begin{bmatrix} 8 \\ -2 \\ -3 \end{bmatrix}$ 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 $C = 2 =$ 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 $A^T A$ is not invertible and the usual formula $P = A(A^T A)^{-1} A^T$ will fail. Replace A in that formula by the matrix B that keeps *only the pivot columns of A* .

29 Only 1 plane contains $0, a_1, a_2$ unless a_1, a_2 are *dependent*. Same test for a_1, \dots, a_{n-1} .

If they are dependent, there is a vector v perpendicular to all the a 's. Then they all lie on the plane $v^T x = 0$ going through $x = (0, 0, \dots, 0)$.

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

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 $\mathbf{q}_1 = (\frac{2}{3}, \frac{2}{3}, -\frac{1}{3})$, $\mathbf{q}_2 = (-\frac{1}{3}, \frac{2}{3}, \frac{2}{3})$. $Q^T Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ but $Q Q^T = \begin{bmatrix} 5/9 & 2/9 & -4/9 \\ 2/9 & 8/9 & 2/9 \\ -4/9 & 2/9 & 5/9 \end{bmatrix}$.

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

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

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

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

- 6 $Q_1 Q_2$ is orthogonal because $(Q_1 Q_2)^T Q_1 Q_2 = Q_2^T Q_1^T Q_1 Q_2 = Q_2^T Q_2 = I$.

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

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

The error $\mathbf{e} = \mathbf{b} - \mathbf{p}$ is orthogonal to \mathbf{q}_1 and \mathbf{q}_2 .

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

$$(b) (QQ^T)(QQ^T) = Q(Q^TQ)Q^T = QQ^T.$$

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

$$(b) Q\mathbf{x} = \mathbf{0} \text{ leads to } Q^TQ\mathbf{x} = \mathbf{0} \text{ which says } \mathbf{x} = \mathbf{0}.$$

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

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

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

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

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

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

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

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

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

$$(c) \hat{\mathbf{x}} = (A^T A)^{-1} A^T(1, 2, 7) = (1, 2).$$

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

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

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

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

- 19** If $A = QR$ then $A^T A = R^T Q^T QR = R^T R =$ lower triangular times upper triangular (this Cholesky factorization of $A^T A$ uses the same R as Gram-Schmidt!). The example

$$\text{has } A = \begin{bmatrix} -1 & 1 \\ 2 & 1 \\ 2 & 4 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} -1 & 2 \\ 2 & -1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 3 & 3 \\ 0 & 3 \end{bmatrix} = QR \text{ and the same } R \text{ appears in}$$

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

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

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

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

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

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

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

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

- 25** This question shows 2 by 2 formulas for QR ; breakdown $R_{22} = 0$ for singular A .

Nonsingular example $\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} \cdot \frac{1}{\sqrt{5}} \begin{bmatrix} 5 & 3 \\ 0 & 1 \end{bmatrix}$.

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

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

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

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

30 The columns of the wavelet matrix W are *orthonormal*. Then $W^{-1} = W^T$. 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 $(\mathbf{a}^T \mathbf{b} / \mathbf{a}^T \mathbf{a})\mathbf{a}$ of $\mathbf{b} = (1, 1, 1, 1)$ onto the first column is $\mathbf{p}_1 = \frac{1}{2}(-1, 1, 1, 1)$. (Check $\mathbf{e} = \mathbf{0}$.) To project onto the plane, add $\mathbf{p}_2 = \frac{1}{2}(1, -1, 1, 1)$ 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) $Q\mathbf{u} = (I - 2\mathbf{u}\mathbf{u}^T)\mathbf{u} = \mathbf{u} - 2\mathbf{u}\mathbf{u}^T\mathbf{u}$. This is $-\mathbf{u}$, provided that $\mathbf{u}^T\mathbf{u}$ equals 1

(b) $Q\mathbf{v} = (I - 2\mathbf{u}\mathbf{u}^T)\mathbf{v} = \mathbf{v} - 2\mathbf{u}\mathbf{u}^T\mathbf{v} = \mathbf{v}$, provided that $\mathbf{u}^T\mathbf{v} = 0$.

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

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

- 36** $[Q, R] = qr(A)$ 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 \mathbf{q}_{n+1} in Gram-Schmidt using the matrix Q with the columns $\mathbf{q}_1, \dots, \mathbf{q}_n$ (instead of using those \mathbf{q} 's separately). Start from \mathbf{a} , subtract its projection $\mathbf{p} = Q^T \mathbf{a}$ onto the earlier \mathbf{q} 's, divide by the length of $\mathbf{e} = \mathbf{a} - Q^T \mathbf{a}$ to get $\mathbf{q}_{n+1} = \mathbf{e}/\|\mathbf{e}\|$.