INTRODUCTION TO LINEAR ALGEBRA

Sixth Edition

SOLUTIONS TO PROBLEM SETS

Gilbert Strang

Massachusetts Institute of Technology

math.mit.edu/weborder.php (orders)

math.mit.edu/linearalgebra (book website)

math.mit.edu/~gs (author website)

www.wellesleycambridge.com (all books)

Wellesley - Cambridge Press
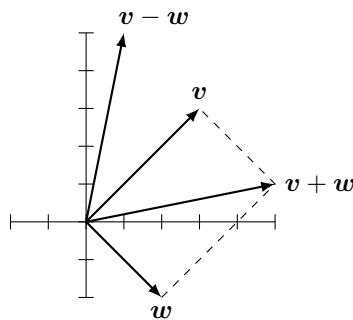
Box 812060

Wellesley, Massachusetts 02482

Problem Set 1.1, page 6

1 $c = ma$ and $d = mb$ lead to $ad = amb = bc$. With no zeros, $ad = bc$ is the equation for a 2×2 matrix to have rank 1.

2 The three edges going around the triangle are $\boldsymbol{u} = (5, 0), \boldsymbol{v} = (-5, 12), \boldsymbol{w} = (0, -12)$. Their sum is $u + v + w = (0, 0)$. Their lengths are $||\boldsymbol{u}|| = 5, ||\boldsymbol{v}|| = 13, ||\boldsymbol{w}|| = 12$. This is a $5 - 12 - 13$ right triangle with $5^2 + 12^2 = 25 + 144 = 169 = 13^2$—the best numbers after the $3 - 4 - 5$ right triangle if we don't count $6 - 8 - 10$.

3 The combinations give (a) a line in $\mathbf{R}^3$ (b) a plane in $\mathbf{R}^3$ (c) all of $\mathbf{R}^3$.

4 $\boldsymbol{v} + \boldsymbol{w} = (2, 3)$ and $\boldsymbol{v} - \boldsymbol{w} = (6, -1)$ will be the diagonals of the parallelogram with $\boldsymbol{v}$ and $\boldsymbol{w}$ as two sides going out from $(0, 0)$.
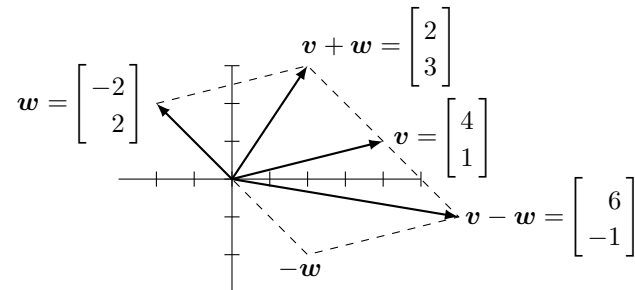
5 This problem gives the diagonals $\boldsymbol{v} + \boldsymbol{w} = (5, 1)$ and $\boldsymbol{v} - \boldsymbol{w} = (1, 5)$ of the parallelogram and asks for the sides $\boldsymbol{v}$ and $\boldsymbol{w}$: The opposite of Problem 4. In this example $\boldsymbol{v} = (3, 3)$ and $\boldsymbol{w} = (2, -2)$. Those come from $\boldsymbol{v} = \frac{1}{2}(\boldsymbol{v} + \boldsymbol{w}) + \frac{1}{2}(\boldsymbol{v} - \boldsymbol{w})$ and $\boldsymbol{w} = \frac{1}{2}(\boldsymbol{v} + \boldsymbol{w}) - \frac{1}{2}(\boldsymbol{v} - \boldsymbol{w})$.

6 $3v + w = (7, 5)$ and $cv + dw = (2c + d, c + 2d)$.

7 $u + v = (-2, 3, 1)$ and $u + v + w = (0, 0, 0)$ and $2u + 2v + w = ($ add first answers $) = (-2, 3, 1)$. The vectors u, v, w are in the same plane because a combination $u + v + w$ gives $(0, 0, 0)$. Stated another way: $u = -v - w$ is in the plane of v and w.

8 The components of every $cv + dw$ add to zero because the components of $v = (1, -2, 1)$ and of $w = (0, 1, -1)$ add to zero. $c = 3$ and $d = 9$ give $3v + 9w = (3, 3, -6)$. There is no solution to $cv + dw = (3, 3, 6)$ because $3 + 3 + 6$ is not zero.

9 The nine combinations $c(2, 1) + d(0, 1)$ with $c = 0, 1, 2$ and $d = 0, 1, 2$ will lie on a lattice. If we took all whole numbers c and d, the lattice would lie over the whole plane.
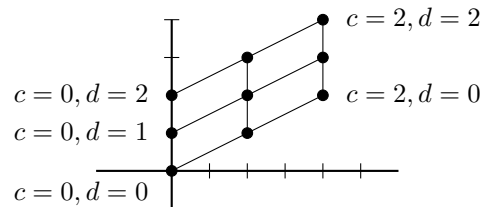
10 The question is whether (a, b, c) is a combination $x_1 u + x_2 v$. Can we solve

$$x_1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} ?$$

Certainly x_1 has to be a. Certainly x_2 has to be c. So the middle components give the **requirement $a + c = b$**.

11 The fourth corner can be $(4, 4)$ or $(4, 0)$ or $(-2, 2)$. Draw 3 possible parallelograms!

12 Four more corners $(1, 1, 0), (1, 0, 1), (0, 1, 1), (1, 1, 1)$. The center point is $\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$. Centers of 6 faces: $\left(\frac{1}{2}, \frac{1}{2}, 0\right), \left(\frac{1}{2}, \frac{1}{2}, 1\right) \& \left(0, \frac{1}{2}, \frac{1}{2}\right), \left(1, \frac{1}{2}, \frac{1}{2}\right) \& \left(\frac{1}{2}, 0, \frac{1}{2}\right), \left(\frac{1}{2}, 1, \frac{1}{2}\right)$. 12 edges.

13 The combinations of $i = (1, 0, 0)$ and $i + j = (1, 1, 0)$ fill the **xy plane** in xyz space.

14 (a) Sum = zero vector.   (b) Sum = -2:00 vector = 8:00 vector.

(c) 2:00 is $30°$ from horizontal = $\left(\cos \frac{\pi}{6}, \sin \frac{\pi}{6}\right) = (\sqrt{3}/2, 1/2)$.

15 Moving the origin to 6:00 adds $j = (0, 1)$ to every vector. So the sum of twelve vectors changes from **0** to $12j = (0, 12)$.

16 First part : u, v, w are all in the same direction.

Second part : Some combination of u, v, w gives the zero vector but those 3 vectors are not on a line. Then their combinations fill a plane in 3D.

17 The two equations are $c + 3d = 14$ and $2c + d = 8$. The solution is $c = 2$ and $d = 4$.

18 The point $\dfrac{3}{4}v + \dfrac{1}{4}w$ is three-fourths of the way to v starting from w. The vector $\dfrac{1}{4}v + \dfrac{1}{4}w$ is halfway to $u = \dfrac{1}{2}v + \dfrac{1}{2}w$. The vector $v + w$ is $2u$ (the far corner of the parallelogram).

19 The combinations $cv + dw$ with $0 \le c \le 1$ and $0 \le d \le 1$ *fill the parallelogram* with sides v and w. For example, if $v = (1, 0)$ and $w = (0, 1)$ then $cv + dw$ fills the unit square. In a special case like $v = (a, 0)$ and $w = (b, 0)$ these combinations only fill a segment of a line.

With $c \ge 0$ and $d \ge 0$ we get the infinite "cone" or "wedge" between v and w. For example, if $v = (1, 0)$ and $w = (0, 1)$, then the cone is the whole first quadrant $x \ge 0$, $y \ge 0$. *Question*: What if $w = -v$? The cone opens to a half-space. But the combinations of $v = (1, 0)$ and $w = (-1, 0)$ only fill a line.

20 (a) $\frac{1}{3}u + \frac{1}{3}v + \frac{1}{3}w$ is the center of the triangle between u, v and w; $\frac{1}{2}u + \frac{1}{2}w$ lies halfway between u and w (b) To fill the triangle keep $c \ge 0$, $d \ge 0$, $e \ge 0$, and $c + d + e = \mathbf{1}$.

21 The sum is $(v - u) + (w - v) + (u - w) = \mathbf{zero\ vector}$. Those three sides of a triangle are in the same plane !

22 The vector $\frac{1}{2}(u + v + w)$ is *outside* the pyramid because $c + d + e = \frac{1}{2} + \frac{1}{2} + \frac{1}{2} > 1$.

23 *All* vectors in 3D are combinations of u, v, w as drawn (not in the same plane). Start by seeing that $cu + dv$ fills a plane, then adding all the vectors ew fills all of $\mathbf{R}^3$. Different answer when u, v, w are in the same plane.

24 A four-dimensional cube has $2^4 = 16$ corners and $2 \cdot 4 = 8$ three-dimensional faces and 24 two-dimensional faces and 32 edges.

25 *Fact*: For any three vectors $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ in the plane, some combination $c\boldsymbol{u} + d\boldsymbol{v} + e\boldsymbol{w}$ is the zero vector (beyond the obvious $c = d = e = 0$). So if there is one combination $C\boldsymbol{u} + D\boldsymbol{v} + E\boldsymbol{w}$ that produces $\boldsymbol{b}$, there will be many more—just add c, d, e or $2c, 2d, 2e$ to the particular solution C, D, E.

The example has $3\boldsymbol{u} - 2\boldsymbol{v} + \boldsymbol{w} = 3(1, 3) - 2(2, 7) + 1(1, 5) = (0, 0)$. It also has $-2\boldsymbol{u} + 1\boldsymbol{v} + 0\boldsymbol{w} = \boldsymbol{b} = (0, 1)$. Adding gives $\boldsymbol{u} - \boldsymbol{v} + \boldsymbol{w} = (0, 1)$. In this case c, d, e equal $3, -2, 1$ and $C, D, E = -2, 1, 0$.

Could another example have $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ that could NOT combine to produce $\boldsymbol{b}$? Yes. The vectors $(1, 1), (2, 2), (3, 3)$ are on a line and no combination produces $\boldsymbol{b}$. We can easily solve $c\boldsymbol{u} + d\boldsymbol{v} + e\boldsymbol{w} = 0$ but not $C\boldsymbol{u} + D\boldsymbol{v} + E\boldsymbol{w} = \boldsymbol{b}$.

26 The combinations of $\boldsymbol{v}$ and $\boldsymbol{w}$ fill the plane *unless $\boldsymbol{v}$ and $\boldsymbol{w}$ lie on the same line through* $(0, 0)$. Four vectors whose combinations fill 4-dimensional space: one example is the "standard basis" $(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0)$, and $(0, 0, 0, 1)$.

27 The equations $c\boldsymbol{u} + d\boldsymbol{v} + e\boldsymbol{w} = \boldsymbol{b}$ are

$$
\begin{array}{llll}
2c \quad -d \qquad\ = 1 & \text{So } d = 2e & c = 3/4 \\
-c +2d \quad -e = 0 & \text{then } c = 3e & d = 2/4 \\
\quad\ \ -d +2e = 0 & \text{then } 4e = 1 & e = 1/4
\end{array}
$$

Problem Set 1.2, page 15

1 $u \cdot v = -2.4 + 2.4 = 0$, $u \cdot w = -.6 + 1.6 = 1$, $u \cdot (v + w) = u \cdot v + u \cdot w = 0 + 1$, $w \cdot v = 4 + 6 = 10 = v \cdot w$.

2 The lengths are $\|u\| = 1$ and $\|v\| = 5$ and $\|w\| = \sqrt{5}$. Then $|u \cdot v| = 0 < (1)(5)$ and $|v \cdot w| = 10 < 5\sqrt{5}$, confirming the Schwarz inequality.

3 Unit vectors $v/\|v\| = (\frac{4}{5}, \frac{3}{5}) = (0.8, 0.6)$ and $w/\|w\| = (1/\sqrt{5}, 2/\sqrt{5})$. The vectors $w, (2, -1)$, and $-w$ make $0°, 90°, 180°$ angles with w. The cosine of θ is $\frac{v}{\|v\|} \cdot \frac{w}{\|w\|} = 10/5\sqrt{5} = 2/\sqrt{5}$.

4 For unit vectors u, v, w: (a) $v \cdot (-v) = -1$ (b) $(v + w) \cdot (v - w) = v \cdot v + w \cdot v - v \cdot w - w \cdot w = 1 + (\ ) - (\ ) - 1 = 0$ so $\theta = 90°$ (notice $v \cdot w = w \cdot v$) (c) $(v - 2w) \cdot (v + 2w) = v \cdot v - 4w \cdot w = 1 - 4 = -3$.

5 $u_1 = v/\|v\| = (1, 3)/\sqrt{10}$ and $u_2 = w/\|w\| = (2, 1, 2)/3$. $U_1 = (3, -1)/\sqrt{10}$ is perpendicular to u_1 (and so is $(-3, 1)/\sqrt{10}$). U_2 could be $(1, -2, 0)/\sqrt{5}$: There is a whole plane of vectors perpendicular to u_2, and a whole circle of unit vectors in that plane.

6 All vectors $w = (c, 2c)$ are perpendicular to $v = (2, -1)$. They lie on a line. All vectors (x, y, z) with $x + y + z = 0$ lie on a *plane*. All vectors perpendicular to both $(1, 1, 1)$ and $(1, 2, 3)$ lie on a *line* in 3-dimensional space.

7 (a) $\cos \theta = v \cdot w/\|v\|\|w\| = 1/(2)(1)$ so $\theta = 60°$ or $\pi/3$ radians   (b) $\cos \theta = 0$ so $\theta = 90°$ or $\pi/2$ radians   (c) $\cos \theta = 2/(2)(2) = 1/2$ so $\theta = 60°$ or $\pi/3$ (d) $\cos \theta = -5/\sqrt{10}\sqrt{5} = -1/\sqrt{2}$ so $\theta = 135°$ or $3\pi/4$ radians.

8 (a) False: v and w are any vectors in the plane perpendicular to u (b) True: $u \cdot (v + 2w) = u \cdot v + 2u \cdot w = 0$ (c) True, $\|u - v\|^2 = (u - v) \cdot (u - v)$ splits into $u \cdot u + v \cdot v = 2$ when $u \cdot v = v \cdot u = 0$.

9 If $v_2 w_2/v_1 w_1 = -1$ then $v_2 w_2 = -v_1 w_1$ or $v_1 w_1 + v_2 w_2 = v \cdot w = 0$: perpendicular! The vectors $(1, 4)$ and $(1, -\frac{1}{4})$ are perpendicular because $1 - 1 = 0$.

10 Slopes $2/1$ and $-1/2$ multiply to give -1. Then $\boldsymbol{v} \cdot \boldsymbol{w} = 0$ and the two vectors (the arrow directions) are perpendicular.

11 $\boldsymbol{v} \cdot \boldsymbol{w} < 0$ means angle $> 90°$; these $\boldsymbol{w}$'s fill half of 3-dimensional space. Draw a picture to show $\boldsymbol{v}$ and the $\boldsymbol{w}$'s.

12 $(1, 1)$ is perpendicular to $(1, 5) - c(1, 1)$ if $(1, 1) \cdot (1, 5) - c(1, 1) \cdot (1, 1) = 6 - 2c = 0$ (then $\boldsymbol{c = 3}$). $\boldsymbol{v} \cdot (\boldsymbol{w} - c\boldsymbol{v}) = 0$ if $c = \boldsymbol{v} \cdot \boldsymbol{w}/\boldsymbol{v} \cdot \boldsymbol{v}$. Subtracting $c\boldsymbol{v}$ is the key to constructing a perpendicular vector $\boldsymbol{w} - c\boldsymbol{v}$.

13 One possibility among many: $\boldsymbol{u} = (1, -1, 0, 0), \boldsymbol{v} = (0, 0, 1, -1), \boldsymbol{w} = (1, 1, -1, -1)$ and $(1, 1, 1, 1)$ are perpendicular to each other. "We can rotate those $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ in their 3D hyperplane and they will stay perpendicular."

14 $\frac{1}{2}(x + y) = (2 + 8)/2 = \boldsymbol{5}$ and $\boldsymbol{5} > 4$; $\cos \theta = 2\sqrt{16}/\sqrt{10}\sqrt{10} = \boldsymbol{8/10}$.

15 $\|\boldsymbol{v}\|^2 = 1 + 1 + \cdots + 1 = 9$ so $\|\boldsymbol{v}\| = \boldsymbol{3}$; $\boldsymbol{u} = \boldsymbol{v}/3 = (\frac{1}{3}, \ldots, \frac{1}{3})$ is a unit vector in 9D; $\boldsymbol{w} = (1, -1, 0, \ldots, 0)/\sqrt{2}$ is a unit vector in the 8D hyperplane perpendicular to $\boldsymbol{v}$.

16 $\cos \alpha = 1/\sqrt{2}$, $\cos \beta = 0$, $\cos \gamma = -1/\sqrt{2}$. For any vector $\boldsymbol{v} = (\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$ the cosines with the 3 axes are $\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = (v_1^2 + v_2^2 + v_3^2)/\|\boldsymbol{v}\|^2 = 1$.

17 $\|\boldsymbol{v}\|^2 = 4^2 + 2^2 = 20$ and $\|\boldsymbol{w}\|^2 = (-1)^2 + 2^2 = 5$. Pythagoras is $\|(3, 4)\|^2 = 25 = 20 + 5$ for the length of the hypotenuse $\boldsymbol{v} + \boldsymbol{w} = (3, 4)$.

18 $\|\boldsymbol{v} + \boldsymbol{w}\|^2 = (\boldsymbol{v} + \boldsymbol{w}) \cdot (\boldsymbol{v} + \boldsymbol{w}) = \boldsymbol{v} \cdot (\boldsymbol{v} + \boldsymbol{w}) + \boldsymbol{w} \cdot (\boldsymbol{v} + \boldsymbol{w})$. This expands to $\boldsymbol{v} \cdot \boldsymbol{v} + 2\boldsymbol{v} \cdot \boldsymbol{w} + \boldsymbol{w} \cdot \boldsymbol{w} = \|\boldsymbol{v}\|^2 + 2\|\boldsymbol{v}\| \, \|\boldsymbol{w}\| \cos \theta + \|\boldsymbol{w}\|^2$.

19 We know that $(\boldsymbol{v} - \boldsymbol{w}) \cdot (\boldsymbol{v} - \boldsymbol{w}) = \boldsymbol{v} \cdot \boldsymbol{v} - 2\boldsymbol{v} \cdot \boldsymbol{w} + \boldsymbol{w} \cdot \boldsymbol{w}$. The Law of Cosines writes $\|\boldsymbol{v}\|\|\boldsymbol{w}\| \cos \theta$ for $\boldsymbol{v} \cdot \boldsymbol{w}$. Here θ is the angle between $\boldsymbol{v}$ and $\boldsymbol{w}$. When $\theta < 90°$ this $\boldsymbol{v} \cdot \boldsymbol{w}$ is positive, so in this case $\boldsymbol{v} \cdot \boldsymbol{v} + \boldsymbol{w} \cdot \boldsymbol{w}$ is larger than $\|\boldsymbol{v} - \boldsymbol{w}\|^2$.

Pythagoras changes from equality $a^2 + b^2 = c^2$ to *inequality* when $\theta < 90°$ or $\theta > 90°$.

20 $2\boldsymbol{v} \cdot \boldsymbol{w} \le 2\|\boldsymbol{v}\|\|\boldsymbol{w}\|$ leads to $\|\boldsymbol{v} + \boldsymbol{w}\|^2 = \boldsymbol{v} \cdot \boldsymbol{v} + 2\boldsymbol{v} \cdot \boldsymbol{w} + \boldsymbol{w} \cdot \boldsymbol{w} \le \|\boldsymbol{v}\|^2 + 2\|\boldsymbol{v}\|\|\boldsymbol{w}\| + \|\boldsymbol{w}\|^2$. This is $(\|\boldsymbol{v}\| + \|\boldsymbol{w}\|)^2$. Taking square roots gives $\|\boldsymbol{v} + \boldsymbol{w}\| \le \|\boldsymbol{v}\| + \|\boldsymbol{w}\|$.

21 $v_1^2 w_1^2 + 2v_1 w_1 v_2 w_2 + v_2^2 w_2^2 \le v_1^2 w_1^2 + v_1^2 w_2^2 + v_2^2 w_1^2 + v_2^2 w_2^2$ is true (cancel 4 terms) because the difference is $v_1^2 w_2^2 + v_2^2 w_1^2 - 2v_1 w_1 v_2 w_2$ which is $(v_1 w_2 - v_2 w_1)^2 \ge 0$.

22 Example 6 gives $|u_1||U_1| \le \frac{1}{2}(u_1^2 + U_1^2)$ and $|u_2||U_2| \le \frac{1}{2}(u_2^2 + U_2^2)$. The whole line

becomes $.96 \le (.6)(.8) + (.8)(.6) \le \frac{1}{2}(.6^2 + .8^2) + \frac{1}{2}(.8^2 + .6^2) = 1$. True : $.96 < 1$.

23 The cosine of θ is $x/\sqrt{x^2 + y^2}$, near side over hypotenuse. Then $|\cos\theta|^2$ is not greater

than $1 : x^2/(x^2 + y^2) \le 1$.

24 These two lines add to $2||\boldsymbol{v}||^2 + 2||\boldsymbol{w}||^2$:

$$||\boldsymbol{v} + \boldsymbol{w}||^2 = (\boldsymbol{v} + \boldsymbol{w}) \cdot (\boldsymbol{v} + \boldsymbol{w}) = \boldsymbol{v} \cdot \boldsymbol{v} + \boldsymbol{v} \cdot \boldsymbol{w} + \boldsymbol{w} \cdot \boldsymbol{v} + \boldsymbol{w} \cdot \boldsymbol{w}$$

$$||\boldsymbol{v} - \boldsymbol{w}||^2 = (\boldsymbol{v} - \boldsymbol{w}) \cdot (\boldsymbol{v} - \boldsymbol{w}) = \boldsymbol{v} \cdot \boldsymbol{v} - \boldsymbol{v} \cdot \boldsymbol{w} - \boldsymbol{w} \cdot \boldsymbol{v} + \boldsymbol{w} \cdot \boldsymbol{w}$$

25 The length $||\boldsymbol{v} - \boldsymbol{w}||$ is between 2 and 8 (triangle inequality when $||\boldsymbol{v}|| = 5$ and $||\boldsymbol{w}|| = 3$). The dot product $\boldsymbol{v} \cdot \boldsymbol{w}$ is between -15 and 15 by the Schwarz inequality.

26 Three vectors in the plane could make angles greater than $90°$ with each other: for

example $(1, 0), (-1, 4), (-1, -4)$. Four vectors could *not* do this ($360°$ total angle).

How many can can be perpendicular to each other in $\mathbf{R}^3$ or $\mathbf{R}^n$? Ben Harris and Greg

Marks showed me that the answer is $n + 1$. The vectors from the center of a regular

simplex in $\mathbf{R}^n$ to its $n+1$ vertices all have negative dot products. If $n+2$ vectors in $\mathbf{R}^n$

had negative dot products, project them onto the plane orthogonal to the last one. Now

you have $n + 1$ vectors in $\mathbf{R}^{n-1}$ with negative dot products. Keep going to 4 vectors in

$\mathbf{R}^2$: no way!

27 The columns of the 4 by 4 "Hadamard matrix" (times $\frac{1}{2}$) are perpendicular unit

vectors:

$$\frac{1}{2}H = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

The columns have

$\frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = 1$.

Their dot products

are all zero.

28 The commands $V = \mathbf{randn}\,(3, 30); D = \mathbf{sqrt}\,(\mathbf{diag}\,(V\,' * V)); U = V \backslash D$; will give

30 random unit vectors in the columns of U. Then $\boldsymbol{u}\,' * U$ is a row matrix of 30 dot

products whose average absolute value should be close to $2/\pi$.

29 The four vectors v_1, v_2, v_3, v_4 must add to zero. Then the four corners of the quadri-lateral could be 0 and v_1 and $v_1 + v_2$ and $v_1 + v_2 + v_3$. We are allowing the side vectors v to cross each other—can you answer if that is not allowed ?

Problem Set 1.3, page 24

1 The column space $\mathbf{C}(A_1)$ is a plane in $\mathbf{R}^3$: the two columns of A_1 are independent

The column space $\mathbf{C}(A_2)$ is all of $\mathbf{R}^3$

The column space $\mathbf{C}(A_3)$ is a line in $\mathbf{R}^3$

2 The combination $A\boldsymbol{x} = \text{column } 1 - 2 \,(\text{column } 2) + \text{column } 3$ is zero for both matrices. This leaves 2 independent columns. So $\mathbf{C}(A)$ is a (2-dimensional) plane in $\mathbf{R}^3$.

3 B has 2 independent columns so its column space is a plane. The matrix C has the same 2 independent columns and the same column space as B.

4 $A\boldsymbol{x} = \begin{bmatrix} 14 \\ 28 \\ 2 \end{bmatrix}$ Typical dot product is $2(1) + 1(2) + 2(5) = 14$ $B\boldsymbol{y} = \begin{bmatrix} 4 \\ 8 \\ 18 \end{bmatrix}$ $I\boldsymbol{z} = \boldsymbol{z} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$

5 $A\boldsymbol{x} = 1 \begin{bmatrix} 2 \\ 4 \\ 0 \end{bmatrix} + 2 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + 5 \begin{bmatrix} 2 \\ 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 14 \\ 28 \\ 2 \end{bmatrix}$

$B\boldsymbol{y} = 4 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + 4 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + 10 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \\ 18 \end{bmatrix}$

$I\boldsymbol{z} = z_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + z_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + z_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$

6 A has **2** independent columns, B has **3**, and $A + B$ has **3**. These are the ranks of A and B and $A + B$. The rule is that $\text{rank}(A + B) \leq \text{rank}(A) + \text{rank}(B)$.

7 (a) $A = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$ $B = \begin{bmatrix} 3 & 1 \\ 4 & 2 \end{bmatrix}$ $A + B = \begin{bmatrix} 4 & 4 \\ 6 & 6 \end{bmatrix} = \text{rank } 1$

(b) $A = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$ $B = \begin{bmatrix} -1 & -3 \\ -2 & -4 \end{bmatrix}$ $A + B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \text{rank } 0$

$$(c)\ A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad A + B = I = \text{rank } 4$$

8 The column space of A is all of $\mathbf{R}^3$. The column space of B is a **line** in $\mathbf{R}^3$. The column space of C is a 2-dimensional plane in $\mathbf{R}^3$. If C had an additional row of zeros, its column space would be a 2-dimensional plane **in $\mathbf{R}^4$**.

9 $A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 1 \\ 1 & 2 & 1 \end{bmatrix}$ **Seven ones** is the maximum for rank 3. With eight ones, two columns will be equal

10 $A = \begin{bmatrix} 3 & 9 \\ 5 & 15 \end{bmatrix}$ has rank 1 : 1 independent column,

                                        1 independent row

      $B = \begin{bmatrix} 1 & 2 & -5 \\ 4 & 8 & -20 \end{bmatrix}$ has 1 independent column in $\mathbf{R}^2$,

                                        1 independent row in $\mathbf{R}^3$

11 (a) If B has an extra zero column, A and B have the **same** column space. Different row spaces because of different row lengths !

(b) If column $3 =$ column $2 -$ column 1, A and B have the same column spaces.

(c) If the new column 3 in B is $(1, 1, 1)$, then the column space is not changed or changed depending whether $(1, 1, 1)$ was already in $\mathbf{C}(A)$.

12 If $\boldsymbol{b}$ is in the column space of A, then $\boldsymbol{b}$ is a combination of the columns of A and *the numbers in that combination* give a solution $\boldsymbol{x}$ to $A\boldsymbol{x} = \boldsymbol{b}$. The examples are solved by $(x_1, x_2) = (1, 1)$ and $(1, -1)$ and $\left(-\frac{1}{2}, \frac{1}{2}\right)$.

13 $A = \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{bmatrix}$ $B = \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ -1 & -2 \end{bmatrix}$ $A + B = \begin{bmatrix} 2 & 0 \\ -1 & 3 \\ -1 & -3 \end{bmatrix}$ has the

same column space as A and B (other examples could have a smaller column space : for example if $B = -A$ in which case $A + B =$ zero matrix).

14 $A = \begin{bmatrix} 1 & 0 & 2 \\ 3 & 1 & 9 \\ 5 & 0 & \mathbf{10} \end{bmatrix}$ has column $3 = 2$ (column 1) $+\, 3$ (column 2)

$A = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & \mathbf{9} \end{bmatrix}$ has column $3 = -1$ (column 1) $+\, 2$ (column 2)

$A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 2 & 4 \\ 0 & 0 & \boldsymbol{q} \end{bmatrix}$ has 2 independent columns if $\boldsymbol{q} \neq \mathbf{0}$

15 If $A\boldsymbol{x} = \boldsymbol{b}$ then the extra column $\boldsymbol{b}$ in $\begin{bmatrix} A & \boldsymbol{b} \end{bmatrix}$ is a combination of the first columns, so the column space and the rank are not changed by including the $\boldsymbol{b}$ column.

16 (a) *False* : B could be $-A$, then $A + B$ has rank zero.

(b) *True* : If the n columns of A are independent, they could not be in a space $\mathbf{R}^m$ with $m < n$. Therefore $m \geq n$.

(c) *True* : If the entries are random and the matrix has $m = n$ (or $m \geq n$), then the columns are almost surely independent.

17 rank 2 : $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ **rank 1** : $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

rank 0 : $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

18 $3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + 4 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + 5 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 7 \\ 12 \end{bmatrix} = S\boldsymbol{x} = \boldsymbol{b}$

$S = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$ and the 3 dot products in $S\boldsymbol{x}$ are $3, 7, 12$

19 Suppose $a = mc$ and $b = md$ (all nonzero). Then $amd = bmc$. Then $a/b = c/d$.

If those ratios are M, then $(a, c) = M(b, d)$.

20 $Sy = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$ is solved by $y = \begin{bmatrix} c_1 \\ c_2 - c_1 \\ c_3 - c_2 \end{bmatrix}$. This is

$y = S^{-1}c = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$. S is square with independent columns. So S

has an inverse with $SS^{-1} = S^{-1}S = I$.

21 To solve $Ax = 0$ we can simplify the 3 equations (this is the subject of Chapter 2).

$$
\begin{array}{l}
\text{Start from } Ax = 0
\end{array}
\quad
\begin{array}{l}
x_1 + 2x_2 + 3x_3 = 0 \\
3x_1 + 5x_2 + 6x_3 = 0 \\
4x_1 + 7x_2 + 9x_3 = 0
\end{array}
\quad
\begin{array}{c}
\text{Row } 2 - 3(\text{row } 1) \\
\text{row } 3 - 4(\text{row } 1)
\end{array}
\quad
\begin{array}{l}
x_1 + 2x_2 + 3x_3 = 0 \\
 - x_2 - 3x_3 = 0 \\
 - x_2 - 3x_3 = 0
\end{array}
$$

If $x_3 = 1$ then $x_2 = -3$ and $x_1 = 3$. Any answer $x = (3c, -3c, c)$ is correct.

22 $\begin{bmatrix} 1 & 1 & 0 \\ 3 & 2 & 1 \\ 7 & 4 & c = \mathbf{3} \end{bmatrix}$ $\begin{bmatrix} 1 & 0 & c = -\mathbf{1} \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ $\begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ 2 & 1 & 5 \\ 3 & 3 & 6 \end{bmatrix}$ $\begin{bmatrix} \mathbf{2} & 1 \\ 4 & \mathbf{2} \\ -2 & 1 \\ 4 & -2 \end{bmatrix}$ $\begin{array}{l} \text{have} \\ \text{dependent} \\ \text{columns} \end{array}$

23 The equation $Ax = 0$ says that x is perpendicular to each row of A (three dot products are zero). So x is perpendicular to all combinations of those rows. In other words, x is perpendicular to the row space (here a plane).

An important fact for linear algebra: Every x in the nullspace of A (meaning $Ax = 0$) is perpendicular to every vector in the row space.

Problem Set 1.4, page 35

1 Here are the 4 ways to multiply AB and the operation counts. A is m by n, B is n by p.

Row i times column k mp dot products, n multiplications each

Matrix A times column k p columns, mn multiplications each

Row i times matrix B m rows, np multiplications each

Column j of A times row j of B n (columns)(rows), mp multiplications each

2 $A = \begin{bmatrix} a & a & a \end{bmatrix}$ factors into $CR = \begin{bmatrix} a \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$

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

$\begin{bmatrix} 1 & 2 & 3 \end{bmatrix}\begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 32 \end{bmatrix}$ $\begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 4 & 8 & 12 \\ 5 & 10 & 15 \\ 6 & 12 & 18 \end{bmatrix}$

4 (a) $\begin{bmatrix} 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} = 2\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$ $= \begin{bmatrix} 2 & 2 & 2 \end{bmatrix}$

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

(b) $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 5 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 9 \\ 0 & 1 \end{bmatrix}$

$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 7 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 9 \\ 0 & 1 \end{bmatrix}$

5 A has 7 columns and 4 rows. Those columns are vectors in 4-dimensional space. We cannot have 5 independent column vectors because we cannot have 5 independent vectors in 4-dimensional space. (This is really just a restatement of the problem. The proof

comes in Section 3.2: Every m by n matrix C, with $m < n$ has a nonzero solution to

$C\boldsymbol{x} = \boldsymbol{0}$. Here $m = 4$ and $n = 5$ and 5 columns of C cannot be independent.)

6 $A = \begin{bmatrix} 2 & -2 & 1 & 6 & 0 \\ 1 & -1 & 0 & 2 & 0 \\ 3 & -3 & 0 & 6 & 1 \end{bmatrix}$ $\qquad C = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 0 \\ 3 & 0 & 1 \end{bmatrix}$

7 $CR = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 0 \\ 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} = A$ in Problem 6.

8 $A = \begin{bmatrix} 2 & 2 & 2 \\ 0 & 4 & 4 \\ 0 & 0 & 6 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 2 \\ 0 & 4 & 4 \\ 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} = AI$ $\begin{array}{l} A = C \\ \text{and} \\ R = I \end{array}$

$B = \begin{bmatrix} 2 & 2 & 2 \\ 0 & 0 & 4 \\ 0 & 0 & 6 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 0 & 4 \\ 0 & 6 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = CR$

9 A random 4 by 4 matrix has independent columns ($C = A$ and $R = I$) with probability 1.
(We could be choosing the 16 entries of A between 0 and 1 with uniform probability
by $A = \mathbf{rand}(4, 4)$. We could be choosing those 16 entries of A from a "bell-shaped"
normal distribution by $A = \mathbf{rand}(4, 4)$. If we were choosing those 16 entries from
a finite list of numbers, then there is a nonzero probability that the columns of A are
dependent. In fact a nonzero probability that all 16 numbers are the same.)

10 If A is a random 4 by 5 matrix, then (using **rand** or **randn** as above) with probability 1
the first 4 columns are independent and go into C. With probability zero (this does not
mean it can't happen !) the first 4 columns will be dependent and C will be different
(C will have r columns with $r \le 4$).

11 $A = \begin{bmatrix} 1 & 0 & a & c \\ 0 & 1 & b & d \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & a & c \\ 0 & 1 & b & d \end{bmatrix} = CR$. Many other possibilities !

12 $A_1 = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ $A_2 = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$

$A_3 = \begin{bmatrix} 2 \\ 6 \end{bmatrix} \begin{bmatrix} 1 & 0.5 & 1.5 \end{bmatrix}$ $A_4 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 1 & 0 \end{bmatrix}$

13 $C = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ and $R = \begin{bmatrix} 2 & 4 \end{bmatrix}$ have $CR = \begin{bmatrix} 2 & 4 \\ 6 & 12 \end{bmatrix}$ and $RC = \begin{bmatrix} 14 \end{bmatrix}$

and $CRC = \begin{bmatrix} 14 \\ 42 \end{bmatrix}$ and $RCR = \begin{bmatrix} 28 & 56 \end{bmatrix}$.

Here is an interesting fact when A is m by n and B is n by m. The m numbers on the main diagonal of AB have the same total as the n numbers on the main diagonal of BA. Example:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 3 \\ 1 & 4 \\ 2 & 5 \end{bmatrix} \quad AB = \begin{bmatrix} 8 & 26 \\ 17 & 62 \end{bmatrix} \quad BA = \begin{bmatrix} 12 & 15 & 18 \\ 17 & 22 & 27 \\ 22 & 29 & 36 \end{bmatrix}$$

$8 + 62 = 12 + 22 + 36$

14 $\begin{bmatrix} 3 & 6 \\ 5 & 10 \end{bmatrix}$ $\begin{bmatrix} 6 & -7 \\ 7 & 6 \end{bmatrix}$ $\begin{bmatrix} 2 & 0 \\ 3 & 6 \end{bmatrix}$ $\begin{bmatrix} 3 & 4 \\ -2 & -3 \end{bmatrix}$

rank one   orthogonal columns   rank 2 $A^2 = I$

15   1. Column j of A equals the matrix C times column j of R.

This is a combination of the **columns** of C.

2. Row i of A is row i of C times the matrix R.

This is a combination of the **rows** of R.

3. (row i of C) $\cdot$ (column j of R) gives A_{ij}

That dot product requires the number of columns of C to equal the number of rows of R.

4. C has r columns so R has r rows (to multiply CR). Those columns of C are independent (by construction). Those rows of R are independent (because R contains the r by r identity matrix).

16 (a) The vector $AB\boldsymbol{x}$ is the matrix A times the vector $B\boldsymbol{x}$. So it is a combination of the columns of A. Therefore $\mathbf{C}(AB) \subseteq \mathbf{C}(A)$.

(b) $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ $B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ give $AB =$ zero matrix and $\mathbf{C}(AB) =$ zero vectors.

17 (a) If A and B have rank 1, then AB has rank 1 or 0. $A = \boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ and $B = \boldsymbol{x}\boldsymbol{y}^{\mathrm{T}}$ give $AB = \boldsymbol{u}(\boldsymbol{v}^{\mathrm{T}}\boldsymbol{x})\boldsymbol{y}^{\mathrm{T}}$ so $AB =$ zero matrix if the dot product $\boldsymbol{v}^{\mathrm{T}}\boldsymbol{x}$ happens to be zero.

(b) If A and B are 3 by 3 matrices of rank 3, then it is **true** that AB has rank 3. *One approach*: If $AB\boldsymbol{x} = \mathbf{0}$ then $B\boldsymbol{x} = \mathbf{0}$ because A has 3 independent columns. But $B\boldsymbol{x} = \mathbf{0}$ only when $\boldsymbol{x} = \mathbf{0}$, because B has 3 independent columns.

(c) Suppose $AB = BA$ for all 2 by 2 matrices B. Choose $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ so that

$$AB = \begin{bmatrix} c & d \\ e & f \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} c & d \\ e & f \end{bmatrix}.$$ This tells us that $\begin{bmatrix} c & 0 \\ e & 0 \end{bmatrix} = \begin{bmatrix} c & d \\ 0 & 0 \end{bmatrix}$

and therefore $d = e = 0$. Now choose $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ so that $AB = \begin{bmatrix} c & 0 \\ 0 & f \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$

$$= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} c & 0 \\ 0 & f \end{bmatrix}.$$ This tells us that $\begin{bmatrix} 0 & c \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & f \\ 0 & 0 \end{bmatrix}$ and $c = f$ and $A = cI$.

18 (a) $AB = \begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix}$ and $BC = \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}$.

(b) $(AB)C =$ column exchange of $AB = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}$

$A(BC) =$ row exchange of $BC = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix} =$ same result ABC.

19 $AB = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \end{bmatrix} +$

$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$

$= \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{1} & \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{2} & \mathbf{2} \\ \mathbf{1} & \mathbf{2} & \mathbf{3} \end{bmatrix}$

$BA = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{3} & \mathbf{2} & \mathbf{1} \\ \mathbf{2} & \mathbf{2} & \mathbf{1} \\ \mathbf{1} & \mathbf{1} & \mathbf{1} \end{bmatrix}$

20 $AB = (4 \times 3)(3 \times 2)$ needs $mnp = (4)(3)(2) = 24$ multiples.

Then $(AB)C = (4 \times 2)(2 \times 1)$ needs $(4)(2)(1) = 8$ more : TOTAL 32.

$BC = (3 \times 2)(2 \times 1)$ needs $mnp = (3)(2)(1) = 6$ multiplies.

Then $A(BC) = (4 \times 3)(3 \times 1)$ needs $(4)(3)(1) = 12$ more : TOTAL 18.

Best to start with $C = $ vector. Multiply by B to get the vector BC, and then the vector $A(BC)$. Vectors need less computing time than matrices !

Problem Set 2.1, page 46

1 Multiply equation 1 by $\ell_{21} = \frac{10}{2} = \mathbf{5}$ and subtract from equation 2 to find $2x + 3y = 1$ (unchanged) and $-6y = 6$. The pivots to circle are 2 and -6. Back substitution in $-6y = 6$ gives $y = -1$. Then $2x + 3y = 1$ gives $x = 2$.

2 The row picture and column picture and coefficient matrix are changed. The solution has not changed.

3 Subtract $-\frac{1}{2}$ (or add $\frac{1}{2}$) times equation 1. The new second equation is $\mathbf{3y = 3}$. Then $y = 1$ and $x = 5$. If the right sides change sign, so does the solution: $(x, y) = (-5, -1)$.

4 Subtract $\ell = \frac{c}{a}$ times equation 1 from equation 2. The new second pivot multiplying y is $d - (cb/a)$ or $(ad - bc)/a$. Then $y = (ag - cf)/(ad - bc)$. Notice the "determinant of A" $= ad - bc$. It must be nonzero for this division.

5 $6x + 4y$ is 2 times $3x + 2y$. There is no solution unless the right side is $2 \cdot 10 = 20$. Then all the points on the line $3x + 2y = 10$ are solutions, including $(0, 5)$ and $(4, -1)$. The two lines in the row picture are the same line, containing all solutions.

6 Singular system if $b = 4$, because $4x + 8y$ is 2 times $2x + 4y$. Then $g = 32$ makes the lines $2x + 4y = 16$ and $4x + 8y = 32$ become the *same*: infinitely many solutions like $(8, 0)$ and $(0, 4)$.

7 If $a = 2$ elimination must fail (two parallel lines in the row picture). The equations have no solution. With $a = 0$, elimination will stop for a row exchange. Then $3y = -3$ gives $y = -1$ and $4x + 6y = 6$ gives $x = 3$.

8 If $k = 3$ elimination must fail: no solution. If $k = -3$, elimination gives $0 = 0$ in equation 2: infinitely many solutions. If $k = 0$ a row exchange is needed: one solution.

9 On the left side, $6x - 4y$ is 2 times $(3x - 2y)$. Therefore we need $b_2 = 2b_1$ on the right side. Then there will be infinitely many solutions (two parallel lines become one single line in the row picture). The column picture has both columns along the same line.

10 The equation $y = 1$ comes from elimination (subtract $x + y = 5$ from $x + 2y = 6$). Then $x = 4$ and $5x - 4y = 20 - 4 = c = 16$.

11 (a) Another solution is $\frac{1}{2}(x + X, y + Y, z + Z)$.   (b) If 25 planes meet at two points, they meet along the whole line through those two points.

12 Elimination leads to an upper triangular system; then comes back substitution.

$$2x + 3y + z = 8 \qquad x = 2$$
$$y + 3z = 4 \text{ gives } y = 1 \quad \text{If a zero is at the start of row 2 or row 3,}$$
$$8z = 8 \qquad z = 1 \quad \text{that avoids a row operation.}$$

$$2x - 3y = 3 \qquad\qquad 2x - 3y = 3 \qquad\qquad 2x - 3y = 3 \qquad x = 3$$
$$4x - 5y + z = 7 \quad \text{gives} \quad y + z = 1 \quad \text{and} \quad y + z = 1 \quad \text{and} \quad y = 1$$
$$2x - y - 3z = 5 \qquad\qquad 2y + 3z = 2 \qquad\qquad -5z = 0 \qquad z = 0$$

13 Subtract 2 times row 1 from row 2 to reach $(d - 10)y - z = 2$ along with $y - z = 3$. If $d = \mathbf{10}$ exchange rows 2 and 3. If $d = \mathbf{11}$ the system becomes singular.

14 The second pivot position will contain $-2 - b$. If $b = -2$ we exchange with row 3. If $b = -1$ (singular case) the second equation is $-y - z = 0$. But equation (3) is the same so there is a *line of solutions* $(x, y, z) = (1, 1, -1)$ when $b = -1$.

15 (a) **Example of 2 exchanges**

$$0x + 0y + 2z = 4$$
$$x + 2y + 2z = 5$$
$$0x + 3y + 4z = 6$$

(exchange 1 and 2, then 2 and 3)

(b) **Exchange but then breakdown**

$$0x + 3y + 4z = 4$$
$$x + 2y + 2z = 5$$
$$0x + 3y + 4z = 6$$

(rows 1 and 3 are not consistent)

16 If row $1 = $ row 2, then row 2 is zero after the first step; exchange the zero row with row 3. The new row 3 has no pivot. If column $2 = $ column 1, then column 2 has no pivot.

17 *Example* $x + 2y + 3z = 0$, $4x + 8y + 12z = 0$, $5x + 10y + 15z = 0$ has 9 different coefficients but rows 2 and 3 become $0 = 0$: infinitely many solutions to $A\boldsymbol{x} = \mathbf{0}$ but almost surely no solution to $A\boldsymbol{x} = \boldsymbol{b}$ for a random $\boldsymbol{b}$.

18 Row 2 becomes $3y - 4z = 5$, then row 3 becomes $(q + 4)z = t - 5$. If $q = -4$ the system is singular—no third pivot. Then if $t = 5$ the third equation is $0 = 0$ which allows infinitely many solutions. Choosing $z = 1$ the equation $3y - 4z = 5$ gives $y = 3$ and equation 1 gives $x = -9$.

19 Elimination fails on $\begin{bmatrix} a & 2 \\ a & a \end{bmatrix}$ if $a = \mathbf{2}$ or $a = \mathbf{0}$. (You could notice that the determinant $a^2 - 2a$ is zero for $a = 2$ and $a = 0$.)

20 $a = 2$ gives equal columns, $a = 4$ gives equal rows, $a = 0$ gives a zero column.

21 Solvable for $s = 10$ (add the two pairs of equations to get $a + b + c + d$ on the left sides, 12 and $2 + s$ on the right sides). So 12 must agree with $2 + s$, which makes $s = 10$. The four equations for a, b, c, d are **singular**! Two solutions are $\begin{bmatrix} 1 & 3 \\ 1 & 7 \end{bmatrix}$ and $\begin{bmatrix} 0 & 4 \\ 2 & 6 \end{bmatrix}$,

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} 4 \\ 2 \\ 8 \\ s \end{bmatrix} \text{ and } U = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

22 $A(2, :) = A(2, :) - 3 * A(1, :)$ subtracts 3 times all of row 1 from all of row 2.

23 The average pivots for rand(3) *without* row exchanges were $\frac{1}{2}, 5, 10$ in one experiment—but pivots 2 and 3 can be arbitrarily large. Their averages are actually infinite ! *With row exchanges* in MATLAB's **lu** code, the averages .75 and .50 and .365 are much more stable (and should be predictable, also for **randn** with normal instead of uniform probability distribution for the numbers in A).

24 If $A(5, 5)$ is 7 not 11, then the last pivot will be 0 not 4.

25 Row j of U is a combination of rows $1, \ldots, j$ of A (when there are no row exchanges). If $A\mathbf{x} = \mathbf{0}$ then $U\mathbf{x} = \mathbf{0}$ (not true if $\mathbf{b}$ replaces $\mathbf{0}$). U just keeps the diagonal of A when A is *lower triangular,* all entries below that diagonal go to zero.

26 The question deals with 100 equations $A\mathbf{x} = \mathbf{0}$ when A is singular.

(a) Some linear combination of the 100 **columns** is **the column of zeros**.

(b) A very singular matrix has all ones: $A = $ **ones** (100). A better example has 99 random rows (or the numbers $1^i, \ldots, 100^i$ in those rows). The 100th row could be the sum of the first 99 rows (or any other combination of those rows with no zeros).

(c) The row picture has 100 planes **meeting along a common line through 0**. The column picture has 100 vectors all in the same 99-dimensional hyperplane.

Problem Set 2.2, page 53

0 If columns 1 and 2 of A are exchanged then rows 1 and 2 of A^{-1} are exchanged. To keep $A^{-1}A = I$, we have to keep

$$\bigl(\text{row } i \text{ of } A^{-1}\bigr)\cdot\bigl(\text{column } i \text{ of } A\bigr)=\mathbf{1} \qquad \bigl(\text{row } i \text{ of } A^{-1}\bigr)\cdot\bigl(\text{column } j \text{ of } A\bigr)=\mathbf{0} \text{ if } i \neq j$$

1 $E_{21} = \begin{bmatrix} 1 & 0 & 0 \\ -\mathbf{5} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $E_{32} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \mathbf{7} & 1 \end{bmatrix}$, $P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & \mathbf{1} & 0 \\ 0 & 0 & \mathbf{1} \\ \mathbf{1} & 0 & 0 \end{bmatrix}$.

2 $E_{32}E_{21}\boldsymbol{b} = (1, -5, -35)$ but $E_{21}E_{32}\boldsymbol{b} = (1, -5, 0)$. When E_{32} comes first, row 3 feels no effect from row 1.

3 $\begin{bmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}$ $\begin{array}{l} \leftarrow E_{21}, E_{31} E_{32} \\ E = E_{32}E_{31}E_{21} \end{array}$ $= \begin{bmatrix} 1 & 0 & 0 \\ -\mathbf{4} & 1 & 0 \\ \mathbf{10} & -\mathbf{2} & 1 \end{bmatrix}$.

Those E's are in the right order to give $EA = U$.

$$E^{-1} = E_{21}^{-1}E_{31}^{-1}E_{32}^{-1} = L = \begin{bmatrix} 1 & 0 & 0 \\ \mathbf{4} & \mathbf{1} & 0 \\ -\mathbf{2} & \mathbf{2} & 1 \end{bmatrix}$$

4 Elimination on column 4: $\boldsymbol{b} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \overset{E_{21}}{\to} \begin{bmatrix} 1 \\ -4 \\ 0 \end{bmatrix} \overset{E_{31}}{\to} \begin{bmatrix} 1 \\ -4 \\ -4 \\ 2 \end{bmatrix} \overset{E_{32}}{\to} \begin{bmatrix} 1 \\ -4 \\ 10 \end{bmatrix}$. The

original $A\boldsymbol{x} = \boldsymbol{b} = (1, 0, 0)$ has become $U\boldsymbol{x} = \boldsymbol{c} = (\mathbf{1}, -\mathbf{4}, \mathbf{10})$. Then back substitution gives $\boldsymbol{z} = -\mathbf{5}, y = \frac{1}{2}, x = \frac{1}{2}$. This solves $A\boldsymbol{x} = (1, 0, 0)$.

5 Changing a_{33} from 7 to 11 will change the third pivot from 5 to 9. Changing a_{33} from 7 to 2 will change the pivot from 5 to *no pivot*.

6 Example: $\begin{bmatrix} 2 & 3 & 7 \\ 2 & 3 & 7 \\ 2 & 3 & 7 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 4 \end{bmatrix}$. If all columns are multiples of column 1, there

is no second pivot.

7 To reverse E_{31}, **add** 7 times row **1** to row **3**. The inverse of the elimination matrix

$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\mathbf{7} & 0 & 1 \end{bmatrix}$ is $E^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \mathbf{7} & 0 & 1 \end{bmatrix}$. Multiplication confirms $EE^{-1} = I$.

8 $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $M^* = \begin{bmatrix} a & b \\ c - \ell a & d - \ell b \end{bmatrix}$. $\det M^* = a(d - \ell b) - b(c - \ell a)$

reduces to $ad - bc$! Subtracting row 1 from row 2 doesn't change $\det M$.

9 $M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix}$ for both parts (a) and (b).

After the exchange, we need E_{31} (not E_{21}) to act on the new row 3.

10 At the same time $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$; $E_{31}E_{13} = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$. Test on the identity matrix!

11 An example with two negative pivots is $A = \begin{bmatrix} 1 & 2 & 2 \\ 1 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix}$. The diagonal entries can

change sign during elimination.

12 For the first, a simple row exchange has $P^2 = I$ so $P^{-1} = P$. For the second,

$P^{-1} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. Always $P^{-1} =$ "transpose" of P, coming in Section 2.4.

13 $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} .5 \\ -.2 \end{bmatrix}$ and $\begin{bmatrix} t \\ z \end{bmatrix} = \begin{bmatrix} -.2 \\ .1 \end{bmatrix}$ so $A^{-1} = \dfrac{1}{10} \begin{bmatrix} 5 & -2 \\ -2 & 1 \end{bmatrix}$. This question

solved $AA^{-1} = I$ column by column, the main idea of Gauss-Jordan elimination.

14 An upper triangular U with $U^2 = I$ is $U = \begin{bmatrix} 1 & a \\ 0 & -1 \end{bmatrix}$ for any a. And also $-U$.

15 (a) Multiply $AB = AC$ by A^{-1} to find $B = C$ (since A is invertible)   (b) As long as

$B - C$ has the form $\begin{bmatrix} x & y \\ -x & -y \end{bmatrix}$, we have $AB = AC$ for $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

16 (a) If $A\boldsymbol{x} = (0, 0, 1)$ then equation $1 + $ equation $2 - $ equation 3 is $0 = 1$

(b) Right sides must satisfy $b_1 + b_2 = b_3$

(c) In elimination, Row 3 becomes a row of zeros—no third pivot.

17 (a) The vector $\boldsymbol{x} = (1, 1, -1)$ solves $A\boldsymbol{x} = \boldsymbol{0}$ (b) After elimination, columns 1

and 2 end in zeros. Then so does column $3 = $ column $1 + 2$: no third pivot.

18 Yes, B is invertible (A was just multiplied by a permutation matrix P). If you exchange

rows 1 and 2 of A to reach B, you exchange **columns** 1 and 2 of A^{-1} to reach B^{-1}. In

matrix notation, $B = PA$ has $B^{-1} = A^{-1}P^{-1} = A^{-1}P$ for this P.

19 (a) If $B = -A$ then A, B can be invertible but $A + B = $ zero matrix is not invertible.

(b) $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ are both singular but $A + B = I$ is invertible.

20 Multiply $C = AB$ on the left by A^{-1} and on the right by C^{-1}. Then $\boldsymbol{A^{-1} = BC^{-1}}$.

21 $M^{-1} = C^{-1}B^{-1}A^{-1}$ so multiply on the left by C and the right by A : $B^{-1} = CM^{-1}A$.

22 $B^{-1} = A^{-1} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{-1} = A^{-1} \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$: subtract *column* 2 of A^{-1} from *column* 1.

23 If A has a column of zeros, so does BA. Then $BA = I$ is impossible. There is no A^{-1}.

24 $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} ad - bc & 0 \\ 0 & ad - bc \end{bmatrix}$.   The inverse of each matrix is
the other divided by $ad - bc$

25 $E_{32}E_{31}E_{21} = \begin{bmatrix} 1 & & \\ & 1 & \\ & -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & & \\ & 1 & \\ -1 & & 1 \end{bmatrix}\begin{bmatrix} 1 & & \\ -1 & 1 & \\ & & 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ -1 & 1 & \\ 0 & -1 & 1 \end{bmatrix} = E.$

Reverse the order and change -1 to $+1$ to get inverses $E_{21}^{-1}E_{31}^{-1}E_{32}^{-1} = \begin{bmatrix} 1 & & \\ 1 & 1 & \\ 1 & 1 & 1 \end{bmatrix} =$

$L = E^{-1}$. The off-diagonal 1's are unchanged by multiplying inverses in this order.

26 $A^2B = I$ can also be written as $A(AB) = I$. Therefore A^{-1} is AB.

27 $A * \text{ones}(4, 1) = \begin{bmatrix} 4 & 4 & 4 & 4 \end{bmatrix}^{\text{T}} - \begin{bmatrix} 4 & 4 & 4 & 4 \end{bmatrix}^{\text{T}} = \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}$ so A cannot be invertible.

28 Six of the sixteen $0 - 1$ matrices are invertible : I and P and all four with three 1's.

29 $\begin{bmatrix} 1 & 3 & 1 & 0 \\ 2 & 7 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & 1 & 0 \\ 0 & 1 & -2 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 7 & -3 \\ 0 & 1 & -2 & 1 \end{bmatrix} = \begin{bmatrix} I & A^{-1} \end{bmatrix};$

$\begin{bmatrix} 1 & 4 & 1 & 0 \\ 3 & 9 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 4 & 1 & 0 \\ 0 & -3 & -3 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -3 & 4/3 \\ 0 & 1 & 1 & -1/3 \end{bmatrix} = \begin{bmatrix} I & A^{-1} \end{bmatrix}.$

30 A can be invertible with diagonal zeros (example to find). B is singular because each row adds to zero. The all-ones vector $x = (1, 1, 1, 1)$ has $Bx = 0$.

31 $\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}^{-1} = \frac{1}{4}\begin{bmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{bmatrix}$; $B\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

so B^{-1} does not exist.

32 $\begin{bmatrix} U & I \end{bmatrix} = \begin{bmatrix} 1 & a & b & 1 & 0 & 0 \\ 0 & 1 & c & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & a & 0 & 1 & 0 & -b \\ 0 & 1 & 0 & 0 & 1 & -c \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$

$\rightarrow \begin{bmatrix} 1 & 0 & 0 & 1 & -a & ac - b \\ 0 & 1 & 0 & 0 & 1 & -c \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} I & U^{-1} \end{bmatrix}.$

33 (a) True (If A has a row of zeros, then so does every AB, and $AB = I$ is impossible).

   (b) False (the matrix of all ones is singular even with diagonal 1's).

   (c) True (the inverse of A^{-1} is A and the inverse of A^2 is $(A^{-1})^2$).

34 Elimination produces the pivots a and $a - b$ and $a - b$. $A^{-1} = \dfrac{1}{a(a-b)} \begin{bmatrix} a & 0 & -b \\ -a & a & 0 \\ 0 & -a & a \end{bmatrix}$.

   The matrix C is not invertible if $c = 0$ or $c = 7$ or $c = 2$.

35 $A^{-1} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ and $\boldsymbol{x} = A^{-1} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 2 \\ 1 \end{bmatrix}$. When the triangular A alternates

   1 and -1 on its diagonals, A^{-1} has 1's on the main diagonal and next diagonal.

36 $\boldsymbol{x} = (1, 1, \ldots, 1)$ has $\boldsymbol{x} = P\boldsymbol{x} = Q\boldsymbol{x}$ so $(P - Q)\boldsymbol{x} = \boldsymbol{0}$. Permutations do not change this all-ones vector. Then $P - Q$ is not invertible.

37 The block inverses are $\begin{bmatrix} I & 0 \\ -C & I \end{bmatrix}$ and $\begin{bmatrix} A^{-1} & 0 \\ -D^{-1}CA^{-1} & D^{-1} \end{bmatrix}$ and $\begin{bmatrix} -D & I \\ I & 0 \end{bmatrix}$.

38 A is invertible when elimination (with row exchanges allowed) produces 3 nonzero pivots.

39 $\left(I - \boldsymbol{uv}^{\mathrm{T}}\right)\left(I + \boldsymbol{uv}^{\mathrm{T}}\left(I - \boldsymbol{v}^{\mathrm{T}}\boldsymbol{u}\right)^{-1}\right)$

   $= I - \boldsymbol{uv}^{\mathrm{T}} + \boldsymbol{uv}^{\mathrm{T}}\left(I - \boldsymbol{v}^{\mathrm{T}}\boldsymbol{u}\right)^{-1} - \left(\boldsymbol{v}^{\mathrm{T}}\boldsymbol{u}\right)\boldsymbol{uv}^{\mathrm{T}}\left(I - \boldsymbol{v}^{\mathrm{T}}\boldsymbol{u}\right)^{-1}$

   $= I - \boldsymbol{uv}^{\mathrm{T}} + \boldsymbol{uv}^{\mathrm{T}} = I$

Problem Set 2.3, page 61

1 $\ell_{21} = 1$ multiplied row 1 and subtracted from row 2; **in reverse** $L = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ times

$$U\boldsymbol{x} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix} = \boldsymbol{c} \text{ is } A\boldsymbol{x} = \boldsymbol{b} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}.$$

In letters, **L multiplies $U\boldsymbol{x} = \boldsymbol{c}$ to give $A\boldsymbol{x} = \boldsymbol{b}$.**

2 $L\boldsymbol{c} = \boldsymbol{b}$ is $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$, solved by $\boldsymbol{c} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$ as elimination goes forward.

$U\boldsymbol{x} = \boldsymbol{c}$ is $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$, solved by $\boldsymbol{x} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ in back substitution.

3 $EA = \begin{bmatrix} 1 & & \\ 0 & 1 & \\ -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 0 & 4 & 2 \\ 6 & 3 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 4 & 2 \\ 0 & 0 & 5 \end{bmatrix} = U.$

With E^{-1} as L, $A = LU = \begin{bmatrix} 1 & & \\ 0 & 1 & \\ 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 0 & 4 & 2 \\ 0 & 0 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 4 & 2 \\ 6 & 3 & 5 \end{bmatrix}.$

4 $\begin{bmatrix} 1 & & \\ 0 & 1 & \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & & \\ -2 & 1 & \\ 0 & 0 & 1 \end{bmatrix} A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & -6 \end{bmatrix} = U.$ Then $A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} U$ is

the same as $E_{21}^{-1} E_{32}^{-1} U = LU$. The multipliers $\ell_{21} = \ell_{32} = 2$ fall into place in L.

5 $E_{32} E_{31} E_{21} \, A = \begin{bmatrix} 1 & & \\ & 1 & \\ & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & & \\ & 1 & \\ -3 & & 1 \end{bmatrix} \begin{bmatrix} 1 & & \\ -2 & 1 & \\ & & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 2 & 2 & 2 \\ 3 & 4 & 5 \end{bmatrix}.$ This is

$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = U.$ Put those multipliers 2, 3, 2 into L. Then $A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{bmatrix} U = LU.$

6 $A = \begin{bmatrix} 2 & 4 \\ 4 & 11 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} \mathbf{1} & 0 \\ \mathbf{2} & \mathbf{1} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{2} \\ 0 & \mathbf{1} \end{bmatrix} = LDU; \; \boldsymbol{U} \text{ is } \boldsymbol{L}^{\mathrm{T}}$

$\begin{bmatrix} 1 & & \\ 4 & 1 & \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 4 & 0 \\ 0 & -4 & 4 \\ 0 & 0 & 4 \end{bmatrix} = \begin{bmatrix} \mathbf{1} & & \\ \mathbf{4} & \mathbf{1} & \\ 0 & -\mathbf{1} & \mathbf{1} \end{bmatrix} \begin{bmatrix} 1 & & \\ & -4 & \\ & & 4 \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{4} & 0 \\ 0 & \mathbf{1} & -\mathbf{1} \\ 0 & 0 & \mathbf{1} \end{bmatrix} = LDL^{\mathrm{T}}.$

7 $\begin{bmatrix} a & a & a & a \\ a & b & b & b \\ a & b & c & c \\ a & b & c & d \end{bmatrix} = \begin{bmatrix} 1 & & & \\ 1 & 1 & & \\ 1 & 1 & 1 & \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} a & a & a & a \\ & b-a & b-a & b-a \\ & & c-b & c-b \\ & & & d-c \end{bmatrix}$. Need $\begin{array}{l} a \neq 0 \text{ All of the} \\ b \neq a \text{ multipliers} \\ c \neq b \text{ are } \ell_{ij} = 1 \\ d \neq c \text{ for this } A \end{array}$

8 **Correction**: Problem 8 has the same L as **Problem 7**.

$\begin{bmatrix} a & r & r & r \\ a & b & s & s \\ a & b & c & t \\ a & b & c & d \end{bmatrix} = \begin{bmatrix} 1 & & & \\ 1 & 1 & & \\ 1 & 1 & 1 & \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} a & r & r & r \\ & b-r & s-r & s-r \\ & & c-s & t-s \\ & & & d-t \end{bmatrix}$. Need $\begin{array}{l} a \neq 0 \\ b \neq r \\ c \neq s \\ d \neq t \end{array}$

9 $\begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix} \boldsymbol{c} = \begin{bmatrix} 2 \\ 11 \end{bmatrix}$ gives $\boldsymbol{c} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$. Then $\begin{bmatrix} 2 & 4 \\ 0 & 1 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ gives $\boldsymbol{x} = \begin{bmatrix} -5 \\ 3 \end{bmatrix}$.

$A\boldsymbol{x} = \boldsymbol{b}$ is $LU\boldsymbol{x} = \begin{bmatrix} 2 & 4 \\ 8 & 17 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 2 \\ 11 \end{bmatrix}$. Eliminate to $\begin{bmatrix} 2 & 4 \\ 0 & 1 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \boldsymbol{c}$.

10 $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \boldsymbol{c} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$ gives $\boldsymbol{c} = \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix}$. Then $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix}$ gives $\boldsymbol{x} = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}$.

Those are forward elimination and back substitution for $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$.

11 (a) L goes to I (b) I goes to L^{-1} (c) LU goes to U. **Elimination multiplies by $\boldsymbol{L^{-1}}$.**

12 (a) Multiply $LDU = L_1 D_1 U_1$ by inverses to get $L_1^{-1} LD = D_1 U_1 U^{-1}$. The left side is lower triangular, the right side is upper triangular $\Rightarrow$ both sides are diagonal.

   (b) L, U, L_1, U_1 have diagonal 1's so $D = D_1$. Then $L_1^{-1} L$ and $U_1 U^{-1}$ are both I.

13
$$\begin{bmatrix} 1 & & \\ 1 & 1 & \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ & 1 & 1 \\ & & 1 \end{bmatrix} = LIU; \quad \begin{bmatrix} a & a & 0 \\ a & a+b & b \\ 0 & b & b+c \end{bmatrix} = L \begin{bmatrix} a & & \\ & b & \\ & & c \end{bmatrix} U.$$

A tridiagonal matrix A has **bidiagonal factors** L and U.

14 For the first matrix A, L keeps the 3 zeros at the start of rows. But U may not have the upper zero where $A_{24} = 0$. For the second matrix B, L keeps the bottom left zero at the start of row 4. U keeps the upper right zero at the start of column 4. **One zero in A and two zeros in B are filled in**.

15 The 2 by 2 upper submatrix A_2 has the first two pivots $5, 9$. Reason: Elimination on A starts in the upper left corner with elimination on A_2.

16
$$\begin{bmatrix} 1 & 2 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 6 & 2 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix} = A = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 5 & 1 \\ 0 & 6 & 4 \end{bmatrix}$$

17 $L^{\mathrm{T}}L = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ and $LL^{\mathrm{T}} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

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

Problem Set 2.4, page 71

1 $A = \begin{bmatrix} 1 & 0 \\ 9 & 3 \end{bmatrix}$ has $A^{\mathrm{T}} = \begin{bmatrix} 1 & 9 \\ 0 & 3 \end{bmatrix}$, $A^{-1} = \begin{bmatrix} 1 & 0 \\ -3 & 1/3 \end{bmatrix}$, $(A^{-1})^{\mathrm{T}} = (A^{\mathrm{T}})^{-1} =$

$\begin{bmatrix} 1 & -3 \\ 0 & 1/3 \end{bmatrix}$; $A = \begin{bmatrix} 1 & c \\ c & 0 \end{bmatrix}$ has $A^{\mathrm{T}} = A$ and $A^{-1} = \dfrac{1}{c^2} \begin{bmatrix} 0 & c \\ c & -1 \end{bmatrix} = (A^{-1})^{\mathrm{T}}$.

2 $(AB)^{\mathrm{T}} = \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix} = B^{\mathrm{T}} A^{\mathrm{T}}$. This answer is different from $A^{\mathrm{T}} B^{\mathrm{T}} = \begin{bmatrix} 7 & 2 \\ 3 & 1 \end{bmatrix}$

(except when $AB = BA$). $AA^{\mathrm{T}} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}$ and $A^{\mathrm{T}} A = \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix}$

3 (a) $((AB)^{-1})^{\mathrm{T}} = (B^{-1}A^{-1})^{\mathrm{T}} = (A^{-1})^{\mathrm{T}}(B^{-1})^{\mathrm{T}}$. This is also $(A^{\mathrm{T}})^{-1}(B^{\mathrm{T}})^{-1}$.

(b) If U is upper triangular, so is U^{-1}: then $(U^{-1})^{\mathrm{T}}$ is *lower* triangular.

4 $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ has $A^2 = 0$. But the diagonal of $A^{\mathrm{T}} A$ has dot products of columns of A

with *themselves*. If $A^{\mathrm{T}} A = 0$, zero dot products $\Rightarrow$ zero columns $\Rightarrow A = $ zero matrix.

5 (a) $\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{y} = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = 5$

(b) This answer 5 is the row $\boldsymbol{x}^{\mathrm{T}} A = \begin{bmatrix} 4 & 5 & 6 \end{bmatrix}$ times $\boldsymbol{y} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$.

(c) This is also the row $\boldsymbol{x}^{\mathrm{T}} = \begin{bmatrix} 0 & 1 \end{bmatrix}$ times $A\boldsymbol{y} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$.

6 $M^{\mathrm{T}} = \begin{bmatrix} A^{\mathrm{T}} & C^{\mathrm{T}} \\ B^{\mathrm{T}} & D^{\mathrm{T}} \end{bmatrix}$; $M^{\mathrm{T}} = M$ needs $A^{\mathrm{T}} = A$ *and* $\boldsymbol{B}^{\mathbf{T}} = \boldsymbol{C}$ *and* $D^{\mathrm{T}} = D$.

7 (a) False: $\begin{bmatrix} 0 & A \\ A & 0 \end{bmatrix}$ is symmetric only if $A = A^{\mathrm{T}}$.

(b) False: The transpose of AB is $B^{\mathrm{T}}A^{\mathrm{T}} = BA$. So $(AB)^{\mathrm{T}} = AB$ needs $BA = AB$.

(c) True: Invertible symmetric matrices have symmetric inverses ! Easiest proof is to transpose $AA^{-1} = I$. So unsymmetric A has unsymmetric A^{-1}.

(d) True: $(ABC)^{\mathrm{T}}$ is $C^{\mathrm{T}}B^{\mathrm{T}}A^{\mathrm{T}}(= CBA$ for symmetric matrices A, B, and C).

8 The 1 in row 1 has n choices; then the 1 in row 2 has $n - 1$ choices ... $(n!$ overall).

9 $P_1 P_2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ but $P_2 P_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

If P_3 and P_4 exchange *different* pairs of rows, then $P_3 P_4 = P_4 P_3 =$ both exchanges.

10 $(3, 1, 2, 4)$ and $(2, 3, 1, 4)$ keep 4 in place; 6 more even P's keep 1 or 2 or 3 in place; $(2, 1, 4, 3)$ and $(3, 4, 1, 2)$ and $(4, 3, 2, 1)$ exchange 2 pairs. $(1, 2, 3, 4)$ makes 12 evens.

11 The "reverse identity" P takes $(1, \ldots, n)$ into $(n, \ldots, 1)$. When rows and also columns are reversed, the $1, 1$ and n, n entries of A change places in PAP. So do the $1, n$ and $n, 1$ entries. In general $(PAP)_{ij}$ is $(A)_{n - i + 1, n - j + 1}$.

12 $(P\boldsymbol{x})^{\mathrm{T}}(P\boldsymbol{y}) = \boldsymbol{x}^{\mathrm{T}}P^{\mathrm{T}}P\boldsymbol{y} = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{y}$ since $P^{\mathrm{T}}P = I$. In general $P\boldsymbol{x} \cdot \boldsymbol{y} = \boldsymbol{x} \cdot P^{\mathrm{T}}\boldsymbol{y} \neq \boldsymbol{x} \cdot P\boldsymbol{y}$:

Non-equality where $P \neq P^{\mathrm{T}}$: $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \neq \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$.

13 $PA = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 6 \\ 1 & 2 & 3 \\ 0 & 4 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$ is upper triangular. Multiplying A *on the right* by a permutation matrix P_2 exchanges the *columns* of A. To make this A lower triangular, we also need P_1 to exchange rows 2 and 3:

$P_1 A P_2 = \begin{bmatrix} 1 & & \\ & & 1 \\ & 1 & \end{bmatrix} A \begin{bmatrix} & & 1 \\ & 1 & \\ 1 & & \end{bmatrix} = \begin{bmatrix} 6 & 0 & 0 \\ 5 & 4 & 0 \\ 3 & 2 & 1 \end{bmatrix}$.

14 A cyclic $P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$ or its transpose will have $P^3 = I : (1,2,3) \to (2,3,1) \to$

$(3,1,2) \to (1,2,3)$. The permutation $\widehat{P} = \begin{bmatrix} 1 & 0 \\ 0 & P \end{bmatrix}$ for the same P has $\widehat{P}^4 = \widehat{P} \neq I$.

Simple row exchanges have $P^2 = I$ and $P^3 = P$.

15 (a) If P sends row 1 to row 4, then P^{T} sends row 4 to row 1   (b) $P = \begin{bmatrix} E & 0 \\ 0 & E \end{bmatrix} = $

P^{T} with $E = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ moves all rows: 1 and 2 are exchanged, 3 and 4 are exchanged.

16 $A^2 - B^2$ and also ABA are symmetric if A and B are symmetric. But $(A+B)(A-B)$

and $ABAB$ are generally *not* symmetric. Transposes $(A-B)(A+B)$ and $BABA$.

17 (a) $5+4+3+2+1 = 15$ independent entries if $S = S^{\mathrm{T}}$ (b) L has 10 and D has 5;

total 15 in LDL^{T} (c) Zero diagonal if $A^{\mathrm{T}} = -A$, leaving $4+3+2+1 = 10$ choices.

(d) The diagonal of $A^{\mathrm{T}}A$ contains $||\text{row } 1||^2, ||\text{row } 2||^2, \ldots \Rightarrow$ never negative.

18 $\begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \mathbf{3} & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -7 \end{bmatrix} \begin{bmatrix} 1 & \mathbf{3} \\ 0 & 1 \end{bmatrix}$; $\begin{bmatrix} 1 & b \\ b & c \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \mathbf{b} & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & c - b^2 \end{bmatrix} \begin{bmatrix} 1 & \mathbf{b} \\ 0 & 1 \end{bmatrix}$

$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & & \\ -\frac{1}{2} & 1 & \\ \mathbf{0} & -\frac{2}{3} & 1 \end{bmatrix} \begin{bmatrix} 2 & & \\ & \frac{3}{2} & \\ & & \frac{4}{3} \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} & \mathbf{0} \\ & 1 & -\frac{2}{3} \\ & & 1 \end{bmatrix} = \boldsymbol{LDL^{\mathrm{T}}}.$

19 $\begin{bmatrix} 1 & & \\ 1 & & \\ & & 1 \end{bmatrix} A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 2 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & & \\ 0 & 1 & \\ 2 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ & 1 & 1 \\ & & -1 \end{bmatrix}$;

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

20 $A = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = P$ and $L = U = I$. Elimination on this $A = P$ exchanges rows 1-2 then rows 2-3 then rows 3-4.

21 One way to decide even vs. odd is to count all pairs that P has in the wrong order. Then P is even or odd when that count is even or odd. Hard step: Show that an exchange always switches that count! Then 3 or 5 exchanges will leave that count odd.

22 $A = \begin{bmatrix} 0 & 1 & 2 & 3 \\ 1 & 2 & 3 & 0 \\ 2 & 3 & 0 & 1 \\ 3 & 0 & 1 & 2 \end{bmatrix} = A^{\mathrm{T}}$ has $0, 1, 2, 3$ in every row. I don't know any rules for a symmetric construction like this "Hankel matrix" with constant antidiagonals.

23 Reordering the rows and/or the columns of $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ will move the entry **a**. So the result cannot be the transpose (which doesn't move **a**).

24 (a) Total currents are $A^{\mathrm{T}}\boldsymbol{y} = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & -1 & -1 \end{bmatrix} \begin{bmatrix} y_{BC} \\ y_{CS} \\ y_{BS} \end{bmatrix} = \begin{bmatrix} y_{BC} + y_{BS} \\ -y_{BC} + y_{CS} \\ -y_{CS} - y_{BS} \end{bmatrix}$.

(b) Either way $(A\boldsymbol{x})^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{x}^{\mathrm{T}}(A^{\mathrm{T}}\boldsymbol{y}) = x_B y_{BC} + x_B y_{BS} - x_C y_{BC} + x_C y_{CS} - x_S y_{CS} - x_S y_{BS}$. Six terms.

25 $P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$ and $P^3 = I$ so three rotations for $360°$; P rotates every $\boldsymbol{v}$ around the $(1, 1, 1)$ line by $120°$.

26 $L(U^{\mathrm{T}})^{-1}$ is lower triangular times lower triangular, so *lower triangular*. The transpose of $U^{\mathrm{T}}DU$ is $U^{\mathrm{T}}D^{\mathrm{T}}U^{\mathrm{T}\,\mathrm{T}} = U^{\mathrm{T}}DU$ again, so $U^{\mathrm{T}}DU$ is *symmetric*. The factorization multiplies lower triangular by symmetric to get LDU which is A.

27 These are groups: Lower triangular with diagonal 1's, diagonal invertible D, permutations P, orthogonal matrices with $Q^{\mathrm{T}} = Q^{-1}$.

28 There are $n!$ permutation matrices of order n. Eventually *two powers of P must be the same permutation*. And if $P^r = P^s$ then $P^{r-s} = I$. Certainly $r - s \leq n!$

$$P = \begin{bmatrix} P_2 & \\ & P_3 \end{bmatrix} \text{ is 5 by 5 with } P_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \text{ and } P_3 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \text{ and } \boldsymbol{P^6 = I}.$$

29 To split the matrix M into (symmetric S) + (anti-symmetric A), the only choice is $S = \frac{1}{2}(M + M^{\mathrm{T}})$ and $A = \frac{1}{2}(M - M^{\mathrm{T}})$.

30 Start from $Q^{\mathrm{T}}Q = I$, as in $\begin{bmatrix} \boldsymbol{q}_1^{\mathrm{T}} \\ \boldsymbol{q}_2^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \boldsymbol{q}_1 & \boldsymbol{q}_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

(a) The diagonal entries give $\boldsymbol{q}_1^{\mathrm{T}}\boldsymbol{q}_1 = 1$ and $\boldsymbol{q}_2^{\mathrm{T}}\boldsymbol{q}_2 = 1$: *unit vectors*

(b) The off-diagonal entry is $\boldsymbol{q}_1^{\mathrm{T}}\boldsymbol{q}_2 = 0$ (and in general $\boldsymbol{q}_i^{\mathrm{T}}\boldsymbol{q}_j = 0$)

(c) The leading example for Q is the rotation matrix $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$.

Problem Set 2.5, page 81

1 $\Delta y = y(x+h) - y(x) = (x+h)^3 - x^3$

$\qquad = x^3 + 3x^2h + 3xh^2 + h^3 - x^3$

First order term $= 3xh^2$ matches $h\, dy/dx$.

Second order term $= 3xh^2$ matches $\frac{1}{2}h^2 d^2y/dx^2 = \frac{1}{2}h^2(6x)$.

2 $y = e^x$ has $dy/dx = e^x = 1$ at the point $x = 0$. The tangent line at $x = 0$ is $y = 1+x$.

3 $y = e^x$ has $d^2y/dx^2 = e^x = 1$ at $x = 0$. The tangent parabola at $x = 0$ is

$y = 1 + x + \frac{1}{2}x^2$. At $x = h$ this parabola is $1 + h + \frac{1}{2}h^2 = e^h$ to second order.

4 $y = e^x$ has all $d^n y/dx^n = e^x = e^0 = 1$ at $x = 0$. So the Taylor series is $y(h) = e^h =$

$1 + h + \frac{1}{2}h^2 + \cdots + \frac{1}{n!}h^n + \cdots$.

5 $\sin h = h - \frac{1}{6}h^3 + \frac{1}{5!}h^5 - \cdots$ so the error in $\sin h - h$ is order h^3.

6 Separate the real and imaginary parts of e^{ix} (even and odd powers of i). Those two parts are exactly the Taylor series for $\cos x$ and $\sin x$. So comparing Taylor series produces $e^{ix} = \cos x + i \sin x$.

7 Centered $\dfrac{(1/10)^3 - (-1/10)^3}{2/10} = \dfrac{1}{100}$ is approximating $\dfrac{dy}{dx} = 3x^2 = 0$ at $x = 0$.

Forward $\dfrac{(1/10)^3 - 0}{1/10} = \dfrac{1}{100}$ $\qquad$ Backward $\dfrac{0 - (-1/10)^3}{1/10} = \dfrac{1}{100}$.

By chance all three give the same answer.

8 Substitute $y(h)$ and also $y(-h)$ in the 3 options of Problem 7,:

Centered gives $\dfrac{2h\, dy/dx(0)}{2h} =$ exact through the h^2 term.

Forward gives $\dfrac{dy}{dx}(0) + \dfrac{1}{2}h\dfrac{d^2y}{dx^2}(0) =$ error of order h.

Backward gives $\dfrac{dy}{dx}(0) - \dfrac{1}{2}h\dfrac{d^2y}{dx^2}(0) =$ error of order h.

9 Compare $\dfrac{e - e^{-1}}{2}$ and $\dfrac{e-1}{1}$ and $\dfrac{1 - e^{-1}}{1}$ as approximations to $dy/dx(0) = e^0 = 1$.

Taking $e \approx 2.8$ and $e^{-1} \approx .35$, the first (centered) difference wins easily.

10 $(e - 2 + e^{-1}) \approx (2.78 - 2 + 0.36) = 1.14$. The correct second derivative of e^x at

$x = 0$ is $1.00 \colon 14\%$ error.

11 The 3rd difference uses numbers $1, -3, 3, -1$.

Test on $y = x^3$ $\qquad 3^3 - 3(2^3) + 3(1^3) - (0^3) = 27 - 24 + 3 = 0 = $ **correct at $x = 0$**.

Test on $y = x^4$ $\qquad 3^4 - 3(2^4) + 3(1^4) - (0^4) = 81 - 48 + 3 \neq 0$ **at $x = 0$**.

12 $\displaystyle\sum_{-\infty}^{\infty} f_i\, g_{i+1} = \sum_{-\infty}^{\infty} f_{i-1}\, g_i$ (just shifting i by 1 in infinite series).

$\displaystyle\sum_{-\infty}^{\infty} f_i\, g_{i-1} = \sum_{-\infty}^{\infty} f_{i+1}\, g_i$ (same idea).

Summation by parts is true $\colon$ Subtract second line from first line.

13 Test $u = x^2$ $\qquad -4 + 8 - 8 + 4 = 0 = $ correct derivative $\dfrac{du}{dx} = 2x$ at $x = 0$.

Test $u = x^4$ $\qquad -2^4 + 8 \cdot 1^4 - 8(-1)^4 + (-2)^4 = 0 = $ also correct $\dfrac{du}{dx} = 4x^3 = 0$

at $x = 0$.

14 The pattern of determinants indicates $\det \mathbf{K_5} = \mathbf{6}$ (correct !). The inverse is

$$
K^{-1} = \frac{1}{6}
\begin{bmatrix}
5 & 4 & 3 & 2 & 1 \\
4 & & & & 2 \\
3 & & & & 3 \\
2 & & & & 4 \\
1 & 2 & 3 & 4 & 5
\end{bmatrix}
= \frac{1}{6}
\begin{bmatrix}
5 & 4 & 3 & 2 & 1 \\
4 & 8 & 6 & 4 & 2 \\
3 & 6 & & 6 & 3 \\
2 & 4 & 6 & 8 & 4 \\
1 & 2 & 3 & 4 & 5
\end{bmatrix}
= \frac{1}{6}
\begin{bmatrix}
5 & 4 & 3 & 2 & 1 \\
4 & 8 & 6 & 4 & 2 \\
3 & 6 & 9 & 6 & 3 \\
2 & 4 & 6 & 8 & 4 \\
1 & 2 & 3 & 4 & 5
\end{bmatrix}
$$

15 Remove column 1 of A_0 to produce A_1 with $T = A_1^T A_1$. Remove columns 1 and 4

to produce A_2 with $K = A_2^T A_2$. Check $\colon$

$$
A_2^T =
\begin{bmatrix}
1 & -1 & 0 \\
0 & 1 & -1
\end{bmatrix}
\begin{bmatrix}
1 & 0 \\
-1 & 1 \\
0 & -1
\end{bmatrix}
=
\begin{bmatrix}
2 & -1 \\
-1 & 2
\end{bmatrix}.
$$

16 $D_4 D_4^T =
\begin{bmatrix}
2 & -1 & 0 & -1 \\
-1 & 2 & -1 & 0 \\
0 & -1 & 2 & -1 \\
-1 & 0 & -1 & 2
\end{bmatrix}
= C_4 = $ periodic circulant matrix (not invertible)

17 The solution to $-\dfrac{d^2 u}{dx^2} = \cos 4\pi x$ is $u(x) = \left(\dfrac{1}{4\pi}\right)^2 \cos 4\pi x + A + Bx$.

The boundary conditions $u = 0$ at $x = 0$ and $x = 1$ give

$$u(0) = \left(\frac{1}{4\pi}\right)^2 + A = 0 \text{ and } u(1) = \left(\frac{1}{4\pi}\right)^2 + A + B = 0.$$

Then $A = -\left(\dfrac{1}{4\pi}\right)^2$ and $B = 0$ and $u(x) = \left(\dfrac{1}{4\pi}\right)^2 (\cos 4\pi x - 1)$.

18 $\Delta_3 = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$ has column $3 = -$ column 1 : not invertible.

$\Delta_4 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}$ has 4 independent columns : invertible.

Problem Set 3.1, page 79

Note An interesting "max-plus" vector space comes from the real numbers **R** combined with $-\infty$. Change addition to give $x + y = \textbf{max}(\boldsymbol{x}, \boldsymbol{y})$ and change multiplication to $xy = \textbf{usual } \boldsymbol{x} + \boldsymbol{y}$. Which y is the zero vector that gives $x + \mathbf{0} = \max(x, \mathbf{0}) = x$ for every x?

1 $x + y \neq y + x$ and $x + (y + z) \neq (x + y) + z$ and $(c_1 + c_2)x \neq c_1x + c_2x$.

2 When $c(x_1, x_2) = (cx_1, 0)$, the only broken rule is 1 times x equals x. Rules (1)-(4) for addition $x + y$ still hold since addition is not changed.

3 (a) cx may not be in our set: not closed under multiplication. Also no $\mathbf{0}$ and no $-x$

   (b) $c(x + y)$ is the usual $(xy)^c$, while $cx + cy$ is the usual $(x^c)(y^c)$. Those are equal. With $c = 3$, $x = 2$, $y = 1$ this is $3(2 + 1) = 8$. The zero vector is the number 1.

4 The zero vector in matrix space **M** is $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$; $\frac{1}{2}A = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$ and $-A = \begin{bmatrix} -2 & 2 \\ -2 & 2 \end{bmatrix}$.
   The smallest subspace of **M** containing the matrix A consists of all matrices cA.

5 (a) One possibility: The matrices cA form a subspace not containing B (b) Yes: the subspace must contain $A - B = I$ (c) Matrices whose main diagonal is all zero.

6 When $\boldsymbol{f}(x) = x^2$ and $\boldsymbol{g}(x) = 5x$, the combination $3\boldsymbol{f} - 4\boldsymbol{g}$ in function space is $\boldsymbol{h}(x) = 3\boldsymbol{f}(x) - 4\boldsymbol{g}(x) = 3x^2 - 20x$.

7 Rule 8 is broken: If $c\boldsymbol{f}(x)$ is defined to be the usual $\boldsymbol{f}(cx)$ then $(c_1 + c_2)\boldsymbol{f} = \boldsymbol{f}((c_1 + c_2)x)$ is not generally the same as $c_1\boldsymbol{f} + c_2\boldsymbol{f} = \boldsymbol{f}(c_1x) + \boldsymbol{f}(c_2x)$.

8 (a) The vectors with integer components allow addition, but not multiplication by $\frac{1}{2}$
   (b) Remove the x axis from the xy plane (but leave the origin). Multiplication by any c is allowed but not all vector additions: $(1, 1) + (-1, 1) = (0, 2)$ is removed.

9 The only subspaces are (a) the plane with $b_1 = b_2$ (d) the linear combinations of v and w (e) the plane with $b_1 + b_2 + b_3 = 0$.

10 (a) All matrices $\begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}$ (b) All matrices $\begin{bmatrix} a & a \\ 0 & 0 \end{bmatrix}$ (c) All diagonal matrices.

11 For the plane $x + y - 2z = 4$, the sum of $(4, 0, 0)$ and $(0, 4, 0)$ is not on the plane. (The key is that this plane does not go through $(0, 0, 0)$.)

12 The parallel plane $\mathbf{P}_0$ has the equation $x + y - 2z = 0$. Pick two points, for example $(2, 0, 1)$ and $(0, 2, 1)$, and their sum $(2, 2, 2)$ is in $\mathbf{P}_0$.

13 The smallest subspace containing a plane $\mathbf{P}$ and a line $\mathbf{L}$ is *either* $\mathbf{P}$ (when the line $\mathbf{L}$ is in the plane $\mathbf{P}$) *or* $\mathbf{R}^3$ (when $\mathbf{L}$ is not in $\mathbf{P}$).

14 (a) The invertible matrices do not include the zero matrix, so they are not a subspace

(b) The sum of singular matrices $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ is not singular: not a subspace.

15 (a) *True*: The symmetric matrices do form a subspace (b) *True*: The matrices with $A^{\mathrm{T}} = -A$ do form a subspace (c) *True*: Any set of vectors from a vector space will span a subspace of that space.

16 The column space of A is the x-axis = all vectors $(x, 0, 0)$: a *line*. The column space of B is the xy plane = all vectors $(x, y, 0)$. The column space of C is the line of vectors $(x, 2x, 0)$.

17 (a) Elimination leads to $0 = b_2 - 2b_1$ and $0 = b_1 + b_3$ in equations 2 and 3: Solution only if $b_2 = 2b_1$ and $b_3 = -b_1$ (b) Elimination leads to $0 = b_1 + b_3$ in equation 3: Solution only if $b_3 = -b_1$.

18 A combination of the columns of C is also a combination of the columns of A. Then $C = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$ and $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ have the same column space. $B = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$ has a different column space. The key word is "space".

19 (a) Solution for every $\boldsymbol{b}$ (b) Solvable only if $b_3 = 0$ (c) Solvable only if $b_3 = b_2$.

20 The extra column $\boldsymbol{b}$ enlarges the column space unless $\boldsymbol{b}$ is *already in* the column space.

$[A \ \boldsymbol{b}] = \begin{bmatrix} 1 & 0 & \mathbf{1} \\ 0 & 0 & \mathbf{1} \end{bmatrix}$ (larger column space) $\begin{bmatrix} 1 & 0 & \mathbf{1} \\ 0 & 1 & \mathbf{1} \end{bmatrix}$ ($\boldsymbol{b}$ is in column space)

(no solution to $A\boldsymbol{x} = \boldsymbol{b}$) ($A\boldsymbol{x} = \boldsymbol{b}$ has a solution)

21 The column space of AB is *contained in* (possibly equal to) the column space of A. The example B = zero matrix and $A \neq 0$ is a case when AB = zero matrix has a smaller column space (it is just the zero space $\mathbf{Z}$) than A.

22 The solution to $Az = b + b^*$ is $z = x + y$. If b and b^* are in $C(A)$ so is $b + b^*$.

23 The column space of any invertible 5 by 5 matrix is R^5. The equation $Ax = b$ is always solvable (by $x = A^{-1}b$) so every b is in the column space of that invertible matrix.

24 (a) *False*: Vectors that are *not* in a column space don't form a subspace.

(b) *True*: Only the zero matrix has $C(A) = \{0\}$. (c) *True*: $C(A) = C(2A)$.

(d) *False*: $C(A - I) \neq C(A)$ when $A = I$ or $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ (or other examples).

25 $A = \begin{bmatrix} 1 & 1 & \mathbf{0} \\ 1 & 0 & \mathbf{0} \\ 0 & 1 & \mathbf{0} \end{bmatrix}$ and $\begin{bmatrix} 1 & 1 & \mathbf{2} \\ 1 & 0 & \mathbf{1} \\ 0 & 1 & \mathbf{1} \end{bmatrix}$ do not have $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ in $C(A)$. $A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & 0 \\ 3 & 6 & 0 \end{bmatrix}$ has $C(A) = $ line in R^3.

26 When $Ax = b$ is solvable for all b, every b is in the column space of A. So that space is $C(A) = R^9$.

27 (a) If u and v are both in $S + T$, then $u = s_1 + t_1$ and $v = s_2 + t_2$. So $u + v = (s_1 + s_2) + (t_1 + t_2)$ is also in $S + T$. And so is $cu = cs_1 + ct_1$: $S + T = $ *subspace*.

(b) If S and T are different lines, then $S \cup T$ is just the two lines (*not a subspace*) but $S + T$ is the whole plane that they span.

28 If $S = C(A)$ and $T = C(B)$ then $S + T$ is the column space of $M = [\,A \quad B\,]$.

29 The columns of AB are combinations of the columns of A. So all columns of $[\,A \quad AB\,]$ are already in $C(A)$. But $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ has a larger column space than $A^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.
For square matrices, the column space is R^n exactly when A is *invertible*.

30 $y - e^{-x}$ and $y = e^x$ are independent solutions to $d^2y/dx^2 = y$. Also $y = \cos x$ and $y = \sin x$ are independent solutions to $d^2y/dx^2 = -y$. The solution space contains all combinations $A\cos x + B\sin x$.

31 If x and y are in the vector space $V \cap W$, then they are in both V and W. So all combinations $cx + dy$ are in both V and W. So all combinations are in $V \cap W$.

Problem Set 3.2, page 100

1 If $Ax = 0$ then $EAx = 0$. If $EAx = 0$, multiply by E^{-1} to find $Ax = 0$.

2 (a) If $c = 4$ then A has rank 1 and column 1 is its pivot column and $(-2, 1, 0)$ and $(-1, 0, 1)$ are special solutions to $Ax = 0$. If $c \neq 4$ then A has rank 2 and columns 1 and 3 are pivot columns and $(-2, 1, 0)$ is a special solution. If $c = 0$ then $B = $ zero matrix with rank 0 and $(1, 0)$ and $(0, 1)$ are special solutions to $Bx = 0$. If $c \neq 0$ then B has rank 1 and column 1 is its pivot column and $(-1, 1)$ is the special solution to $Bx = 0$.

3 $R = \begin{bmatrix} 1 & 3 & 0 & 2 \\ 0 & 0 & 1 & 6 \end{bmatrix}$. All matrices $A = CR$ with $C = 2$ by 2 invertible matrix have the same nullspace as R.

4 (a) $R = \begin{bmatrix} 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 3 \end{bmatrix}$ Free variables x_2, x_4, x_5
Pivot variables x_1, x_3 (b) $R = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix}$ Free x_3
Pivot x_1, x_2

5 Free variables x_2, x_4, x_5 and solutions $(-2, 1, 0, 0, 0)$, $(0, 0, -2, 1, 0)$, $(0, 0, -3, 0, 1)$.

6 (a) *False*: Any singular square matrix would have free variables   (b) *True*: An invertible square matrix has *no* free variables.   (c) *True* (only n columns to hold pivots)   (d) *True* (only m rows to hold pivots)

7 $A = \begin{bmatrix} C \end{bmatrix} \begin{bmatrix} I & I \end{bmatrix}$ (notice that $F = I$). The r special solutions to $Ax = 0$ are the r columns of $\begin{bmatrix} -I \\ I \end{bmatrix}$.

8 $R = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.

Notice the identity matrix in the pivot columns of these *reduced* row echelon forms R.

9 If column 4 of a 3 by 5 matrix is all zero then x_4 is a *free* variable. Its special solution is $x = (0, 0, 0, 1, 0)$, because 1 will multiply that zero column to give $Ax = 0$.

10 If column 1 = column 5 then x_5 is a free variable. Its special solution is $(-1, 0, 0, 0, 1)$.

11 The nullspace contains only $\boldsymbol{x} = \boldsymbol{0}$ when A has 5 pivots. Also the column space is $\mathbf{R}^5$, because we can always solve $A\boldsymbol{x} = \boldsymbol{b}$ and every $\boldsymbol{b}$ is in the column space.

12 If a matrix has n columns and r pivots, there are $\boldsymbol{n} - \boldsymbol{r}$ special solutions. The nullspace contains only $\boldsymbol{x} = \boldsymbol{0}$ when $\boldsymbol{r} = \boldsymbol{n}$. The column space is all of $\mathbf{R}^m$ when $r = m$. All those statements are important!

13 Fill in **12** then **3** then **1** to get the complete solution in $\mathbf{R}^3$ to $x - 3y - z = 12$

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \mathbf{12} \\ 0 \\ 0 \end{bmatrix} + y \begin{bmatrix} \mathbf{3} \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} \mathbf{1} \\ 0 \\ 1 \end{bmatrix} = \text{one particular solution} + \text{all nullspace solutions.}$$

14 Column 5 is sure to have no pivot since it is a combination of earlier columns. With 4 pivots in the other columns, the special solution is $\boldsymbol{s} = (1, 0, 1, 0, 1)$. The nullspace contains all multiples of this vector $\boldsymbol{s}$ (this nullspace is a line in $\mathbf{R}^5$).

15 To produce special solutions $(2, 2, 1, 0)$ and $(3, 1, 0, 1)$ with free variables x_3, x_4:

$$R = \begin{bmatrix} 1 & 0 & -2 & -3 \\ 0 & 1 & -2 & -1 \end{bmatrix} \text{ and } A \text{ can be any invertible 2 by 2 matrix times this } R.$$

16 The nullspace of $A = \begin{bmatrix} 1 & 0 & 0 & -4 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & -2 \end{bmatrix}$ $\quad$ The rank is 3 $\quad$ is the line through the special solution $\begin{bmatrix} 4 \\ 3 \\ 2 \\ 1 \end{bmatrix}$.

17 $A = \begin{bmatrix} 1 & 0 & -1/2 \\ 1 & 3 & -2 \\ 5 & 1 & -3 \end{bmatrix}$ has $\begin{bmatrix} 1 \\ 1 \\ 5 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix}$ in $\mathbf{C}(A)$ and $\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$ in $\mathbf{N}(A)$. *Which other A's?*

18 $A = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & -1 \\ 0 & 1 & 0 \end{bmatrix}$

19 $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ has $\mathbf{N}(A) = \mathbf{C}(A)$. Notice that $\text{rref}(A^\mathrm{T}) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ is not A^T.

20 If nullspace = column space (with r pivots) then $n - r = r$. If $n = 3$ then $3 = 2r$ is impossible. Only possible when n is even.

21 If A times every column of B is zero, the column space of B is contained in the *nullspace* of A. An example is $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$. Here $\mathbf{C}(B)$ *equals* $\mathbf{N}(A)$.
For $B = 0, \mathbf{C}(B)$ is smaller than $\mathbf{N}(A)$.

22 For $A = $ random 3 by 3 matrix, R is almost sure to be I. For 4 by 3, R is most likely to be I with a fourth row of zeros. What is R for a random 3 by 4 matrix?

23 If $\mathbf{N}(A) = $ line through $\boldsymbol{x} = (2, 1, 0, 1)$, A has *three pivots* (4 columns and 1 special solution). Its reduced echelon form can be $R = \begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ (add any zero rows).

24 $R = \begin{bmatrix} 1 & -2 & -3 \end{bmatrix}$, $R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, $R = I$. Any zero rows come after those rows.

25 (a) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ (b) All 8 matrices are R's !

26 The nullspace of $B = \begin{bmatrix} A & A \end{bmatrix}$ contains all vectors $\boldsymbol{x} = \begin{bmatrix} \boldsymbol{y} \\ -\boldsymbol{y} \end{bmatrix}$ for $\boldsymbol{y}$ in $\mathbf{R}^4$.

One reason that R is the same for A and $-A$: They have the same nullspace. (They also have the same row space. They also have the same column space, but that is not required for two matrices to share the same R. R tells us the nullspace and row space.)

27 If $C\boldsymbol{x} = \mathbf{0}$ then $A\boldsymbol{x} = \mathbf{0}$ and $B\boldsymbol{x} = \mathbf{0}$. So $\mathbf{N}(C) = \mathbf{N}(A) \cap \mathbf{N}(B) = $ *intersection*.

28 A has $R_0 = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \end{bmatrix}$ and $R = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$. B and C have $R_0 = \begin{bmatrix} 1 & 2 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

And $R = \begin{bmatrix} 1 & 2 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 & 3 \end{bmatrix}$.

29 $R = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and $N = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

30 A and A^{T} have the same rank $r =$ number of pivots. But the pivot column is column 2

for this matrix A and column 1 for A^{T}: $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

31 The new entries keep rank 1 : $A = \begin{bmatrix} a & b & c \\ d & \frac{bd}{a} & \frac{cd}{a} \\ g & \frac{bg}{a} & \frac{cg}{a} \end{bmatrix}$ if $a \neq 0$, $B = \begin{bmatrix} \mathbf{3} & 9 & -\mathbf{4.5} \\ 1 & \mathbf{3} & -\mathbf{1.5} \\ 2 & 6 & -3 \end{bmatrix}$,

$M = \begin{bmatrix} a & b \\ c & \mathbf{bc/a} \end{bmatrix}$ if $a \neq 0$.

32 With rank 1, the second row of R does not exist !

33 Invertible r by r submatrices $\quad S = \begin{bmatrix} 1 & 3 \\ 1 & 4 \end{bmatrix}$ and $S = [\,1\,]$ and $S = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.
Use pivot rows and columns

34 (a) A and B will both have the same nullspace and row space as the R they share.

(b) A equals an *invertible* matrix times B, when they share the same R. A key fact!

35 CORRECTED : $A^{\mathrm{T}} \mathbf{y} = \mathbf{0}$: $y_1 - y_3 + y_4 = -y_1 + y_2 + y_5 = -y_2 + y_3 + y_6 =$

$-y_4 - y_5 - y_6 = 0$.

These equations add to $0 = 0$. Free variables y_3, y_5, y_6: watch for flows around loops.

The solutions to $A^{\mathrm{T}} \mathbf{y} = \mathbf{0}$ are combinations of $(-1, 0, 0, 1, -1, 0)$ and $(0, 0, -1, -1, 0, 1)$

and $(0, -1, 0, 0, 1, -1)$. Those are flows around the 3 small loops.

36 $C = \begin{bmatrix} 1 & 3 \\ 2 & 6 \\ 2 & 7 \end{bmatrix}$ C^{T} has pivot columns $\begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix}$. The invertible S inside C is $\begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix}$

37 The column space of AB contains all vectors $(AB)\mathbf{x}$. Those vectors are the same as

$A(B\mathbf{x})$ so they are also in the column space of A.

38 By matrix multiplication, each column of AB is A times the corresponding column of B. So if column j of B is a combination of earlier columns of B, then column j of AB is the same combination of earlier columns of AB. Then rank $(AB) \leq$ rank (B). No new pivot columns !

39 We are given $AB = I$ which has rank n. Then $\text{rank}(AB) \leq \text{rank}(A)$ forces $\text{rank}(A) = n$. This means that A is invertible. The right-inverse B is also a left-inverse: $BA = I$ and $B = A^{-1}$.

40 Certainly A and B have at most rank 2. Then their product AB has at most rank 2. Since BA is 3 by 3, it cannot be I even if $AB = I$. *Example* $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$,

$B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, BA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

41 $A = \begin{bmatrix} I & I \end{bmatrix}$ has $N = \begin{bmatrix} I \\ -I \end{bmatrix}$; $B = \begin{bmatrix} I & I \\ 0 & 0 \end{bmatrix}$ has the same N; $C = \begin{bmatrix} I & I & I \end{bmatrix}$ has

$N = \begin{bmatrix} -I & -I \\ I & 0 \\ 0 & I \end{bmatrix}$.

42 The m by n matrix Z has r ones to start its main diagonal. Otherwise Z is all zeros.

43 $R_0 = \begin{bmatrix} I & F \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} r \text{ by } r & r \text{ by } n-r \\ m-r \text{ by } r & m-r \text{ by } n-r \end{bmatrix}$; (b) $B = \begin{bmatrix} I \\ 0 \end{bmatrix}$ (c) $C = \begin{bmatrix} I & 0 \end{bmatrix}$

$\mathbf{rref}(R_0^\mathrm{T}) = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$; $\mathbf{rref}(R_0^\mathrm{T} R_0) =$ same R_0

44 $R_0 = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ has $R_0^{\mathrm{T}} R_0 = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and this matrix row reduces to $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = $

$\begin{bmatrix} R_0 \\ \text{zero row} \end{bmatrix}$. Always $R_0^{\mathrm{T}} R_0$ has the same nullspace as R_0, so its row reduced form must be R_0 with $n - m$ extra zero rows. R_0 is determined by its nullspace and shape !

45 $A = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 2 & 5 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \end{bmatrix}$

Notice 2 rows of A are in the matrix B.

46 Multiply block row 1 by JW^{-1} to produce row 2.

Problem Set 3.3, page 111

1
$$\begin{bmatrix} 2 & 4 & 6 & 4 & \mathbf{b}_1 \\ 2 & 5 & 7 & 6 & \mathbf{b}_2 \\ 2 & 3 & 5 & 2 & \mathbf{b}_3 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 4 & 6 & 4 & \mathbf{b}_1 \\ 0 & 1 & 1 & 2 & \mathbf{b}_2 - \mathbf{b}_1 \\ 0 & -1 & -1 & -2 & \mathbf{b}_3 - \mathbf{b}_1 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 4 & 6 & 4 & \mathbf{b}_1 \\ 0 & 1 & 1 & 2 & \mathbf{b}_2 - \mathbf{b}_1 \\ 0 & 0 & 0 & 0 & \mathbf{b}_3 + \mathbf{b}_2 - 2\mathbf{b}_1 \end{bmatrix} \begin{matrix} 4 \\ -1 \\ 0 \end{matrix}$$

$A\boldsymbol{x} = \boldsymbol{b}$ has a solution when $b_3 + b_2 - 2b_1 = 0$; the column space contains all combinations of $(2, 2, 2)$ and $(4, 5, 3)$. **This is the plane** $b_3 + b_2 - 2b_1 = 0$ (!). The nullspace contains all combinations of $\boldsymbol{s}_1 = (-1, -1, 1, 0)$ and $\boldsymbol{s}_2 = (2, -2, 0, 1)$; $\boldsymbol{x}_{complete} = \boldsymbol{x}_p + c_1\boldsymbol{s}_1 + c_2\boldsymbol{s}_2$;

$$\begin{bmatrix} R_0 & \boldsymbol{d} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & -2 & 4 \\ 0 & 1 & 1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \text{ gives the particular solution } x_p = (4, -1, 0, 0).$$

2
$$\begin{bmatrix} 2 & 1 & 3 & \mathbf{b}_1 \\ 6 & 3 & 9 & \mathbf{b}_2 \\ 4 & 2 & 6 & \mathbf{b}_3 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 1 & 3 & \mathbf{b}_1 \\ 0 & 0 & 0 & \mathbf{b}_2 - 3\mathbf{b}_1 \\ 0 & 0 & 0 & \mathbf{b}_3 - 2\mathbf{b}_1 \end{bmatrix} \quad \text{Then } \begin{bmatrix} R_0 & \boldsymbol{d} \end{bmatrix} = \begin{bmatrix} 1 & 1/2 & 3/2 & \mathbf{5} \\ 0 & 0 & 0 & \mathbf{0} \\ 0 & 0 & 0 & \mathbf{0} \end{bmatrix}$$

$A\boldsymbol{x} = \boldsymbol{b}$ has a solution when $b_2 - 3b_1 = 0$ and $b_3 - 2b_1 = 0$; $\mathbf{C}(A) =$ line through $(2, 6, 4)$ which is the intersection of the planes $b_2 - 3b_1 = 0$ and $b_3 - 2b_1 = 0$; the nullspace contains all combinations of $\boldsymbol{s}_1 = (-1/2, 1, 0)$ and $\boldsymbol{s}_2 = (-3/2, 0, 1)$; particular solution $\boldsymbol{x}_p = \boldsymbol{d} = (5, 0, 0)$ and complete solution $\boldsymbol{x}_p + c_1\boldsymbol{s}_1 + c_2\boldsymbol{s}_2$.

3 (a)
$$\begin{matrix} x + 3y = 7 \\ 2x + 6y = 14 \end{matrix} \qquad \begin{matrix} x + 3y = 7 \\ 0 = 0 \end{matrix} \qquad \boldsymbol{x}_p = \begin{bmatrix} 7 \\ 0 \end{bmatrix} \qquad \boldsymbol{x}_n = c\boldsymbol{s} = c\begin{bmatrix} -3 \\ 1 \end{bmatrix} \text{ for any } c.$$

(b) $\boldsymbol{x}_{complete} = \begin{bmatrix} 7 \\ 0 \end{bmatrix} + c\begin{bmatrix} -3 \\ 1 \end{bmatrix};$ $\qquad \boldsymbol{x}_{complete} = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} + c\begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix}.$

4 $\boldsymbol{x}_{complete} = \boldsymbol{x}_p + \boldsymbol{x}_n = (\frac{1}{2}, 0, \frac{1}{2}, 0) + x_2(-3, 1, 0, 0) + x_4(0, 0, -2, 1).$

5
$$\begin{bmatrix} 1 & 2 & -2 & b_1 \\ 2 & 5 & -4 & b_2 \\ 4 & 9 & -8 & b_3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & -2 & b_1 \\ 0 & 1 & 0 & b_2 - 2b_1 \\ 0 & 0 & 0 & b_3 - 2b_1 - b_2 \end{bmatrix}$$ solvable if $b_3 - 2b_1 - b_2 = 0$.

Back-substitution gives the particular solution to $A\boldsymbol{x} = \boldsymbol{b}$ and the special solution to

$$A\boldsymbol{x} = \boldsymbol{0}: \boldsymbol{x} = \begin{bmatrix} 5b_1 - 2b_2 \\ b_2 - 2b_1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}. \quad \begin{bmatrix} 2 & 0 & 2 & b_1 \\ 4 & 4 & 0 & b_2 \\ 8 & 8 & 0 & b_3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 1 & b_1/2 \\ 0 & 1 & -1 & b_2/4 - b_1/2 \\ 0 & 0 & 0 & b_3 - 2b_2 \end{bmatrix}$$

is solvable if $b_3 = 2b_2$. Then $\boldsymbol{x} = \begin{bmatrix} b_1/2 \\ b_2/4 - b_1/2 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$.

6 (a) Solvable if $b_2 = 2b_1$ and $3b_1 - 3b_3 + b_4 = 0$. Then $\boldsymbol{x} = \begin{bmatrix} 5b_1 - 2b_3 \\ b_3 - 2b_1 \end{bmatrix} = \boldsymbol{x}_p$

(b) Solvable if $b_2 = 2b_1$ and $3b_1 - 3b_3 + b_4 = 0$. $\boldsymbol{x} = \begin{bmatrix} 5b_1 - 2b_3 \\ b_3 - 2b_1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$.

7
$$\begin{bmatrix} 1 & 3 & 1 & b_1 \\ 3 & 8 & 2 & b_2 \\ 2 & 4 & 0 & b_3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & 1 & b_2 \\ 0 & -1 & -1 & b_2 - 3b_1 \\ 0 & -2 & -2 & b_3 - 2b_1 \end{bmatrix}$$
One more step gives $\begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix} =$ row 3 $-$ 2 (row 2) $+$ 4(row 1) **provided $b_3 - 2b_2 + 4b_1 = 0$.**

8 (a) Every $\boldsymbol{b}$ is in $\mathbf{C}(A)$: *independent rows*, only the zero combination gives $\boldsymbol{0}$.

(b) We need $b_3 = 2b_2$, because $(\text{row } 3) - 2(\text{row } 2) = \boldsymbol{0}$.

9 (a) $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 0 \end{bmatrix} + y \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ (b) $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 0 \end{bmatrix} + z \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$. The second

equation in part (b) removed one special solution from the nullspace.

10 $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$ has $\boldsymbol{x}_p = (2, 4, 0)$ and $\boldsymbol{x}_{\text{null}} = (c, c, c)$. Many possible A !

11 A 1 by 3 system has at least **two** free variables. But $\boldsymbol{x}_{\text{null}}$ in Problem 10 only has **one**.

12 (a) If $A\boldsymbol{x}_1 = \boldsymbol{b}$ and $A\boldsymbol{x}_2 = \boldsymbol{b}$ then $\boldsymbol{x} = \boldsymbol{x}_1 - \boldsymbol{x}_2$ and also $\boldsymbol{x} = \boldsymbol{0}$ solve $A\boldsymbol{x} = \boldsymbol{0}$

(b) $A(2\boldsymbol{x}_1 - 2\boldsymbol{x}_2) = \boldsymbol{0}, A(2\boldsymbol{x}_1 - \boldsymbol{x}_2) = \boldsymbol{b}$

13 (a) The particular solution x_p is always multiplied by 1. $2x_p$ would solve $Ax = 2b$

(b) Any solution can be x_p. If A has rank $= m$, the only x_p is $\mathbf{0}$.

(c) $\begin{bmatrix} 3 & 3 \\ 3 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 6 \\ 6 \end{bmatrix}$. Then $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is shorter (length $\sqrt{2}$) than $\begin{bmatrix} 2 \\ 0 \end{bmatrix}$ (length 2)

(d) The only "homogeneous" solution in the nullspace is $x_n = \mathbf{0}$ when A is invertible.

14 If column 5 has no pivot, x_5 is a *free* variable. The zero vector *is not* the only solution to $Ax = \mathbf{0}$. If this system $Ax = b$ has a solution, it has *infinitely many* solutions.

15 If row 3 of U has no pivot, that is a *zero row*. $Ux = c$ is only solvable provided $c_3 = 0$. $Ax = b$ *might not be solvable*, because U may have other zero rows needing more $c_i = 0$.

16 The largest rank is 3. Then there is a pivot in every *row*. The solution *always exists*. The column space is $\mathbf{R}^3$. An example is $A = \begin{bmatrix} I & F \end{bmatrix}$ for any 3 by 2 matrix F.

17 The largest rank of a 6 by 4 matrix is 4. Then there is a pivot in every *column*. The columns are independent. The solution is *unique* (if there is a solution). The nullspace contains only the *zero vector*. Then $\mathbf{R_0} = \mathbf{rref}(A) = \begin{bmatrix} I & (4 \text{ by } 4) \\ 0 & (2 \text{ by } 4) \end{bmatrix}$.

18 Rank $= 2$; rank $= 3$ unless $q = 2$ (then rank $= 2$). Transpose has the same rank!

19 If $Ax_1 = b$ and also $Ax_2 = b$ then $A(x_1 - x_2) = \mathbf{0}$ and we can add $x_1 - x_2$ to any solution of $Ax = B$: the solution x is not unique. But there will be **no solution** to $Ax = B$ if B is not in the column space.

20 For $A, q = 3$ gives rank 1, every other q gives rank 2. For $B, q = 6$ gives rank 1, every other q gives rank 2. These matrices cannot have rank 3.

21 (a) $\begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$ has 0 or 1 solutions, depending on b (b) $\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b \end{bmatrix}$ has infinitely many solutions for every b (c) There are 0 or ∞ solutions when A has rank $r < m$ and $r < n$: the simplest example is a zero matrix. (d) *one* solution for all b when A is square and invertible (like $A = I$).

22 (a) $r < m$, always $r \leq n$ (b) $r = m, r < n$ (c) $r < m, r = n$ (d) $r = m = n$.

23 $\begin{bmatrix} 2 & 4 & 4 \\ 0 & 3 & 6 \\ 0 & 0 & 0 \end{bmatrix} \to R_0 = \begin{bmatrix} \mathbf{1} & 0 & -2 \\ 0 & \mathbf{1} & 2 \\ 0 & 0 & 0 \end{bmatrix}$ and $\begin{bmatrix} 2 & 4 & 4 \\ 0 & 3 & 6 \\ 0 & 0 & 5 \end{bmatrix} \to R_0 = I = R$ and

$\begin{bmatrix} 0 & 0 & 4 \\ 0 & 1 & 0 \end{bmatrix} \to R_0 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = R.$

24 $R_0 = I$ when A is square and invertible—so for a triangular matrix, all diagonal entries must be nonzero.

25 $\begin{bmatrix} 1 & 2 & 3 & \mathbf{0} \\ 0 & 0 & 4 & \mathbf{0} \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 0 & \mathbf{0} \\ 0 & 0 & 1 & \mathbf{0} \end{bmatrix}$; $\boldsymbol{x}_n = \begin{bmatrix} -2 \\ \mathbf{1} \\ 0 \end{bmatrix}$; $\begin{bmatrix} 1 & 2 & 3 & \mathbf{5} \\ 0 & 0 & 4 & \mathbf{8} \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 0 & -\mathbf{1} \\ 0 & 0 & 1 & \mathbf{2} \end{bmatrix}.$

Free $x_2 = 0$ gives $\boldsymbol{x}_p = (-1, 0, 2)$ because the pivot columns contain I. Note: $R_0 = R.$

26 $[\, R_0 \ \boldsymbol{d} \,] = \begin{bmatrix} 1 & 0 & 0 & \mathbf{0} \\ 0 & 0 & 1 & \mathbf{0} \\ 0 & 0 & 0 & \mathbf{0} \end{bmatrix}$ leads to $\boldsymbol{x}_n = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$; $[\, R_0 \ \boldsymbol{d} \,] = \begin{bmatrix} 1 & 0 & 0 & -\mathbf{1} \\ 0 & 0 & 1 & \mathbf{2} \\ 0 & 0 & 0 & \mathbf{5} \end{bmatrix}$

leads to no solution because of the 3rd equation $0 = 5$.

27 $\begin{bmatrix} 1 & 0 & 2 & 3 & \mathbf{2} \\ 1 & 3 & 2 & 0 & \mathbf{5} \\ 2 & 0 & 4 & 9 & \mathbf{10} \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 2 & 3 & \mathbf{2} \\ 0 & 3 & 0 & -3 & \mathbf{3} \\ 0 & 0 & 0 & 3 & \mathbf{6} \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 2 & 0 & -\mathbf{4} \\ 0 & 1 & 0 & 0 & \mathbf{3} \\ 0 & 0 & 0 & 1 & \mathbf{2} \end{bmatrix}$; $\begin{bmatrix} -4 \\ 3 \\ 0 \\ 2 \end{bmatrix}$; $\boldsymbol{x}_n = x_3 \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \end{bmatrix}.$

28 For $A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \\ 0 & 3 \end{bmatrix}$, the only solution to $A\boldsymbol{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ is $\boldsymbol{x} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$

B cannot exist since 2 equations in 3 unknowns cannot have a unique solution.

29 $A = \begin{bmatrix} 1 & 3 & 1 \\ 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 1 & 5 \end{bmatrix}$ factors into $LU = \begin{bmatrix} 1 & & & \\ 1 & 1 & & \\ 2 & 2 & 1 & \\ 1 & 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 1 \\ 0 & -1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ and the rank is

$r = 2$. The special solution to $A\boldsymbol{x} = \boldsymbol{0}$ and $U\boldsymbol{x} = \boldsymbol{0}$ is $\boldsymbol{s} = (-7, 2, 1)$. Since $\boldsymbol{b} = (1, 3, 6, 5)$ is also the last column of A, a particular solution to $A\boldsymbol{x} = \boldsymbol{b}$ is $(0, 0, 1)$ and the complete solution is $\boldsymbol{x} = (0, 0, 1) + c\boldsymbol{s}$. (Another particular solution is $\boldsymbol{x}_p = (7, -2, 0)$ with free variable $x_3 = 0$.)

For $\boldsymbol{b} = (1, 0, 0, 0)$ elimination leads to $U\boldsymbol{x} = (1, -1, 0, 1)$ and the fourth equation is $0 = 1$. No solution for this $\boldsymbol{b}$.

30 If the complete solution to $A\boldsymbol{x} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ is $\boldsymbol{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ c \end{bmatrix}$ then $A = \begin{bmatrix} 1 & 0 \\ 3 & 0 \end{bmatrix}$.

31 (a) If $\boldsymbol{s} = (2, 3, 1, 0)$ is the only special solution to $A\boldsymbol{x} = \boldsymbol{0}$, the complete solution is $\boldsymbol{x} = c\boldsymbol{s}$ (a line of solutions). The rank of A must be $4 - 1 = 3$.

(b) The fourth variable x_4 is *not free* in $\boldsymbol{s}$, and R_0 must be $\begin{bmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

(c) $A\boldsymbol{x} = \boldsymbol{b}$ can be solved for all $\boldsymbol{b}$, because A and R_0 have *full row rank $r = 3$.*

32 If $A\boldsymbol{x} = \boldsymbol{b}$ and $C\boldsymbol{x} = \boldsymbol{b}$ have the same solutions, A and C have the same shape and the same nullspace (take $\boldsymbol{b} = \boldsymbol{0}$). If $\boldsymbol{b} = $ column 1 of A, $\boldsymbol{x} = (1, 0, \ldots, 0)$ solves $A\boldsymbol{x} = \boldsymbol{b}$ so it solves $C\boldsymbol{x} = \boldsymbol{b}$. Then A and C share column 1. Other columns too: $\boldsymbol{A} = \boldsymbol{C}$!

33 The column space of R_0 (m by n with rank r) is spanned by its r pivot columns (the first r columns of an m by m identity matrix). The column space of R (after $m - r$ zero rows are removed from R_0) is r-dimensional space $\mathbf{R}^r$.

Problem Set 3.4, page 124

1 $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = 0$ gives $c_3 = c_2 = c_1 = 0$. So those 3 column vectors are independent: no other combination gives $\mathbf{0}$

$\begin{bmatrix} 1 & 1 & 1 & 2 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 1 & 4 \end{bmatrix} \boldsymbol{c} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ is solved by $\boldsymbol{c} = \begin{bmatrix} 1 \\ 1 \\ -4 \\ 1 \end{bmatrix}$. Then $\boldsymbol{v}_1 + \boldsymbol{v}_2 - 4\boldsymbol{v}_3 + \boldsymbol{v}_4 = \mathbf{0}$ (dependent).

2 $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ are independent (the -1's are in different positions). All six vectors in $\mathbf{R}^4$ are on the plane $(1, 1, 1, 1) \cdot \boldsymbol{v} = 0$ so no four of these six vectors can be independent.

3 If $a = 0$ then column $1 = \mathbf{0}$; if $d = 0$ then $b(\text{column } 1) - a(\text{column } 2) = \mathbf{0}$; if $f = 0$ then all columns end in zero (they are all in the xy plane, they must be dependent).

4 $U\boldsymbol{x} = \begin{bmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ gives $z = 0$ then $y = 0$ then $x = 0$ (by back substitution). A square triangular matrix has independent columns (invertible matrix) *when its diagonal has no zeros*.

5 (a) $\begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 3 \\ 0 & -5 & -7 \\ 0 & -1 & -5 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 3 \\ 0 & -5 & -7 \\ 0 & 0 & -18/5 \end{bmatrix}$: invertible $\Rightarrow$ independent columns.

(b) $\begin{bmatrix} 1 & 2 & -3 \\ -3 & 1 & 2 \\ 2 & -3 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & -3 \\ 0 & 7 & -7 \\ 0 & -7 & 7 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & -3 \\ 0 & 7 & -7 \\ 0 & 0 & 0 \end{bmatrix}$; $A \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ columns add to $\mathbf{0}$.

6 Columns 1, 2, 4 are independent. Also 1, 3, 4 and 2, 3, 4 and others (but not 1, 2, 3). Same column numbers (not same columns!) for A. This is because $EA = U$ for the matrix E that subtracts 2 times row 1 from row 4. So A and U have the same nullspace (same dependencies of columns).

7 The sum $v_1 - v_2 + v_3 = 0$ because $(w_2 - w_3) - (w_1 - w_3) + (w_1 - w_2) = 0$. So the

differences are *dependent* and the difference matrix is singular: $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix}$.

8 If $c_1(w_2 + w_3) + c_2(w_1 + w_3) + c_3(w_1 + w_2) = 0$ then $(c_2 + c_3)w_1 + (c_1 + c_3)w_2 + (c_1 + c_2)w_3 = 0$. Since the w's are independent, $c_2 + c_3 = c_1 + c_3 = c_1 + c_2 = 0$. The only solution is $c_1 = c_2 = c_3 = 0$. Only this combination of v_1, v_2, v_3 gives 0.

(changing -1's to 1's for the matrix A in solution **7** above makes A invertible.)

9 (a) The four vectors in $\mathbf{R}^3$ are the columns of a 3 by 4 matrix A. There is a nonzero solution to $Ax = 0$ because there is at least one free variable   (b) Two vectors are dependent if $[\, v_1 \ v_2 \,]$ has rank 0 or 1. (OK to say "they are on the same line" or "one is a multiple of the other" but *not* "v_2 is a multiple of v_1" —since v_1 might be 0.)

(c) A nontrivial combination of v_1 and 0 gives 0: $0v_1 + 3(0,0,0) = (0,0,0)$.

10 The plane is the nullspace of $A = [1 \ 2 \ -3 \ -1]$. Three free variables give three independent solutions $(x, y, z, t) = (-2, 1, 0, 0)$ and $(3, 0, 1, 0)$ and $(1, 0, 0, 1)$. Combinations of those special solutions give more solutions (all solutions).

11 (a) Line in $\mathbf{R}^3$ (b) Plane in $\mathbf{R}^3$ (c) All of $\mathbf{R}^3$ (d) All of $\mathbf{R}^3$.

12 b is in the column space when $Ax = b$ has a solution; c is in the row space when $A^{\mathrm{T}}y = c$ has a solution. *False* because the zero vector is always in the row space.

13 The column space and row space of A and U all have the same dimension = 2. *The row spaces of A and U are the same*, because the rows of U are combinations of the rows of A (and vice versa !).

14 $v = \frac{1}{2}(v + w) + \frac{1}{2}(v - w)$ and $w = \frac{1}{2}(v + w) - \frac{1}{2}(v - w)$. The two pairs *span* the same space. They are a basis for the same space when v and w are *independent*.

15 The n independent vectors span a space of dimension n. They are a *basis* for that space. If they are the columns of A then m is *not less* than n $(m \geq n)$. *Invertible* if $m = n$.

16 These bases are not unique!   (a) $(1,1,1,1)$ for the space of all constant vectors
(c, c, c, c) (b) $(1, -1, 0, 0), (1, 0, -1, 0), (1, 0, 0, -1)$ for the space of vectors with
sum of components $= 0$ (c) $(1, -1, -1, 0), (1, -1, 0, -1)$ for the space perpendic-
ular to $(1, 1, 0, 0)$ and $(1, 0, 1, 1)$ (d) The columns of I are a basis for its column
space, the empty set is a basis (by convention) for $\mathbf{N}(I) = \mathbf{Z} = \{\text{zero vector}\}$.

17 The column space of $U = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix}$ is $\mathbf{R}^2$ so take any bases for $\mathbf{R}^2$;
(row 1 and row 2) or (row 1 and row $1 +$ row 2) or (row 1 and $-$ row 2) are bases
for the row space of U.

18 (a) The 6 vectors *might not* span $\mathbf{R}^4$ (b) The 6 vectors *are not* independent
(c) Any four *might be* a basis.

19 n independent columns $\Rightarrow$ rank n. Columns span $\mathbf{R}^m \Rightarrow$ rank m. Columns are basis
for $\mathbf{R}^m \Rightarrow rank = m = n$. The rank counts the number of *independent* columns.

20 One basis is $(2, 1, 0)$, $(-3, 0, 1)$. A basis for the intersection with the xy plane is
$(2, 1, 0)$. The normal vector $(1, -2, 3)$ is a basis for the line perpendicular to the plane.

21 (a) The only solution to $A\boldsymbol{x} = \mathbf{0}$ is $\boldsymbol{x} = \mathbf{0}$ because *the columns are independent*
(b) $A\boldsymbol{x} = \boldsymbol{b}$ is solvable because *the columns span* $\mathbf{R}^5$. Their combinations give every $\boldsymbol{b}$.
Key point: A basis gives exactly one solution for every $\boldsymbol{b}$.

22 (a) True   (b) False because the basis vectors for $\mathbf{R}^6$ might not be in $\mathbf{S}$.

23 Columns 1 and 2 are bases for the (**different**) column spaces of A and U; rows 1 and
2 are bases for the (**equal**) row spaces of A and U; $(1, -1, 1)$ is a basis for the (**equal**)
nullspaces. **Row spaces and nullspaces** stay fixed in elimination.

24 (a) *False* $A = \begin{bmatrix} 1 & 1 \end{bmatrix}$ has dependent columns, independent row   (b) *False* Column
space $\neq$ row space for $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ (c) *True*: Both dimensions $= 2$ if A is
invertible, dimensions $= 0$ if $A = 0$, otherwise dimensions $= 1$ (d) *False*, columns
may be dependent, in that case not a basis for $\mathbf{C}(A)$.

25 (a) Make $v_1, \ldots v_k$ the columns of A. Then find the first n independent columns (we are told they span $\mathbf{R}^n$).

(b) Make $v_1, \ldots v_j$ the rows of A and then include the n rows of the identity matrix. Row elimination will keep the first j independent rows and find $n - j$ more rows to form a basis for $\mathbf{R}^n$.

26 A has rank 2 if $c = 0$ and $d = 2$; $B = \begin{bmatrix} c & d \\ d & c \end{bmatrix}$ has rank 2 except when $c = d$ or $c = -d$.

27 (a) Basis for all diagonal matrices : $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(b) Add $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ = basis for symmetric matrices.

(c) $\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$.

These are simple bases (among many others) for (a) diagonal matrices (b) symmetric matrices (c) skew-symmetric matrices. The dimensions are $3, 6, 3$.

28 $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$;

Echelon matrices do *not* form a subspace; they **span** the upper triangular matrices (not every U is an echelon matrix).

29 $\begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 1 & 0 \\ 0 & -1 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & -1 \end{bmatrix}$; $\begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix}$.

30 (a) The invertible matrices span the space of all 3 by 3 matrices (b) The rank one matrices also span the space of all 3 by 3 matrices (c) I by itself spans the space of all multiples cI.

31 $\begin{bmatrix} -1 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $\begin{bmatrix} -1 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 0 & 0 \\ -1 & 2 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 2 \end{bmatrix}$. **Dimension $= 4$**.

32 (a) $y(x) =$ constant C (b) $y(x) = 3x$.   (c) $y(x) = 3x + C = y_p + y_n$ solves $y' = 3$.

33 $y(0) = 0$ requires $A + B + C = 0$. One basis is $\cos x - \cos 2x$ and $\cos x - \cos 3x$.

34 (a) $y(x) = e^{2x}$ is a basis for all solutions to $y' = 2y$ (b) $y = x$ is a basis for all solutions to $dy/dx = y/x$ (First-order linear equation $\Rightarrow$ 1 basis function in solution space).

35 $y_1(x), y_2(x), y_3(x)$ can be $x, 2x, 3x$ (dim 1) or $x, 2x, x^2$ (dim 2) or x, x^2, x^3 (dim 3).

36 Basis $1, x, x^2, x^3$, for cubic polynomials; basis $x - 1, x^2 - 1, x^3 - 1$ for the subspace with $p(1) = 0$. (4-dimensional space and 3-dimensional subspace).

37 Basis for **S**: $(1, 0, -1, 0), (0, 1, 0, 0), (1, 0, 0, -1)$; basis for **T**: $(1, -1, 0, 0)$ and $(0, 0, 2, 1)$; **S** $\cap$ **T** $=$ multiples of $(3, -3, 2, 1) =$ nullspace for 3 equations in $\mathbf{R}^4$ has dimension 1.

38 If the 5 by 5 matrix $[\,A \quad b\,]$ is invertible, b is not a combination of the columns of A: no solution to $Ax = b$. If $[\,A \quad b\,]$ is singular, and the 4 columns of A are independent (rank 4), b *is* a combination of those columns. In this case $Ax = b$ has a solution.

39 One basis for $y'' = y$ is $y = e^x$ and $y = e^{-x}$. One basis for $y'' = -y$ is $y = \cos x$ and $y = \sin x$.

40 $\mathbf{I} = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} - \begin{bmatrix} & 1 & \\ & & 1 \\ 1 & & \end{bmatrix} + \begin{bmatrix} & & 1 \\ & 1 & \\ 1 & & \end{bmatrix} + \begin{bmatrix} 1 & & \\ & & 1 \\ & 1 & \end{bmatrix} - \begin{bmatrix} & & 1 \\ 1 & & \\ & 1 & \end{bmatrix}$. The six P's are dependent.

Those five are independent: The 4th has $P_{11} = 1$ and cannot be a combination of the others. Then the 3rd cannot be (from $P_{22} = 1$) and also 1st ($P_{33} = 1$). Continuing, a nonzero combination of all five could not be zero. Further challenge: How many independent 4 by 4 permutation matrices?

41 The dimension of S spanned by all rearrangements of x is   (a) zero when $x = 0$ (b) one when $x = (1, 1, 1, 1)$ (c) three when $x = (1, 1, -1, -1)$ because all rearrangements of this x are perpendicular to $(1, 1, 1, 1)$ (d) four when the x's are not

equal and don't add to zero. **No x gives** dim $S = 2$. I owe this nice problem to Mike Artin—the answers are the same in higher dimensions: 0 or 1 or $n - 1$ or n.

42 The problem is to show that the $\boldsymbol{u}$'s, $\boldsymbol{v}$'s, $\boldsymbol{w}$'s together are independent. We know the $\boldsymbol{u}$'s and $\boldsymbol{v}$'s together are a basis for $\boldsymbol{V}$, and the $\boldsymbol{u}$'s and $\boldsymbol{w}$'s together are a basis for $\boldsymbol{W}$. Suppose a combination of $\boldsymbol{u}$'s, $\boldsymbol{v}$'s, $\boldsymbol{w}$'s gives $\boldsymbol{0}$. **To be proved**: All coefficients $=$ zero.

Key idea: In that combination giving $\boldsymbol{0}$, the part $\boldsymbol{x}$ from the $\boldsymbol{u}$'s and $\boldsymbol{v}$'s is in $\boldsymbol{V}$. So the part from the $\boldsymbol{w}$'s is $-\boldsymbol{x}$. This part is now in $\boldsymbol{V}$ and also in $\boldsymbol{W}$. But if $-\boldsymbol{x}$ is in $\boldsymbol{V} \cap \boldsymbol{W}$ it is a combination of $\boldsymbol{u}$'s only. Now the combination giving $\boldsymbol{0}$ uses only $\boldsymbol{u}$'s and $\boldsymbol{v}$'s (independent in $\boldsymbol{V}$!) so all coefficients of $\boldsymbol{u}$'s and $\boldsymbol{v}$'s must be zero. Then $\boldsymbol{x} = \boldsymbol{0}$ and the coefficients of the $\boldsymbol{w}$'s are also zero.

43 If the left side of $\dim(\mathbf{V}) + \dim(\mathbf{W}) = \dim(\mathbf{V} \cap \mathbf{W}) + \dim(\mathbf{V} + \mathbf{W})$ is greater than n, then $\dim(\mathbf{V} \cap \mathbf{W})$ must be greater than zero. So $\mathbf{V} \cap \mathbf{W}$ contains nonzero vectors.

Here is a more basic approach: Put a basis for $\mathbf{V}$ and then a basis for $\mathbf{W}$ in the columns of a matrix A. Then A has more columns than rows and there is a nonzero solution to $A\boldsymbol{x} = \boldsymbol{0}$. That $\boldsymbol{x}$ gives a combination of the $\mathbf{V}$ columns $=$ a combination of the $\mathbf{W}$ columns.

44 If $A^2 = $ zero matrix, this says that each column of A is in the nullspace of A. If the column space has dimension r, the nullspace has dimension $10 - r$ by the Counting Theorem. So we must have $r \leq 10 - r$ and this leads to $r \leq 5$.

Problem Set 3.5, page 137

1 (a) Row and column space dimensions $9-5=5$, nullspace dimension $=4$, $\dim(\mathbf{N}(A^{\mathrm{T}}))$

$= 9 - 7 = 2$ sum $5 + 5 + 4 + 2 = 16 = m + n$

(b)   Column space is $\mathbf{R}^3$; left nullspace contains only $\mathbf{0}$ (dimension zero).

2 A: Row space basis $=$ row $1 = (1, 2, 4)$; nullspace $(-2, 1, 0)$ and $(-4, 0, 1)$; column

space basis $=$ column $1 = (1, 2)$; left nullspace $(-2, 1)$. B: Row space basis $=$

both rows $= (1, 2, 4)$ and $(2, 5, 8)$; column space basis $=$ two columns $= (1, 2)$ and

$(2, 5)$; nullspace $(-4, 0, 1)$; left nullspace basis is empty because the space contains

only $\boldsymbol{y} = \mathbf{0}$: the rows of B are independent.

3 Row space basis $=$ first two rows of R; column space basis $=$ pivot columns (of A not R)

$= (1, 1, 0)$ and $(3, 4, 1)$; nullspace basis $(1, 0, 0, 0, 0)$, $(0, 2, -1, 0, 0)$, $(0, 2, 0, -2, 1)$;

left nullspace $(1, -1, 1) =$ last row of the elimination matrix $E^{-1} = L$.

4 (a) $\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$ (b) Impossible: $r+(n-r)$ must be 3   (c) $\begin{bmatrix} 1 & 1 \end{bmatrix}$ (d) $\begin{bmatrix} 9 & -3 \\ 3 & -1 \end{bmatrix}$

(e) *Impossible* Row space $=$ column space requires $m = n$. Then $m - r = n - r$;

nullspaces have the same dimension.   Section 4.1 will prove $\mathbf{N}(A)$ and $\mathbf{N}(A^{\mathrm{T}})$

orthogonal to the row and column spaces respectively—here those are the same space.

5 $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 0 \end{bmatrix}$ has those rows spanning its row space. $B = \begin{bmatrix} 1 & -2 & 1 \end{bmatrix}$ has the

same vectors spanning its nullspace and $AB^{\mathrm{T}} =$ zero matrix (*not AB*).

6 A: dim $\mathbf{2, 2, 2, 1}$: Rows $(0, 3, 3, 3)$ and $(0, 1, 0, 1)$; columns $(3, 0, 1)$ and $(3, 0, 0)$;

nullspace $(1, 0, 0, 0)$ and $(0, -1, 0, 1)$; $\mathbf{N}(A^{\mathrm{T}})$ $(0, 1, 0)$. B: dim $\mathbf{1, 1, 0, 2}$ Row space

(1), column space $(1, 4, 5)$, nullspace: empty basis, $\mathbf{N}(A^{\mathrm{T}})$ $(-4, 1, 0)$ and $(-5, 0, 1)$.

7 Invertible 3 by 3 matrix A: row space basis = column space basis = $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$; nullspace basis and left nullspace basis are *empty*. Matrix $B = \begin{bmatrix} A & A \end{bmatrix}$: row space basis $(1, 0, 0, 1, 0, 0)$, $(0, 1, 0, 0, 1, 0)$ and $(0, 0, 1, 0, 0, 1)$; column space basis $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$; nullspace basis $(-1, 0, 0, 1, 0, 0)$ and $(0, -1, 0, 0, 1, 0)$ and $(0, 0, -1, 0, 0, 1)$; left nullspace basis is empty.

8 $\begin{bmatrix} I & 0 \end{bmatrix}$ and $\begin{bmatrix} I & I; & 0^{\mathrm{T}} & 0^{\mathrm{T}} \end{bmatrix}$ and $\begin{bmatrix} 0 \end{bmatrix}$ = 3 by 2 have *row space dimensions* = $3, 3, 0$ = column space dimensions; *nullspace dimensions* $2, 3, 2$; *left nullspace dimensions* $0, 2, 3$.

9 (a) Same row space and nullspace. So rank (dimension of row space) is the same

(b) Same column space and left nullspace. Same rank (dimension of column space).

10 For **rand**(3), almost surely rank= 3, nullspace and left nullspace contain only $(0, 0, 0)$. For **rand**$(3, 5)$ the rank is almost surely 3 and the dimension of the nullspace is 2.

11 (a) No solution means that $r < m$. Always $r \leq n$. Can't compare m and n here.

(b) Since $m - r > 0$, the left nullspace must contain a nonzero vector.

12 A neat choice is $\begin{bmatrix} 1 & 1 \\ 0 & 2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 1 \\ 2 & 4 & 0 \\ 1 & 0 & 1 \end{bmatrix}$; $r + (n - r) = n = 3$ does not match $2 + 2 = 4$. Only $\boldsymbol{v} = \mathbf{0}$ is in both $\mathbf{N}(A)$ and $\mathbf{C}(A^{\mathrm{T}})$.

13 (a) *False*: Usually row space $\neq$ column space.

(b) *True*: A and $-A$ have the same four subspaces

(c) *False* (choose A and B same size and invertible: then they have the same four subspaces)

14 Row space basis can be the nonzero rows of U: $(1, 2, 3, 4)$, $(0, 1, 2, 3)$, $(0, 0, 1, 2)$; nullspace basis $(0, 1, -2, 1)$ as for U; column space basis $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ (happen to have $\mathbf{C}(A) = \mathbf{C}(U) = \mathbf{R}^3$); left nullspace has empty basis.

15 After a row exchange, the row space and nullspace stay the same; $(2, 1, 3, 4)$ is in the new left nullspace after the row exchange.

16 If $A\boldsymbol{v} = \mathbf{0}$ and $\boldsymbol{v}$ is a row of A then $\boldsymbol{v} \cdot \boldsymbol{v} = 0$. So $\boldsymbol{v}$ is perpendicualr to $\boldsymbol{v}$: $\boldsymbol{v} = \mathbf{0}$.

17 Row space of $A = yz$ plane; column space of $A = xy$ plane; nullspace $= x$ axis; left nullspace $= z$ axis. For $I + A$: Row space $=$ column space $= \mathbf{R}^3$, both nullspaces contain only the zero vector.

18 $a_{11} = 1, a_{12} = 0, a_{13} = 1, a_{22} = 0, a_{32} = 1, a_{31} = 0, a_{23} = 1, a_{33} = 0, a_{21} = 1$.
(Need to specify the five moves).

19 Row $3 - 2$ row $2 +$ row $1 =$ zero row so the vectors $c(1, -2, 1)$ are in the left nullspace. The same vectors happen to be in the nullspace (an accident for this matrix).

20 The steps from A to R_0 are described on page 96 (Section 3.2). I don't think I can do better—but you could put those ideas into different words. By all means give an example that needs row exchanges.

21 (a) $\boldsymbol{u}$ and $\boldsymbol{w}$ (b) $\boldsymbol{v}$ and $\boldsymbol{z}$ (c) rank < 2 if $\boldsymbol{u}$ and $\boldsymbol{w}$ are dependent or if $\boldsymbol{v}$ and $\boldsymbol{z}$ are dependent   (d) The rank of $\boldsymbol{uv}^{\mathrm{T}} + \boldsymbol{wz}^{\mathrm{T}}$ is 2.

22 $A = \begin{bmatrix} \boldsymbol{u} & \boldsymbol{w} \end{bmatrix} \begin{bmatrix} \boldsymbol{v}^{\mathrm{T}} \\ \boldsymbol{z}^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 2 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 4 & 2 \\ 5 & 1 \end{bmatrix}$ $\boldsymbol{u}, \boldsymbol{w}$ span column space; $\boldsymbol{v}, \boldsymbol{z}$ span row space

23 As in Problem 22: Row space basis $(3, 0, 3), (1, 1, 2)$; column space basis $(1, 4, 2)$, $(2, 5, 7)$; the rank of (3 by 2) times (2 by 3) cannot be larger than the rank of either factor, so rank ≤ 2 and the 3 by 3 product is not invertible.

24 $A^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{d}$ puts $\boldsymbol{d}$ in the *row space* of A; unique solution if the *left nullspace* (nullspace of A^{T}) contains only $\boldsymbol{y} = \boldsymbol{0}$.

25 (a) *True* (A and A^{T} have the same rank)   (b) *False* $A = \begin{bmatrix} 1 & 0 \end{bmatrix}$ and A^{T} have very different left nullspaces   (c) *False* (A can be invertible and unsymmetric even if $C(A) = C(A^{\mathrm{T}})$)   (d) *True* (The subspaces for A and $-A$ are always the same. If $A^{\mathrm{T}} = A$ or $A^{\mathrm{T}} = -A$ they are also the same for A^{T})

26 Choose $d = bc/a$ to make $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ a rank-1 matrix. Then the row space has basis (a, b) and the nullspace has basis $(-b, a)$. Those two vectors are perpendicular !

27 B and C (checkers and chess) both have rank 2 if $p \neq 0$. Row 1 and 2 are a basis for the row space of C, $B^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{0}$ has 6 special solutions with -1 and 1 separated by a zero; $\mathbf{N}(C^{\mathrm{T}})$ has $(-1, 0, 0, 0, 0, 0, 0, 1)$ and $(0, -1, 0, 0, 0, 0, 1, 0)$ and columns $3, 4, 5, 6$ of I; $\mathbf{N}(C)$ is a challenge : one vector in $\mathbf{N}(C)$ is $(1, 0, \ldots, 0, -1)$.

28 The subspaces for $A = \boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ are pairs of orthogonal lines ($\boldsymbol{v}$ and $\boldsymbol{v}^{\perp}$, $\boldsymbol{u}$ and $\boldsymbol{u}^{\perp}$). If B has those same four subspaces then $B = cA$ with $c \neq 0$.

29 (a) $AX = 0$ if each column of X is a multiple of $(1, 1, 1)$; $\dim(\text{nullspace}) = 3$.

(b) If $AX = B$ then all columns of B add to zero; dimension of the B's $= 6$.

(c) $3 + 6 = \dim(M^{3 \times 3}) = 9$ entries in a 3 by 3 matrix.

30 The key is equal row spaces. First row of $A =$ combination of the rows of B : the only possible combination (notice I) is 1 (row 1 of B). Same for each row so $F = G$.

31 $A = \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 \\ -1 & 0 & 0 & 1 \end{bmatrix}$ $\mathbf{N}(A) \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ Row space $\mathbf{C}(A^{\mathrm{T}}) \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$

$\mathbf{C}(A) \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ $\mathbf{N}(A^{\mathrm{T}}) \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

32 (a) $\mathbf{N}(BA)$ contains $\mathbf{N}(A)$.

(b) $\mathbf{C}(AB)$ is contained in $\mathbf{C}(A)$.

33 (a) $\mathbf{N}(A)$ and $\mathbf{N}(B)$ contain $\mathbf{N}(T)$.

(b) Row spaces of A and B are contained in the row space of T.

34 Fundamental subspaces for $A \quad (m \times n)$.

Row space $\mathbf{C}(A^{\mathrm{T}})$ perpendicular to Nullspace $\mathbf{N}(A)$: Dimensions r and $n - r$.

Column space $\mathbf{C}(A)$ perpendicular to $\mathbf{N}(A^{\mathrm{T}})$: Dimensions r and $m - r$.

Subspaces for $W = \begin{bmatrix} A & A \end{bmatrix}$ − same rank r.

Row space of W contains all $\begin{bmatrix} \boldsymbol{v} & \boldsymbol{v} \end{bmatrix} \quad \boldsymbol{v}$ in $\mathbf{C}(A^{\mathrm{T}})$ (Dimension r).

Nullspace of W contains all $\begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{z} \end{bmatrix}$ with $\boldsymbol{y} + \boldsymbol{z}$ in $\mathbf{N}(W)$ (Dimension $2n - r$).

Column space of $W = $ Column space of A (Dimension r).

Nullspace of $W^{\mathrm{T}} = $ Nullspace of A (Dimension $m - r$).

35 Please send a proof or counterexample. Thank you.

Problem Set 4.1, page 148

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 $x_n = \mathbf{0}$, 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 $A = \begin{bmatrix} 1 & 2 & -3 \\ 2 & -3 & 1 \\ -3 & 5 & -2 \end{bmatrix}$
and make col $3 = -$ col $1-$ col 2.

(b) Impossible because $\mathbf{N}(A)$ and $\mathbf{C}(A^{\mathrm{T}})$ 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 $\mathbf{C}(A)$ and $\mathbf{N}(A^{\mathrm{T}})$ is impossible: not perpendicular

(d) Rows orthogonal to columns makes A times $A =$ zero matrix. 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 $\mathbf{C}(A)$ is orthogonal to $\mathbf{N}(A^{\mathrm{T}})$.

(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 $0 = (\boldsymbol{y}^{\mathrm{T}} A)\boldsymbol{x} = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{b}$ 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 $\mathbf{N}(A^{\mathrm{T}})$ 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 4.1 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 $\mathbf{C}(A)$.

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 space and row space 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 $\mathbf{N}(A^{\mathrm{T}})$ 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 $\boldsymbol{x}_r + \boldsymbol{x}_n = (1, -1) + (1, 1)$.

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

14 $A\boldsymbol{x} = B\widehat{\boldsymbol{x}}$ means that $\begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ -\widehat{\boldsymbol{x}} \end{bmatrix} = \boldsymbol{0}$. 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 $p + q > n$. (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 $\boldsymbol{W}$.)

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

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 $(1, 1, 1)$ and $(1, 1, -1)$, then $\mathbf{S}^{\perp}$ is the line spanned by $(1, -1, 0)$.

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

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 the hyperplane $\mathbf{P}$.

$A = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$ has $\mathbf{P}$ as its nullspace and $\mathbf{P}^{\perp}$ as its row space.

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

24 $AA^{-1} = I$: Column 1 of A^{-1} is orthogonal to rows $2, 3, \ldots, n$ of A and therefore it is orthogonal 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 $(-\mathbf{2}, \mathbf{1})$. The nullspace of the 2 by 2

matrix is the line $3x + y = \mathbf{0}$. One particular vector in the nullspace is $(-\mathbf{1}, \mathbf{3})$.

28 (a) $(1, -1, 0)$ is in both planes. Normal vectors are perpendicular to each other,

but planes can 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 $\mathbf{C}(B) \leq$

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

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

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

Problem Set 4.2, page 159

0 (a) $P^2 = P$ is true but $P^{\mathrm{T}} \neq P$. This question is about nonsymmetric projections (the error e is no longer perpendicular to the projection p). Same for $I - P$.

(b) $v = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is in $\mathbf{C}(P)$ and $w = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is in $\mathbf{C}(I - P)$. But v is not perpendicular to w.

(c) If $P^2 = P$ and $P^{\mathrm{T}} = P$, then the column spaces of P and $I - P$ are perpendicular:

$$(Pv)^{\mathrm{T}}(I - P)w = v^{\mathrm{T}}P(I - P)w = 0 \text{ since } P - P^2 = 0.$$

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

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

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

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

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

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

4 $P_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $P_2 = \dfrac{aa^{\mathrm{T}}}{a^{\mathrm{T}}a} = \dfrac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$. 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 $P_1 = \dfrac{1}{9} \begin{bmatrix} 1 & -2 & -2 \\ -2 & 4 & 4 \\ -2 & 4 & 4 \end{bmatrix}$ and $P_2 = \dfrac{1}{9} \begin{bmatrix} 4 & 4 & -2 \\ 4 & 4 & -2 \\ -2 & -2 & 1 \end{bmatrix}$.

P_1 and P_2 are the projection matrices onto the lines through $a_1 = (-1, 2, 2)$ and $a_2 = (2, 2, -1)$. $P_1 P_2 =$ *zero matrix because* $a_1 \perp a_2$.

6 $p_1 = (\frac{1}{9}, -\frac{2}{9}, -\frac{2}{9})$ and $p_2 = (\frac{4}{9}, \frac{4}{9}, -\frac{2}{9})$ and $p_3 = (\frac{4}{9}, -\frac{2}{9}, \frac{4}{9})$. So $p_1 + p_2 + p_3 = b$.

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

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 Here P is the usual projection $(P^2 = P = P^{\mathrm{T}})$ of the whole space $\mathbf{R}^m$ onto its subspace **S**.

S is the row space and column space of P.

$S^\perp$ is the nullspace of P and P^{T}.

12 (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.

13 $P_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} =$ 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.

14 $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, $P = $ square matrix $= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, $\boldsymbol{p} = P \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 0 \end{bmatrix}$.

15 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 $\boldsymbol{b} = 2$(column 1 of A) :

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

16 $2A$ 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.

17 $\frac{1}{2}(1, 2, -1) + \frac{3}{2}(1, 0, 1) = (2, 1, 1)$. So $\boldsymbol{b}$ is in the plane. Projection shows $P\boldsymbol{b} = \boldsymbol{b}$.

18 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*.

19 (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 $(1, 1, 1)$.

20 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 $\begin{bmatrix} 5/6 & 1/6 & 1/3 \\ 1/6 & 5/6 & -1/3 \\ 1/3 & -1/3 & 1/3 \end{bmatrix}$.

21 $\boldsymbol{e} = \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}$, $Q = \frac{\boldsymbol{e}\boldsymbol{e}^{\mathrm{T}}}{\boldsymbol{e}^{\mathrm{T}}\boldsymbol{e}} = \begin{bmatrix} 1/6 & -1/6 & -1/3 \\ -1/6 & 1/6 & 1/3 \\ -1/3 & 1/3 & 2/3 \end{bmatrix}$, $I - Q = \begin{bmatrix} 5/6 & 1/6 & 1/3 \\ 1/6 & 5/6 & -1/3 \\ 1/3 & -1/3 & 1/3 \end{bmatrix}$.

22 $\left(A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\right)^2 = A(A^{\mathrm{T}}A)^{-1}(A^{\mathrm{T}}A)(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}} = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$. So $P^2 = P$. $P\boldsymbol{b}$ is in the column space (where P projects). Then its projection $P(P\boldsymbol{b})$ is also $P\boldsymbol{b}$.

23 $P^{\mathrm{T}} = (A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}})^{\mathrm{T}} = A((A^{\mathrm{T}}A)^{-1})^{\mathrm{T}}A^{\mathrm{T}} = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}} = P$. ($A^{\mathrm{T}}A$ is symmetric!)

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

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

26 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 $S = n$.

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

28 If $A^{\mathrm{T}}A\boldsymbol{x} = \boldsymbol{0}$ 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*: $A^{\mathrm{T}}A\boldsymbol{x} = \boldsymbol{0}$ exactly when $A\boldsymbol{x} = \boldsymbol{0}$.

29 Start from $P^2 = P$. The $(2,2)$ entry of P^2 is the dot product (row 2 of P) $\cdot$ (column 2 of P). Since P is symmetric this is $||\text{row } 2 \text{ of } P||^2$.

$$P = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \text{ has } P^2 = P \text{ and } P_{22} = \frac{1}{2} \text{ and column } 2 = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} \text{ has } ||\text{column } 2||^2 = \frac{1}{2}$$

30 If $BB^{\mathrm{T}}\boldsymbol{x} = \boldsymbol{0}$ then $0 = \boldsymbol{x}^{\mathrm{T}}BB^{\mathrm{T}}\boldsymbol{x} = ||B^{\mathrm{T}}\boldsymbol{x}||^2$. But then $B^{\mathrm{T}}\boldsymbol{x} = \boldsymbol{0}$. Since B has independent rows, this only happens if $\boldsymbol{x} = \boldsymbol{0}$. So BB^{T} is invertible.

Problem Set 4.3, page 161

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 $\begin{bmatrix} 1 \\ 5 \\ 13 \\ 17 \end{bmatrix}$; When $\boldsymbol{p}$ replaces $\boldsymbol{b}$, Project $\boldsymbol{b}$ to $\boldsymbol{p} = P\boldsymbol{b} =$

$\widehat{\boldsymbol{x}} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$ 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 $a = (1, 1, 1, 1)$ and $b = (0, 8, 8, 20)$ give $\hat{x} = a^{\mathrm{T}}b/a^{\mathrm{T}}a = 9$ and the projection is $\hat{x}a = p = (9, 9, 9, 9)$. Then $e^{\mathrm{T}}a = (-9, -1, -1, 11)^{\mathrm{T}}(1, 1, 1, 1) = 0$ and the shortest distance from b to the line through a is $\|e\| = \sqrt{204}$.

7 Now the 4 by 1 matrix in $Ax = b$ is $A = \begin{bmatrix} 0 & 1 & 3 & 4 \end{bmatrix}^{\mathrm{T}}$. Then $A^{\mathrm{T}}A = \begin{bmatrix} 26 \end{bmatrix}$ and $A^{\mathrm{T}}b = \begin{bmatrix} 112 \end{bmatrix}$. Best $D = 112/26 = 56/13$.

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

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

10
$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 3 & 9 & 27 \\ 1 & 4 & 16 & 64 \end{bmatrix} \begin{bmatrix} C \\ D \\ E \\ F \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ 8 \\ 20 \end{bmatrix} . \quad \text{Then} \quad \begin{bmatrix} C \\ D \\ E \\ F \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 0 \\ 47 \\ -28 \\ 5 \end{bmatrix} .$$
Exact cubic so $p = b$, $e = 0$.
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 $\hat{b} = 9$ at center time, $\hat{t} = 2$.

(b) The first equation $Cm + D \sum t_i = \sum b_i$ divided by m gives $C + D\hat{t} = \hat{b}$. This shows : The best line goes through $\hat{b}$ at time $\hat{t}$.

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

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

(c) $p = (3, 3, 3)$ and $e = (-2, -1, 3)$ $p^{\mathrm{T}}e = 0$. 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.

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 $\widehat{\boldsymbol{x}} = \begin{bmatrix} 9 \\ 4 \end{bmatrix}$ comes from $\begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix}\begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 35 \\ 42 \end{bmatrix}$.

18 $\boldsymbol{p} = A\widehat{\boldsymbol{x}} = (5, 13, 17)$ 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 $\boldsymbol{b} = 9$ (column 1) $+ 4$ (column 2) is *in the column space of* A.

21 $\boldsymbol{e}$ is in $\mathbf{N}(A^{\mathrm{T}})$; $\boldsymbol{p}$ is in $\mathbf{C}(A)$; $\widehat{\boldsymbol{x}}$ is in $\mathbf{C}(A^{\mathrm{T}})$; $\mathbf{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 $\|\boldsymbol{e}\|^2 = \boldsymbol{e}^{\mathrm{T}}(\boldsymbol{b} - \boldsymbol{p}) = \boldsymbol{e}^{\mathrm{T}}\boldsymbol{b} = \boldsymbol{b}^{\mathrm{T}}\boldsymbol{b} - \boldsymbol{b}^{\mathrm{T}}\boldsymbol{p}$.

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 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 $\boldsymbol{y} = (t_2 - t_3, t_3 - t_1, t_1 - t_2)$ 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

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} C \\ D \\ E \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 3 \\ 4 \end{bmatrix} . \; A^{\mathrm{T}}A = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} ; A^{\mathrm{T}}\boldsymbol{b} = \begin{bmatrix} 8 \\ -2 \\ -3 \end{bmatrix} ; \begin{bmatrix} C \\ D \\ E \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ -3/2 \end{bmatrix} .$$

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 invertible 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 $\boldsymbol{a}_1, \dots, \boldsymbol{a}_{n-1}$. If they are dependent, there is a vector $\boldsymbol{v}$ perpendicular to all the $\boldsymbol{a}$'s. Then they all (including $\mathbf{0}$) lie on the plane $\boldsymbol{v}^{\mathrm{T}}\boldsymbol{x} = 0$ going through $\boldsymbol{x} = (0, 0, \dots, 0)$.

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

Problem Set 4.4, page 186

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

$q_1 = (\frac{2}{3}, \frac{2}{3}, -\frac{1}{3})$. $q_2 = (-\frac{1}{3}, \frac{2}{3}, \frac{2}{3})$.

$Q^{\mathrm{T}}Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ 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 $q_1 = (1, 1, 1)/\sqrt{3}$ my favorite is $q_2 = (1, -1, 0)/\sqrt{2}$ and $q_3 = (1, 1, -2)/\sqrt{6}$.

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 $Q_1 Q_2$ is orthogonal because $(Q_1 Q_2)^{\mathrm{T}} Q_1 Q_2 = Q_2^{\mathrm{T}} Q_1^{\mathrm{T}} Q_1 Q_2 = Q_2^{\mathrm{T}} Q_2 = I$. Another approach is to see that $(Q_1 Q_1)^{-1} = Q_2^{-1} Q_1^{-1} = Q_2^{\mathrm{T}} Q_1^{\mathrm{T}} = (Q_1 Q_2)^{\mathrm{T}}$.

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

8 If q_1 and q_2 are *orthonormal* vectors in $\mathbf{R}^5$ then $p = (q_1^{\mathrm{T}} b) q_1 + (q_2^{\mathrm{T}} b) q_2$ is closest to b. The error $e = b - p$ is orthogonal to q_1 and 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 q_1, q_2, q_3 are *orthonormal* then the dot product of q_1 with $c_1 q_1 + c_2 q_2 + c_3 q_3 =$ **0** gives $c_1 = 0$. Similarly $c_2 = c_3 = 0$. This proves : *Independent q's*

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

11 (a) Two *orthonormal* vectors are $q_1 = \frac{1}{10}(1, 3, 4, 5, 7)$ and $q_2 = \frac{1}{10}(-7, 3, 4, -5, 1)$

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

12 Multiply $b = x_1 a_1 + x_2 a_2 + x_3 a_3$ by $a x_1^{\mathrm{T}}$:

$$a_1^{\mathrm{T}} b = x_1 a_1^{\mathrm{T}} a_1 + x_2 a_1^{\mathrm{T}} a_2 + x_3 a_1^{\mathrm{T}} a_3 = 0 + 0 + x_1 a_1^{\mathrm{T}} a_1$$

Divide by $a_1^{\mathrm{T}} a_1$ (not necessarily equal to 1) to find $x_1 = a_1^{\mathrm{T}} b / a_1^{\mathrm{T}} a_1$.

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

14 $\begin{bmatrix} 1 & 4 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} q_1 & q_2 \end{bmatrix} \begin{bmatrix} \|a\| & q_1^{\mathrm{T}} b \\ 0 & \|B\| \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{2} & 2\sqrt{2} \\ 0 & 2\sqrt{2} \end{bmatrix} = QR$.

15 (a) Gram-Schmidt chooses $q_1 = a/\|a\| = \frac{1}{3}(1, 2, -2)$ and $q_2 = \frac{1}{3}(2, 1, 2)$. Then $q_3 = \frac{1}{3}(2, -2, -1)$.

(b) The nullspace of A^{T} contains q_3

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

16 $p = (a^{\mathrm{T}} b / a^{\mathrm{T}} a) a = 14a/49 = 2a/7$ is the projection of b onto a. $q_1 = a/\|a\| = a/7$ is $(4, 5, 2, 2)/7$. $B = b - p = (-1, 4, -4, -4)/7$ has $\|B\| = 1$ so $q_2 = B$.

17 $p = (a^{\mathrm{T}} b / a^{\mathrm{T}} a) a = (3, 3, 3)$ and $e = (-2, 0, 2)$. Then Gram-Schmidt will choose $q_1 = (1, 1, 1)/\sqrt{3}$ and $q_2 = (-1, 0, 1)/\sqrt{2}$.

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

Gram-Schmidt would go on to normalize $q_1 = A/\|A\|, q_2 = B/\|B\|, q_3 = C/\|C\|$.

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** From $Q^{\mathrm{T}} = Q^{-1}$ we find $(Q^{\mathrm{T}})^{-1} = (Q^{-1})^{-1}$. And always $(Q^{\mathrm{T}})^{-1} = (Q^{-1})^{\mathrm{T}}$.

(b) **True** So the transpose of $Q^{-1} = $ inverse of Q^{-1}. We are given orthonormal columns: $Q^{\mathrm{T}}Q = I$. Then $||Qx||^2 = x^{\mathrm{T}}Q^{\mathrm{T}}Qx = x^{\mathrm{T}}x = ||x||^2$.

Notice: Q might be rectangular (and then not an orthogonal matrix).

21 The orthonormal vectors are $q_1 = (1,1,1,1)/2$ and $q_2 = (-5,-1,1,5)/\sqrt{52}$. Then $b = (-4,-3,3,0)$ projects to $p = (q_1^{\mathrm{T}}b)q_1 + (q_2^{\mathrm{T}}b)q_2 = (-7,-3,-1,3)/2$. And $b - p = (-1,-3,7,-3)/2$ is orthogonal to both q_1 and q_2.

22 $A = (1,1,2)$, $B = (1,-1,0)$, $C = (-1,-1,1)$. These are not yet unit vectors. Gram-Schmidt will divide by $||A|| = \sqrt{6}$ and $||B|| = \sqrt{2}$ and $||C|| = \sqrt{3}$.

23 You can see why $q_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, q_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, q_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$. $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) 1 equation, 4 unknowns, 3 independent solutions.

We could choose the solutions $(1,0,0,1)$, $(0,1,0,1)$, and $(0,0,1,1)$.

(b) Those vectors are orthogonal to $S^{\perp} = $ line through $(1,1,1-1)$.

(c) The component of $b = (1,1,1,1)$ in $S^{\perp}$ is

$$b_2 = \frac{(1,1,1,1)^{\mathrm{T}}(1,1,1,-1)}{(1,1,1,-1)^{\mathrm{T}}(1,1,1,-1)}(1,1,1,-1) = \frac{2}{4}(1,1,1,-1) = \frac{1}{2}(1,1,1,-1)$$

Then $b_1 = b - b_2 = (1,1,1,1) - \dfrac{1}{2}(1,1,1,-1) = \dfrac{1}{2}(1,1,1,3)$. Check $1+1+1-3 = 0$.

25
$$\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = \frac{\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}}{\sqrt{5}} \ \frac{\begin{bmatrix} 5 & 3 \\ 0 & 1 \end{bmatrix}}{\sqrt{5}}$$

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \frac{\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}}{\sqrt{2}} \ \frac{\begin{bmatrix} 2 & 2 \\ 0 & 0 \end{bmatrix}}{\sqrt{2}} \text{ and } R_{22} \text{ has become zero. } A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$
had dependent columns.

26 $(q_2^{\mathrm{T}} C^*) q_2 = \dfrac{B^{\mathrm{T}} c}{B^{\mathrm{T}} B} B$ because $q_2 = \dfrac{B}{\|B\|}$ and the extra q_1 in C^* is orthogonal to q_2.

27 When a and b are not orthogonal, the projections onto these lines *do not add* to the projection onto the plane of a and b. We must use the orthogonal A and B (or orthonormal q_1 and 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 $q_1 = \frac{1}{3}(2, 2, -1)$, $q_2 = \frac{1}{3}(2, -1, 2)$, $q_3 = \frac{1}{3}(1, -2, -2)$.

30 W has orthonormal columns so $W^{\mathrm{T}} W = I$ and $W^{\mathrm{T}} = W^{-1}$.

31 Choose $c = \frac{1}{2}$ to give orthonormal columns. The projection of $b = (1, 1, 1, 1)$ onto the first column $q_1 = \frac{1}{2}(1, -1, -1, -1)$ is $(q_1^{\mathrm{T}} b) q_1 = -q_1$.

The projection of b on q_2 is $(q_2^{\mathrm{T}} b) q_2 = -q_2$.

Since q_1 is orthogonal to q_2, we add to find the projection $-q_1 - q_2$ onto the plane of q_1 and q_2.

32 $Q = I - 2uu^{\mathrm{T}}$ is a **reflection matrix** $(Q^2 = I)$ if u is a unit vector $\|u\| = 1$.

$$u = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \text{ gives } Q = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \qquad u = \begin{bmatrix} 0 \\ \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix} \text{ gives } Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix}.$$

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

34 (a) $Qu = (I - 2uu^{\mathrm{T}})u = u - 2uu^{\mathrm{T}} u$. This is $-u$, provided that $u^{\mathrm{T}} u$ equals 1

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

35 Starting from $A = (1, -1, 0, 0)$, the orthogonal (not orthonormal) vectors $B = (1, 1, -2, 0)$ and $C = (1, 1, 1, -3)$ and $D = (1, 1, 1, 1)$ are in the directions of q_2, q_3, 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 q_{n+1} in Gram-Schmidt using the matrix Q with the columns $q_1, \dots, q_n$ (instead of using those q's separately). Start from a, *subtract its projection* $p = QQ^{\mathrm{T}}a$ onto the earlier q's, *divide by the length* of $e = a - QQ^{\mathrm{T}}a$ to get the next $q_{n+1} = e/\|e\|$.

Problem Set 4.5, page 196

1 $A_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix} = CR$ so $A_1^+ = R^{\mathrm{T}}(C^{\mathrm{T}}AR^{\mathrm{T}})^{-1}C^{\mathrm{T}}$ from

equation (7), page 195.

$$A_1^+ = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \frac{1}{4} \begin{bmatrix} 1 & 1 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \qquad A_2^+ = \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 1/4 & 0 \end{bmatrix} \qquad A_3^+ = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/4 \\ 0 & 0 \end{bmatrix}$$

2 $A^+A = \dfrac{r c^{\mathrm{T}} c r^{\mathrm{T}}}{(r^{\mathrm{T}}r)\,(c^{\mathrm{T}}c)} = \dfrac{r r^{\mathrm{T}}}{r^{\mathrm{T}}r} \quad$ and $\quad AA^+ = \dfrac{c c^{\mathrm{T}}}{c^{\mathrm{T}}c}$ are symmetric

The first two Penrose conditions are also easy to check :

$$AA^+A = (c r^{\mathrm{T}})\frac{r r^{\mathrm{T}}}{r^{\mathrm{T}}r} = c r^{\mathrm{T}} = A \quad \text{and} \quad A^+AA^+ = \frac{r r^{\mathrm{T}}}{r^{\mathrm{T}}r}\,\frac{r c^{\mathrm{T}}}{(r^{\mathrm{T}}r)\,(c^{\mathrm{T}}c)} = A^+$$

3 Problem : Check that $B = \begin{bmatrix} A \\ 0 \end{bmatrix}$ has $B^+ = \begin{bmatrix} A^+ & 0 \end{bmatrix}$.

Solution : Start from $B^+B = A^+A$. Then the Penrose conditions are satisfied by B^+.

4 The column space of A is $\mathbf{R}^2$. The row space is the x-y plane in $\mathbf{R}^3$. Then $A^+ = A^{\mathbf{T}}$.

5 $A^+ = \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 1/2 & 0 \end{bmatrix} \quad$ and $\quad (A^{\mathrm{T}})^+ = \begin{bmatrix} 0 & 0 & 1/2 \\ 1 & 0 & 0 \end{bmatrix}$.

It is always true that $(A^{\mathrm{T}})^+$ is the transpose of A^+. The straightforward proof quickly checks the Penrose conditions. We are simply reversing the left side and right side of the "Big Picture" of 4 subspaces.

6 Given that $P^{\mathrm{T}} = P = P^2$, the pseudoinverse P^+ is the same as P. The first two Penrose conditions become $P^3 = P$ (**true**). The last two conditions become $(P^2)^{\mathrm{T}} = P^{\mathrm{T}} = P$ (**true**).

7 We are asked to verify $(AA^+)^2 = AA^+$. Watch $C^T A R^T$ cancel its inverse !

$$(AA^+)^2 = [AR^T(C^T A R^T)^{-1}C^T][AR^T(C^T A R^T)^{-1}C^T] = AR^T(C^T A R^T)^{-1}C^T = AA^+$$

8 $A^+ A = [R^T(C^T A R^T)^{-1}C^T]CR = R^T(C^T C\, R R^T)^{-1}C^T CR = R^T(R R^T)^{-1}R$ is symmetric.

9 Delete edges d and e. The graph becomes a triangle with 3 by 3 incidence matrix.

$$A = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix} = CR = \begin{bmatrix} -1 & 1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix}.$$

Then $\boldsymbol{C^T A R^T} = \begin{bmatrix} -1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}\begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{bmatrix} = \boldsymbol{3I}$

From formula (7): $\boldsymbol{A^+} = R^T\left(\dfrac{1}{3}\right)C^T = \dfrac{1}{3}A^T = \dfrac{1}{3}\begin{bmatrix} -1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix}.$

10 $A = \begin{bmatrix} 1 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ give $AB = \begin{bmatrix} 1 \end{bmatrix}$ and $BA = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$. Certainly

$(AB)^+ = \begin{bmatrix} 1 \end{bmatrix}$ is not $B^+ A^+ = \begin{bmatrix} \dfrac{1}{2} & \dfrac{1}{2} \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, so pseudoinverses *don't copy true inverses* (where $(AB)^{-1} = B^{-1}A^{-1}$). But they do copy inverses when ranks are right—as they are for

$$(BA)^+ = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}^+ = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}\begin{bmatrix} 1 & 1 \end{bmatrix} = A^+ B^+.$$

11 The four Penrose conditions for $(A^+)^+$ are all satisfied by A. (Also $(A^+)^+$ takes the column space of A^+ to its row space. This means $(A^+)^+$ takes the row space of A to its column space—just like A !)

Problem Set 5.1, page 203

1 $\det(2A) = 2^4 \det A = 8$; $\det(-A) = (-1)^4 \det A = \frac{1}{2}$; $\det(A^2) = \frac{1}{4}$; $\det(A^{-1}) = 2$.

2 $\det(\frac{1}{2}A) = (\frac{1}{2})^3 \det A = -\frac{1}{8}$ and $\det(-A) = (-1)^3 \det A = 1$; $\det(A^2) = 1$; $\det(A^{-1}) = -1$. If $\det A = 0$ then $\det A/2 = \det(-A) = \det A^2 = 0$; no A^{-1}.

3 (a) *False*: $\det(I + I)$ is not $1 + 1$ (except when $n = 1$)

   (b) *False*: $\det(4A)$ is $4^n \det A$

   (c) *False*: $A = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $AB - BA = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ is invertible.

4 Exchange rows 1 and 3 to show $\det J_3 = -1$. Exchange rows 1 and 4, then rows 2 and 3 to show $\det J_4 = 1$. Two exchanges = even permutation.

5 $|J_5| = 1$ by exchanging row 1 with 5 and row 2 with 4. $|J_6| = -1$, $|J_7| = -1$. Determinants $1, 1, -1, -1$ repeat in cycles of length 4 so the determinant of J_{101} is $+1$.

6 $\det A = 4, \det B = 0, \det C = 0$.

7 The 6 terms become $a(q + b)z - b(p + a)z + \cdots$ (4 more). The approach in the display (using linearity to split up row 2) is better. Result: $\det$ does not change if row 2 is added to row 1.

8 $\det A^{\mathrm{T}} = \begin{bmatrix} a & p & x \\ b & q & y \\ c & r & z \end{bmatrix} = \begin{array}{l} aqz + cpy + brx \\ - ary - bpz - cqx \end{array} = $ **same six terms as $\det A$**

   Key point: $\det P^{\mathrm{T}} = \det P$ for every permutation, because the number of row exchanges is the same (just done in reverse order). Then P is even when P^{T} is even.

9 $\det A = 1$ from two row exchanges. $\det B = 2$ (subtract rows 1 and 2 from row 3, then columns 1 and 2 from column 3). $\det C = 0$ and $\det D = 0$ (equal rows).

10 If the entries in every row add to zero, then $(1, 1, \ldots, 1)$ is in the nullspace: singular A has $\det = 0$. (The columns add to the zero column so they are linearly dependent.) If every row adds to one, then rows of $A - I$ add to zero (not necessarily $\det A = 1$).

11 If P_1 needs n exchanges to reach I and P_2 needs N exchanges then $P_1 P_2$ reaches I after those $n+N$ exchanges. So $\det(P_1 P_2) = (-1)^{n+N} = (-1)^n (-1)^N = (\det P_1)(\det P_2)$.

12 We can pair off even permutations with odd permutations: odd $=$ even followed by exchanging 1 and 2. Number of even permutations $=\frac{1}{2}n!=$number of odd permutations.

13 A singular rank one matrix has determinant $= 0$. The skew-symmetric A also has $\det A = 0$. A skew-symmetric matrix A of odd order 3: Changing every sign will multiply $\det A$ by $(-1)^3$ but also keep the same $\det A = \det A^{\mathrm{T}}$. So $\det A = 0$.

14 When the i, j entry is i times j, row $2 = 2$ times row 1 so $\det A = 0$.

When the ij entry is $i + j$, row $3 - $ row $2 = $ row2 $-$ row 1 so A is singular: $\det A = 0$.

15 Fill a row (or column) by 4 zeros to guarantee $\det = 0$. Leave only the main diagonal (12 zeros) to allow $\det A \neq 0$.

16 The cofactor formula $\det A = a_{11}C_{11} + \cdots + a_{1n}C_{1n}$ gives $\det = 0$ if all cofactors are zero. The 2 by 2 matrix of 1's has $\det = 0$ even though no cofactors are zero.

17 Two equal rows imply **det** $= \mathbf{0}$. Proof for 3×3 if row $1 = $ row 2. Then $a = p, b = q$, $c = r$. Then $aqz+brx+cpy-ary-bpz-cqx = abz+bcx+cay-acy-baz-cbx = 0$.

18 If A has two equal rows then A^{T} has two equal columns (say columns j and k). Then the columns are not independent. So $\det A^{\mathrm{T}} = 0$ and $\det A = 0$. Other proofs also reach this conclusion.

19 Start from $AC^{\mathrm{T}} = (\det A)I$. Take determinants of both sides:
$$(\det A)(\det C) = (\det A)^n \quad \text{and} \quad \det C = (\det A)^{n-1}$$
Note: If $\det A = 0$ (singular matrix) then A is the limit of invertible matrices $A_1, A_2, \ldots$ Apply $\det C_i = (\det A_i)^{n-1}$ and take the limit as $i \to \infty$. (How would you define C for a 1 by 1 matrix ??)

20 If you know C and if $\det A = 1$ then you know $A^{-1} = C^{\mathrm{T}}/1$. Then invert A^{-1} to find A.

Problem Set 5.2, page 209

1 If $\det A = 2$ then $\det A^{-1} = \frac{1}{2}$, $\det A^n = 2^n$, and $\det A^{\mathrm{T}} = 2$.

2 $\det A = -2$, independent columns; $\det B = 0$, dependent columns; $\det C = -1$, independent columns but $\det D = 0$ because its submatrix B has dependent rows (and dependent columns).

3 The problem suggests 3 ways to see that $\det A = 0$: All cofactors of row 1 are zero. A has rank ≤ 2. Each of the 6 terms in $\det A$ is zero. Notice also that column 2 has no pivot.

4 (a) $A = \begin{bmatrix} 0.9 & -0.9 \\ 0.9 & 0.9 \end{bmatrix}$ has $\det A = 1.62$ and $\det A^n = (1.62)^n \to \infty$.

(b) $A = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$ has $\det A = 0$ and $\det A^n = 0$ even if $A_{ij} = 2$.

5 (a) $|A| = \begin{vmatrix} 2 & 5 \\ 1 & 4 \end{vmatrix} = 3$, $|B_1| = \begin{vmatrix} 1 & 5 \\ 2 & 4 \end{vmatrix} = -6$, $|B_2| = \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = 3$ so $x_1 = -6/3 = -2$ and $x_2 = 3/3 = 1$ (b) $|A| = 4, |B_1| = 3, |B_2| = -2, |B_3| = 1$. Therefore $x_1 = 3/4$ and $x_2 = -1/2$ and $x_3 = 1/4$.

6 (a) $y = \begin{vmatrix} a & 1 \\ c & 0 \end{vmatrix} / \begin{vmatrix} a & b \\ c & d \end{vmatrix} = -c/(ad - bc)$ (b) $y = \det B_2 / \det A = (fg - id)/D$. That is because B_2 with $(1, 0, 0)$ in column 2 has $\det B_2 = fg - id$.

7 (a) $x_1 = 3/0$ and $x_2 = -2/0$: *no solution*   (b) $x_1 = x_2 = \mathbf{0}/\mathbf{0}$: *undetermined*.

8 The determinant is linear in its first column so $|x_1\,\boldsymbol{a}_1 + x_2\,\boldsymbol{a}_2 + x_3\,\boldsymbol{a}_3\,\boldsymbol{a}_2\,\boldsymbol{a}_3|$ splits into $x_1|\boldsymbol{a}_1\,\boldsymbol{a}_2\,\boldsymbol{a}_3| + x_2|\boldsymbol{a}_2\,\boldsymbol{a}_2\,\boldsymbol{a}_3| + x_3|\boldsymbol{a}_3\,\boldsymbol{a}_2\,\boldsymbol{a}_3|$. The last two determinants are zero because of repeated columns, leaving $x_1|\boldsymbol{a}_1\,\boldsymbol{a}_2\,\boldsymbol{a}_3|$ which is $x_1 \det A$.

9 If the first column in A is also the right side b then $\det A = \det B_1$. Both B_2 and B_3 are singular since a column is repeated. Therefore $x_1 = |B_1|/|A| = 1$ and $x_2 = x_3 = 0$.

10 The pattern $\det = 1, 0, -1, -1, 0, 1$ repeats as in $E_{n+6} = E_n$. So $E_{100} = E_4$ after 16 repeats of length 6. And $E_4 = \mathbf{-1}$.

11 If the entries are 1 to 9, the maximum determinant **may be** $412 = \begin{vmatrix} 9 & 3 & 5 \\ 4 & 8 & 1 \\ 2 & 6 & 7 \end{vmatrix}$.

12 **True**. We know that $\det AB = (\det A)(\det B)$. And we know $\det(ABC) = (\det AB)(\det C)$.

Put these together to prove $\det(ABC) = (\det A)(\det B)(\det C)$.

13 The combinations of $i = (1, 0, 0)$ and $i + j = (1, 1, 0)$ fill the xy **plane** in xyz space.

14 (a) Sum = zero vector. (b) Sum = -2:00 vector = 8:00 vector.

(c) 2:00 is $30°$ from horizontal = $(\cos \frac{\pi}{6}, \sin \frac{\pi}{6}) = (\sqrt{3}/2, 1/2)$.

15 Moving the origin to 6:00 adds $j = (0, 1)$ to every vector. So the sum of twelve vectors changes from $\mathbf{0}$ to $12j = (0, 12)$.

16 (a) $\det P = 1$ because columns $2, 1, 4, 3$ have two exchanges from $1, 2, 3, 4$.

(b) $\det P = -1$ because columns $3, 2, 1, 4$ have only one exchange (of 3 and 1).

(c) $\det P = -1$ because columns $1, 2, 4, 3$ have one exchange (3 and 4).

17 The sum is $(v - u) + (w - v) + (u - w) = $ **zero vector**. Those three sides of a triangle are in the same plane !

18 *All* vectors in 3D are combinations of u, v, w as drawn (not in the same plane). Start by seeing that $cu + dv$ fills a plane, then adding all the vectors ew fills all of $\mathbf{R}^3$. Different answer when u, v, w are in the same plane.

19 The only 4×4 column orders that start with $3, 2$ are $3, 2, 1, 4$ and $3, 2, 4, 1$ (so 2 terms in $\det A$).

Problem Set 5.3, page 214

1 Edge matrix $E = $ identity matrix I for a unit cube. With edges e_1, e_2, e_3 out from $(0, 0, 0)$, the other 7 corners are $e_1, e_2, e_3, e_1 + e_2, e_1 + e_3, e_2 + e_3, e_1 + e_2 + e_3$.

2 Perpendicular e's give an ordinary box with side lengths $||e_1||, ||e_2||, ||e_3||$.

3 The largest box with edge lengths $1, 2, 3$ is the ordinary 90-degree box with volume 6. This is Hadamard's inequality (see Problem 4).

4 $|\det E| \leq (||e_1||)(||e_2||) \dots (||e_n||)$ is "Hadamard's inequality". One proof starts with $E = \begin{bmatrix} e_1 & \dots & e_n \end{bmatrix} = QR$ from Section 4.4 (Gram-Schmidt producing orthogonal q's from independent e's). Each e_j is a combination of orthogonal q_i with $||q_i|| = 1$:

$$a_j = \sum r_{ij} q_i \qquad ||a_j||^2 = \sum r_{ij}^2 ||q_i||^2 \geq r_{jj}^2$$

$$|\det E| = |\det Q| \, |\det R| = (r_{11}) \dots (r_{nn}) \leq ||e_1|| \dots ||e_n||.$$

Wikipedia proves Hadamard's inequality from "geometric mean" $\leq$ "arithmetic mean".

5 (a) The parallelogram area with edges $(3, 2)$ and $(1, 4)$ is the determinant of $\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix} =$

$12 - 2 = 10$.   (b) Triangle area $= \frac{1}{2}$(parallelogram area) $= \dfrac{1}{2} \begin{vmatrix} 3 & 2 \\ 4 & 6 \end{vmatrix} = 5$.

(c) The triangle area is $\frac{1}{2}$(parallelogram area) $= \dfrac{1}{2} \begin{vmatrix} 3 & 2 \\ 1 & 4 \end{vmatrix} = 5$.

(b)

(c)
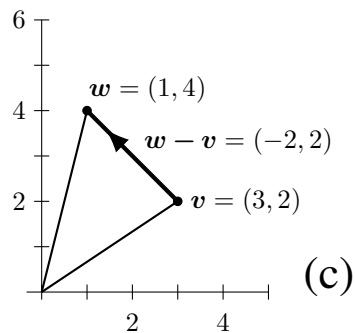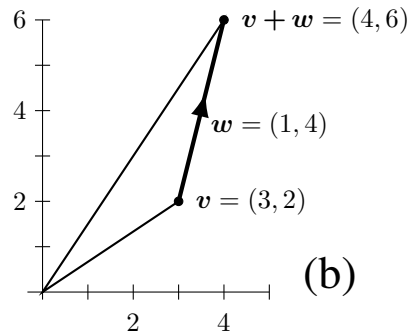
6 (a) Two sides are $(3,4) - (2,1) = (1,3)$ and $(0,5) - (2,1) = (-2,4)$. The triangle

area is half the parallelogram area $= \dfrac{1}{2} \begin{vmatrix} 1 & 3 \\ -2 & 4 \end{vmatrix} = 5$.
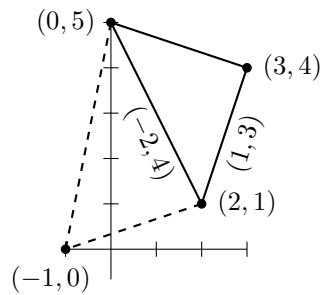
(b) has an additional triangle (dashed lines) of area $\dfrac{1}{2} \begin{vmatrix} 3 & 1 \\ 1 & 5 \end{vmatrix} = 7$.

7 The hypercube in 4-dimensions has perpendicular sides of length $\sqrt{1+1+1+1} = 2$.
So its volume is $2^4 = 16$. This must be $|\det H|$.

8 An n-dimensional cube has 2^n corners. This is the case $m = 0$ in Wikipedia's formula:

An n-dimensional cube has $2^{n-m} \dfrac{n!}{m!(n-m)!}$ m-dimensional sides, edges, corners,...

For edges $(m = 1)$ this rule gives 2^{n-1} times n. 12 edges for a 3D cube $(n = 3)$.

For faces of dimension $m = n - 1$ this rule gives $2n$. Six faces for a 3D cube.

The cube in $\mathbf{R}^n$ whose edges come from $2I$ has volume $\det(2I) = 2^n$.

9 The 3-dimensional "unit pyramid" in $\mathbf{R}^3$ has volume $\dfrac{1}{3!} = \dfrac{1}{6}$. I believe the 4-dimensional
"unit pyramid" has volume $\dfrac{1}{4!} = \dfrac{1}{24}$.

Problem Set 6.1, page 226

1 The eigenvalues of A are $\lambda = 1$ and 0.5 $\left(\text{or } \frac{1}{2}\right)$.

The eigenvalues of A^n are $\lambda = 1$ and $\left(\frac{1}{2}\right)^n$.

The eigenvalues of A^∞ are $\lambda = 1$ and 0.

(a) A row exchange leaves this A with $\lambda = 1$ and -0.5 $\left(\text{or } -\frac{1}{2}\right)$.

(b) Every A has $n - r$ zero eigenvalues ($r = $ rank) : not changed by elimination.

2 A has $\lambda_1 = -1$ and $\lambda_2 = 5$ with eigenvectors $x_1 = (-2, 1)$ and $x_2 = (1, 1)$. The matrix $A + I$ has the same eigenvectors, with eigenvalues increased by 1 to **0** and **6**. That zero eigenvalue correctly indicates that $A + I$ is singular.

3 A has $\lambda_1 = 2$ and $\lambda_2 = -1$ (check trace and determinant) with $x_1 = (1, 1)$ and $x_2 = (2, -1)$. A^{-1} has the same eigenvectors, with eigenvalues $1/\lambda = \frac{1}{2}$ and -1.

4 $\det(A - \lambda I) = \lambda^2 + \lambda - 6 = (\lambda + 3)(\lambda - 2)$. Then A has $\lambda_1 = -3$ and $\lambda_2 = 2$ (check trace $= -1$ and determinant $= -6$) with $x_1 = (3, -2)$ and $x_2 = (1, 1)$. A^2 has the *same eigenvectors* as A, with eigenvalues $\lambda_1^2 = 9$ and $\lambda_2^2 = 4$.

5 A and B have eigenvalues 1 and 3 (their diagonal entries : triangular matrices). $A + B$ has $\lambda^2 + 8\lambda + 15 = 0$ and $\lambda_1 = 3$, $\lambda_2 = 5$. Eigenvalues of $A + B$ *are not equal* to eigenvalues of A plus eigenvalues of B.

6 A and B have $\lambda_1 = 1$ and $\lambda_2 = 1$. AB and BA have $\lambda^2 - 4\lambda + 1 = 0$ and the quadratic formula gives $\lambda = 2 \pm \sqrt{3}$. Eigenvalues of AB *are not equal* to eigenvalues of A times eigenvalues of B. ***Eigenvalues of AB and BA are equal*** (this is proved at the end of Section 6.2).

7 The eigenvalues of U (on its diagonal) are the *pivots* of A. The eigenvalues of L (on its diagonal) are all 1's. The eigenvalues of A *are not* the same as the pivots.

8 (a) Multiply Ax to see λx which reveals λ (b) Solve $(A - \lambda I)x = 0$ to find x.

9 (a) Multiply $A\boldsymbol{x} = \lambda\boldsymbol{x}$ by A: $A(A\boldsymbol{x}) = A(\lambda\boldsymbol{x}) = \lambda A\boldsymbol{x}$ gives $A^2\boldsymbol{x} = \boldsymbol{\lambda}^2\boldsymbol{x}$

  (b) Multiply by A^{-1}: $\boldsymbol{x} = A^{-1}A\boldsymbol{x} = A^{-1}\lambda\boldsymbol{x} = \lambda A^{-1}\boldsymbol{x}$ gives $A^{-1}\boldsymbol{x} = \frac{1}{\boldsymbol{\lambda}}\boldsymbol{x}$

  (c) Add $I\boldsymbol{x} = \boldsymbol{x}$: $(A + I)\boldsymbol{x} = (\boldsymbol{\lambda} + 1)\boldsymbol{x}$.

10 $\det(A - \lambda I) = \lambda^2 - 1.4\lambda + 0.4$ so A has $\lambda_1 = 1$ and $\lambda_2 = 0.4$ with $\boldsymbol{x}_1 = (1, 2)$ and $\boldsymbol{x}_2 = (1, -1)$. A^∞ has $\lambda_1 = 1$ and $\lambda_2 = 0$ (same eigenvectors as A). A^{100} has $\lambda_1 = 1$ and $\lambda_2 = (0.4)^{100}$ which is near zero. So A^{100} is very near A^∞: same eigenvectors and close eigenvalues.

11 Proof 1. $A - \lambda_1 I$ is singular so its two columns are in the same direction. Also $(A - \lambda_1 I)\boldsymbol{x}_2 = (\lambda_2 - \lambda_1)\boldsymbol{x}_2$. So $\boldsymbol{x}_2$ is in the column space and both columns must be multiples of $\boldsymbol{x}_2$. Here is also a **second proof**: Columns of $A - \lambda_1 I$ are in the nullspace of $A - \lambda_2 I$ because $M = (A - \lambda_2 I)(A - \lambda_1 I)$ is the zero matrix [this is the *Cayley-Hamilton Theorem* in Problem 6.2.30]. Notice that M has *zero eigenvalues* $(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_1) = 0$ and $(\lambda_2 - \lambda_2)(\lambda_2 - \lambda_1) = 0$. So those columns solve $(A - \lambda_2 I)\boldsymbol{x} = \boldsymbol{0}$, they are eigenvectors.

12 The projection matrix P has $\lambda = 1, 0, 1$ with eigenvectors $(1, 2, 0)$, $(2, -1, 0)$, $(0, 0, 1)$. Add the first and last vectors: $(1, 2, 1)$ also has $\lambda = 1$. The whole column space of P contains eigenvectors with $\lambda = 1$! *Note* $P^2 = P$ leads to $\lambda^2 = \lambda$ so $\lambda = 0$ or 1.

13 (a) $P\boldsymbol{u} = (\boldsymbol{u}\boldsymbol{u}^\mathrm{T})\boldsymbol{u} = \boldsymbol{u}$ times $\boldsymbol{u}^\mathrm{T}\boldsymbol{u} = \boldsymbol{u}$ times 1. So $\lambda = 1$.
  (b) $P\boldsymbol{v} = (\boldsymbol{u}\boldsymbol{u}^\mathrm{T})\boldsymbol{v} = \boldsymbol{u}(\boldsymbol{u}^\mathrm{T}\boldsymbol{v}) = \boldsymbol{0}$.
  (c) $\boldsymbol{x}_1 = (-1, 1, 0, 0)$, $\boldsymbol{x}_2 = (-3, 0, 1, 0)$, $\boldsymbol{x}_3 = (-5, 0, 0, 1)$ all have $P\boldsymbol{x} = 0\boldsymbol{x} = \boldsymbol{0}$.

14 $\det(Q - \lambda I) = \lambda^2 - 2\lambda\cos\theta + 1 = 0$ when $\lambda = \cos\theta \pm i\sin\theta = e^{i\theta}$ and $e^{-i\theta}$. Check $\lambda_1\lambda_2 = \cos^2\theta + \sin^2\theta = 1$ and $\lambda_1 + \lambda_2 = 2\cos\theta$. Two eigenvectors of this rotation matrix are $\boldsymbol{x}_1 = (1, i)$ and $\boldsymbol{x}_2 = (1, -i)$ (or $c\boldsymbol{x}_1$ and $d\boldsymbol{x}_2$ with $cd \neq 0$).

15 The other two eigenvalues are $\lambda = \frac{1}{2}(-1 \pm i\sqrt{3})$. Those three eigenvalues add to $0 = $ trace of P. The three eigenvalues of the second P are $1, 1, -1$.

16 Set $\lambda = 0$ in $\det(A - \lambda I) = (\lambda_1 - \lambda)\ldots(\lambda_n - \lambda)$ to find $\det A = (\lambda_1)(\lambda_2)\cdots(\lambda_n)$.

17 Comparing $\lambda^2 - (a+d)\lambda + (ad-bc)$ with $(\lambda - \lambda_1)(\lambda - \lambda_2) = \lambda^2 - (\lambda_1 + \lambda_2)\lambda + \lambda_1\lambda_2$

shows :
$$a + d = \lambda_1 + \lambda_2 = \text{ trace} \qquad ad - bc = \lambda_1\lambda_2 = \text{ determinant}$$

If $\lambda_1 = 3$ and $\lambda_2 = 4$ then $\det(A - \lambda I) = \lambda^2 - 7\lambda + 12$.

18 Trace $= 9$. Three possibilities are $A = \begin{bmatrix} 4 & 0 \\ 0 & 5 \end{bmatrix}$, $\begin{bmatrix} 10 & -1 \\ 30 & -1 \end{bmatrix}$, $\begin{bmatrix} 4 & 6 \\ 0 & 5 \end{bmatrix}$.

19 (a) rank $= 2$ (b) $\det(B^{\mathrm{T}}B) = 0$ (d) eigenvalues of $(B^2 + I)^{-1}$ are $1, \frac{1}{2}, \frac{1}{5}$.

20 $A = \begin{bmatrix} 0 & 1 \\ -\mathbf{28} & \mathbf{11} \end{bmatrix}$ has trace 11 and determinant 28, so $\lambda = 4$ and 7. Moving to a 3 by

3 companion matrix, for eigenvalues $1, 2, 3$ we want $\det(C - \lambda I) = (1 - \lambda)(2 - \lambda)$

$(3 - \lambda)$. Multiply out to get $-\lambda^3 + 6\lambda^2 - 11\lambda + 6$. To get those numbers $6, -11, 6$

from a companion matrix you just put them into the last row :

$C = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \mathbf{6} & -\mathbf{11} & \mathbf{6} \end{bmatrix}$ Notice the trace $6 = 1 + 2 + 3$ and determinant $6 = (1)(2)(3)$.

21 $(A - \lambda I)$ has the same determinant as $(A - \lambda I)^{\mathrm{T}}$ because every square matrix has

$\det M = \det M^{\mathrm{T}}$. Pick $M = A - \lambda I$.

$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \text{ have } \textit{different eigenvectors } \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

22 We can choose $M = \begin{bmatrix} .1 & 0 & 0 \\ .2 & .4 & 0 \\ .7 & .6 & 1 \end{bmatrix}$. Its eigenvalues $\lambda = .1, \; .4, \; 1.0$ are on the

diagonal. Clearly M^{T} has rows adding to 1 so M^{T} times the column $v = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^{\mathrm{T}}$

equals v. Challenge : A 3 by 3 singular Markov matrix with trace $\frac{1}{2}$ has $\boldsymbol{\lambda = 0, \; 1, \; -\frac{1}{2}}$.

23 $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $\begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}$.   Always A^2 is the zero matrix if $\lambda = 0$ and 0, by the Cayley-Hamilton Theorem in Problem 6.2.30.

24 $\lambda = \mathbf{0, 0, 6}$ (*notice rank* 1 *and trace* 6). Two eigenvectors of uv^{T} are perpendicular to

v and the third eigenvector is u : $x_1 = (0, -2, 1)$, $x_2 = (1, -2, 0)$, $x_3 = (1, 2, 1)$.

25 When A and B have the same n λ's and $\boldsymbol{x}$'s, look at any combination $\boldsymbol{v} = c_1\boldsymbol{x}_1 + \cdots + c_n\boldsymbol{x}_n$. Multiply by A and B: $A\boldsymbol{v} = c_1\lambda_1\boldsymbol{x}_1 + \cdots + c_n\lambda_n\boldsymbol{x}_n$ **equals** $B\boldsymbol{v} = c_1\lambda_1\boldsymbol{x}_1 + \cdots + c_n\lambda_n\boldsymbol{x}_n$ **for all vectors** $\boldsymbol{v}$. So $A = B$.

26 A has eigenvalues 1 and 2 from block B (with eigenvectors ending in $0, 0$). A also has eigenvalues 5 and 7 from block D because A^{T} has eigenvalues $5, 7$ from block D^{T} (and transposing doesn't change eigenvalues).

27 A has rank 1 with eigenvalues $0, 0, 0, 4$ (the 4 comes from the trace of A). C has rank 2 (ensuring two zero eigenvalues) and $(1, 1, 1, 1)$ is an eigenvector with $\lambda = 2$. With trace 4, the other eigenvalue is also $\lambda = 2$, and its eigenvector is $(1, -1, 1, -1)$.

28 The 4 by 4 matrix A of 1's has $\lambda = 0, 0, 0, 4$. Then $B = A - I$ has $\lambda = -1, -1, -1, 3$. And $C = I - A$ has $\lambda = 1, 1, 1, -3$.

29 A is triangular: $\lambda(A) = 1, 4, 6$; $\lambda(B) = 2, \sqrt{3}, -\sqrt{3}$; C has rank one: $\lambda(C) = 0, 0, 6$.

30 $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = (a + b) \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ when $a + b = c + d$. Thus $\lambda_1 = a + b$.

Then $\lambda_2 = \text{trace} - \lambda_1 = (a + d) - (a + b) = d - b$.

31 If PA exchanges rows 1 and 2 of A, then AP^{T} exchanges columns 1 and 2. In fact

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = P^{\mathrm{T}} = P^{-1} \quad \text{and} \quad B = PAP^{\mathrm{T}} = PAP^{-1}.$$

Then B is **similar** to A and they have the same eigenvalues. In this rank 1 and trace 11 example, the eigenvalues of A and B are $0, 0, 11$. From $A - 11I = \begin{bmatrix} -10 & 2 & 1 \\ 3 & -5 & 3 \\ 4 & 8 & -7 \end{bmatrix}$

the eigenvector for $\lambda = 11$ is $\begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}$.

32 (a) u is a basis for the nullspace (we know $Au = 0u$); v and w give a basis for the column space (we know Av and Aw are in the column space).

(b) $A(v/3 + w/5) = 3v/3 + 5w/5 = v + w$. So $x = v/3 + w/5$ is a particular solution to $Ax = v + w$. Add any cu from the nullspace to find all solutions.

(c) If $Ax = u$ had a solution, u would be in the column space : wrong dimension 3.

33 Always $(uv^{\mathrm{T}})\,u = u(v^{\mathrm{T}}u)$ so u is an eigenvector of uv^{T} with $\lambda = v^{\mathrm{T}}u$. (Watch numbers $v^{\mathrm{T}}u$, vectors u, matrices uv^{T} !!) If $v^{\mathrm{T}}u = 0$ then $A^2 = u(v^{\mathrm{T}}u)v^{\mathrm{T}}$ is the zero matrix and $\lambda^2 = 0, 0$ and $\lambda = 0, 0$ and trace $(A) = 0$. This zero trace also comes from adding the diagonal entries of $A = uv^{\mathrm{T}}$:

$$A = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{bmatrix} u_1v_1 & u_1v_2 \\ u_2v_1 & u_2v_2 \end{bmatrix} \quad \text{has trace } u_1v_1 + u_2v_2 = v^{\mathrm{T}}u = 0$$

34 The vector $(1, 1, 1, 1)$ is not changed by P. It is the eigenvector for $\lambda = 1$. The other 3 eigenvectors (discussed in detail in Section 6.4) are

$$x_2, x_3, x_4 = \begin{bmatrix} 1 \\ i \\ i^2 \\ i^3 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} \begin{bmatrix} 1 \\ -i \\ (-i)^2 \\ (-i)^3 \end{bmatrix}.$$

35 The six 3 by 3 permutation matrices include $P = I$ and three single row exchange matrices P_{12}, P_{13}, P_{23} and two double exchange matrices like $P_{12}P_{13}$. Since $P^{\mathrm{T}}P = I$ gives $(\det P)^2 = 1$, the determinant of P is 1 or -1. The pivots are always 1 (but there may be row exchanges). The trace of P can be 3 (for $P = I$) or 1 (for row exchange) or 0 (for double exchange). The possible eigenvalues are 1 and -1 and $e^{2\pi i/3}$ and $e^{-2\pi i/3}$.

36 $AB - BA = I$ can happen only for infinite matrices. If $A^{\mathrm{T}} = A$ and $B^{\mathrm{T}} = -B$ then

$$x^{\mathrm{T}}x = x^{\mathrm{T}}\,(AB - BA)\,x = x^{\mathrm{T}}\,(A^{\mathrm{T}}B + B^{\mathrm{T}}A)\,x \leq ||Ax||\,||Bx|| + ||Bx||\,||Ax||.$$

Therefore $||Ax||\,||Bx|| \geq \frac{1}{2}||x||^2$ and $(||Ax||/||x||)\,(||Bx||/||x||) \geq \frac{1}{2}$.

37 $\lambda_1 = e^{2\pi i/3}$ and $\lambda_2 = e^{-2\pi i/3}$ give $\det \lambda_1\lambda_2 = 1$ and trace $\lambda_1 + \lambda_2 = -1$.

$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ with $\theta = \dfrac{2\pi}{3}$ has this trace and det. So does every $M^{-1}AM$!

38 (a) Since the columns of A add to 1, one eigenvalue is $\lambda = 1$ and the other is $c - 0.6$

(to give the correct trace $c + 0.4$).

(b) If $c = 1.6$ then both eigenvalues are 1, and all solutions to $(A - I)\,x = 0$ are

multiples of $x = (1, -1)$. In this case A has rank 1.

(c) If $c = 0.8$, the eigenvectors for $\lambda = 1$ are multiples of $(1, 3)$. Since all powers A^n

also have column sums $= 1$, A^n will approach $\dfrac{1}{4}\begin{bmatrix} 1 & 1 \\ 3 & 3 \end{bmatrix} = $ rank-1 matrix A^∞ with

eigenvalues $1, 0$ and correct eigenvectors. $(1, 3)$ and $(1, -1)$.

Problem Set 6.2, page 242

1 Eigenvectors in X and eigenvalues 1 and 3 in Λ. Then $A = X\Lambda X^{-1}$ is

$$\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}.$$ The second matrix has $\lambda = 0$ (rank 1) and

$\lambda = 4$ (trace $= 4$). Then $A = X\Lambda X^{-1}$ is $\begin{bmatrix} 1 & 1 \\ 3 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 3 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} \frac{3}{4} & -\frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} \end{bmatrix}.$

$A^3 = X\Lambda^3 X^{-1}$ and $A^{-1} = X\Lambda^{-1}X^{-1}.$

2   Put the eigenvectors in X $A = X\Lambda X^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 0 & 5 \end{bmatrix}.$

and eigenvalues $2, 5$ in Λ.

3 If $A = X\Lambda X^{-1}$ then the eigenvalue matrix for $A + 2I$ is $\Lambda + 2I$ and the eigenvector

matrix is still X. So $A + 2I = X(\Lambda + 2I)X^{-1} = X\Lambda X^{-1} + X(2I)X^{-1} = A + 2I.$

4 (a) False: We are not given the λ's (b) True (c) True since X has independent columns.

(d) False: For this we would need the eigenvectors of X.

5 With $X = I$, $A = X\Lambda X^{-1} = \Lambda$ is a diagonal matrix. If X is triangular, then X^{-1} is

triangular, so $X\Lambda X^{-1}$ is also triangular.

6 The columns of X are nonzero multiples of $(2,1)$ and $(0,1)$: either order. The same

eigenvector matrices diagonalize A and A^{-1}.

7 Every matrix that has eigenvectors $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ has the form

$$A = X\Lambda X^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix}/2 = \frac{1}{2}\begin{bmatrix} \lambda_1 + \lambda_2 & \lambda_1 - \lambda_2 \\ \lambda_1 - \lambda_2 & \lambda_1 + \lambda_2 \end{bmatrix}.$$

You could check trace $= \lambda_1 + \lambda_2$ and $\det = \frac{1}{4}\, 4\lambda_1\lambda_2 = \lambda_1\lambda_2.$

8 $A = X\Lambda X^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \dfrac{1}{\lambda_1 - \lambda_2}\begin{bmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}\begin{bmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{bmatrix}.$

$X\Lambda^k X^{-1} = \dfrac{1}{\lambda_1 - \lambda_2}\begin{bmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} \lambda_1^k & 0 \\ 0 & \lambda_2^k \end{bmatrix}\begin{bmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix}.$

The second component is $F_k = (\lambda_1^k - \lambda_2^k)/(\lambda_1 - \lambda_2)$.

9 (a) The equations are $\begin{bmatrix} G_{k+2} \\ G_{k+1} \end{bmatrix} = A \begin{bmatrix} G_{k+1} \\ G_k \end{bmatrix}$ with $A = \begin{bmatrix} .5 & .5 \\ 1 & 0 \end{bmatrix}$. This matrix

has $\lambda_1 = 1$, $\lambda_2 = -\frac{1}{2}$ with $x_1 = (1, 1)$, $x_2 = (1, -2)$

(b) $A^n = X\Lambda^n X^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 1^n & 0 \\ 0 & (-.5)^n \end{bmatrix} \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{bmatrix} \to A^\infty = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}$

10 The rule $F_{k+2} = F_{k+1} + F_k$ produces the pattern: even, odd, odd, even, odd, odd, ...

11 (a) *True* (no zero eigenvalues)   (b) *False* (repeated $\lambda = 2$ may have only one line of

eigenvectors)   (c) *False* (repeated λ may have a full set of eigenvectors)

12 (a) False: don't know if $\lambda = 0$ or not.

(b) True: an eigenvector is missing, which can only happen for a repeated eigenvalue.

(c) True: We know there is only one line of eigenvectors.

13 $A = \begin{bmatrix} 8 & 3 \\ -3 & 2 \end{bmatrix}$ (or other), $A = \begin{bmatrix} 9 & 4 \\ -4 & 1 \end{bmatrix}$, $A = \begin{bmatrix} 10 & 5 \\ -5 & 0 \end{bmatrix}$; only eigenvectors are $x = (c, -c)$.

14 The rank of $A - 3I$ is $r = 1$. Changing any entry except $a_{12} = 1$ makes A

diagonalizable (the new A will have two different eigenvalues)

15 $A^k = X\Lambda^k X^{-1}$ approaches zero **if and only if every** $|\lambda| < 1$; A_1 is a Markov matrix

so $\lambda_{\max} = 1$ and $A_1^k \to A_1^\infty$, A_2 has $\lambda = .6 \pm .3$ so $A_2^k \to 0$.

16 $\begin{bmatrix} .6 & .9 \\ .4 & .1 \end{bmatrix} = X\Lambda X^{-1}$ with $\Lambda = \begin{bmatrix} 1 & 0 \\ 0 & .2 \end{bmatrix}$ and $X = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$; $\Lambda^k \to \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

Then $A_1^k = X\Lambda^k X^{-1} \to \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$: *steady state*.

17 A_2 is $X\Lambda X^{-1}$ with $\Lambda = \begin{bmatrix} .9 & 0 \\ 0 & .3 \end{bmatrix}$ and $X = \begin{bmatrix} 3 & -3 \\ 1 & 1 \end{bmatrix}$; $A_2^{10} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = (.9)^{10} \begin{bmatrix} 3 \\ 1 \end{bmatrix}$.

$A_2^{10} \begin{bmatrix} 3 \\ -1 \end{bmatrix} = (.3)^{10} \begin{bmatrix} 3 \\ -1 \end{bmatrix}$. Then $A_2^{10} \begin{bmatrix} 6 \\ 0 \end{bmatrix} = (.9)^{10} \begin{bmatrix} 3 \\ 1 \end{bmatrix} + (.3)^{10} \begin{bmatrix} 3 \\ -1 \end{bmatrix}$ because

$u_0 = \begin{bmatrix} 6 \\ 0 \end{bmatrix}$ is the sum of $\begin{bmatrix} 3 \\ 1 \end{bmatrix} + \begin{bmatrix} 3 \\ -1 \end{bmatrix}$.

18 $\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = X\Lambda X^{-1} = \dfrac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ and

$A^k = X\Lambda^k X^{-1} = \dfrac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3^k \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$

Multiply those last three matrices to get $A^k = \dfrac{1}{2} \begin{bmatrix} 1 + 3^k & 1 - 3^k \\ 1 - 3^k & 1 + 3^k \end{bmatrix}.$

19 $B^k = X\Lambda^k X^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix}^k \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 5^k & 5^k - 4^k \\ 0 & 4^k \end{bmatrix}.$

20 $\det A = (\det X)(\det \Lambda)(\det X^{-1}) = \det \Lambda = \lambda_1 \cdots \lambda_n$. This proof ($\det$ = product of λ's) works when A is *diagonalizable*. The formula is always true.

21 trace $XY = (aq + bs) + (cr + dt)$ is equal to $(qa + rc) + (sb + td) = $ trace YX. Diagonalizable case: the trace of $X\Lambda X^{-1} = $ trace of $(\Lambda X^{-1})X = $ trace of $\Lambda = \Sigma \lambda_i$. $AB - BA = I$ is impossible since the left side has **trace $= 0$**.

22 If $A = X\Lambda X^{-1}$ then $B = \begin{bmatrix} A & 0 \\ 0 & 2A \end{bmatrix} = \begin{bmatrix} X & 0 \\ 0 & X \end{bmatrix} \begin{bmatrix} \Lambda & 0 \\ 0 & 2\Lambda \end{bmatrix} \begin{bmatrix} X^{-1} & 0 \\ 0 & X^{-1} \end{bmatrix}$. So B has the original λ's from A and the additional eigenvalues $2\lambda_1, \ldots, 2\lambda_n$ from $2A$.

23 The A's form a subspace since cA and $A_1 + A_2$ all have the same X. When $X = I$ the A's with those eigenvectors give the subspace of **diagonal matrices**. The dimension of that matrix space is 4 since the matrices are 4 by 4.

24 If A has columns $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ then column by column, $A^2 = A$ means every $A\boldsymbol{x}_i = \boldsymbol{x}_i$. All vectors in the column space (combinations of those columns $\boldsymbol{x}_i$) are eigenvectors with $\lambda = 1$. Always the nullspace has $\lambda = 0$ (A might have dependent columns, so there could be less than n eigenvectors with $\lambda = 1$). Dimensions of those spaces $\mathbf{C}(A)$ and $\mathbf{N}(A)$ add to n by the Fundamental Theorem, so A *is diagonalizable* (n independent eigenvectors altogether).

25 Two problems: The nullspace and column space can overlap, so $\boldsymbol{x}$ could be in both. There may not be r independent eigenvectors in the column space.

26 $R = X\sqrt{\Lambda}X^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 3 & \\ & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} /2 = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ has $\boldsymbol{R^2 = A}$.

$\sqrt{B}$ needs $\lambda = \sqrt{9}$ and $\sqrt{-1}$, the trace (their sum) is not real so $\sqrt{B}$ cannot be real.

Note that the square root of $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$ has *two* imaginary eigenvalues $\sqrt{-1} = i$ and

$-i$, real trace 0, real square root $R = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$.

27 The factorizations of A and B into $X\Lambda X^{-1}$ are the same. So $A = B$.

28 $A = X\Lambda_1 X^{-1}$ and $B = X\Lambda_2 X^{-1}$. Diagonal matrices always give $\Lambda_1\Lambda_2 = \Lambda_2\Lambda_1$.

Then $AB = BA$ from

$$X\Lambda_1 X^{-1} X\Lambda_2 X^{-1} = X\boldsymbol{\Lambda_1\Lambda_2}X^{-1} = X\boldsymbol{\Lambda_2\Lambda_1}X^{-1} = X\Lambda_2 X^{-1} X\Lambda_1 X^{-1} = BA.$$

29 (a) $A = \begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$ has $\lambda = a$ and $\lambda = d$: $(A - aI)(A - dI) = \begin{bmatrix} 0 & b \\ 0 & d - a \end{bmatrix} \begin{bmatrix} a - d & b \\ 0 & 0 \end{bmatrix}$

$= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.   (b) $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ has $A^2 = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ and $A^2 - A - I = 0$ is true,

matching $\det(A - \lambda I) = \lambda^2 - \lambda - 1 = 0$ as the Cayley-Hamilton Theorem predicts.

30 When $A = X\Lambda X^{-1}$ is diagonalizable, the matrix $A - \lambda_j I = X(\Lambda - \lambda_j I)X^{-1}$ will

have 0 in the j, j diagonal entry of $\Lambda - \lambda_j I$. The product $p(A)$ becomes

$$p(A) = (A - \lambda_1 I) \cdots (A - \lambda_n I) = X(\Lambda - \lambda_1 I) \cdots (\Lambda - \lambda_n I)X^{-1}.$$

That product is the zero matrix because the factors produce a zero in each

diagonal position. Then $p(A) = $ zero matrix, which is the Cayley-Hamilton Theorem.

(If A is not diagonalizable, one proof is to take a sequence of diagonalizable matrices

approaching A.)

Comment   I have also seen the following Cayley-Hamilton proof but I am not con-

vinced:

Apply the formula $AC^{\mathrm{T}} = (\det A)I$ from Section 5.1 to $A - \lambda I$ with variable λ. Its

cofactor matrix C will be a polynomial in λ, since cofactors are determinants:

$$(A - \lambda I)C^{\mathrm{T}}(\lambda) = \det(A - \lambda I)I = p(\lambda)I.$$

"For fixed A, this is an identity between two matrix polynomials." Set $\lambda = A$ to find the zero matrix on the left, so $p(A) =$ zero matrix on the right—which is the Cayley-Hamilton Theorem.

I am not certain about the key step of substituting a matrix A for λ. If other matrices B are substituted for λ, does the identity remain true? If $AB \neq BA$, even the order of multiplication seems unclear . . .

31 If $AB = BA$, then B has the same eigenvectors $(1,0)$ and $(0,1)$ as A. So B is also diagonal $b = c = 0$. The nullspace for the following equation is 2-dimensional:

$$AB - BA = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} - \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} \mathbf{0} & -\boldsymbol{b} \\ \boldsymbol{c} & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

Those 4 equations $0 = 0, -b = 0, c = 0, 0 = 0$ have a 4 by 4 coefficient matrix with rank $= 4 - 2 = 2$.

32 B has $\lambda = i$ and $-i$, so B^4 has $\lambda^4 = 1$ and 1. Then $B^4 = I$ and $B^{1024} = I$.

C has $\lambda = (1 \pm \sqrt{3}i)/2$. This λ is $\exp(\pm \pi i/3)$ so $\lambda^3 = -1$ and -1. Then $C^3 = -I$ which leads to $C^{1024} = (-I)^{341}C = -C$.

33 The eigenvalues of $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ are $\lambda = e^{i\theta}$ and $e^{-i\theta}$ (trace $2\cos\theta$ and determinant $\lambda_1 \lambda_2 = 1$). Their eigenvectors are $(1, -i)$ and $(1, i)$:

$$A^n = X\Lambda^n X^{-1} = \begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix} \begin{bmatrix} e^{in\theta} & \\ & e^{-in\theta} \end{bmatrix} \begin{bmatrix} i & -1 \\ i & 1 \end{bmatrix} /2i$$

$$= \begin{bmatrix} (e^{in\theta} + e^{-in\theta})/2 & \cdots \\ (e^{in\theta} - e^{-in\theta})/2i & \cdots \end{bmatrix} = \begin{bmatrix} \cos n\theta & -\sin n\theta \\ \sin n\theta & \cos n\theta \end{bmatrix}.$$

Geometrically, n rotations by θ give one rotation by $n\theta$.

34 Columns of X times rows of ΛX^{-1} gives a sum of r rank-1 matrices ($r =$ rank of A). Those matrices are $\lambda_1 \boldsymbol{x}_1 \boldsymbol{y}_1^{\mathrm{T}}$ to $\lambda_r \boldsymbol{x}_r \boldsymbol{y}_r^{\mathrm{T}}$.

35 Multiply $\text{ones}(n) * \text{ones}(n) = n * \text{ones}(n)$. Then

$$AA^{-1} = (\text{eye}(n) + \text{ones}(n)) * (\text{eye}(n) + C * \text{ones}(n))$$

$$= \text{eye}(n) + (1 + C + Cn) * \text{ones}(n) = \text{eye}(n) \text{ for } \boldsymbol{C = -1/(n+1)}.$$

36 $B = A_1^{-1}$ leads to $A_2 A_1 = B(A_1 A_2)B^{-1}$. Then $A_2 A_1$ is similar to $A_1 A_2$: they have the same eigenvectors (not zero because A_1 and A_2 are invertible).

37 Choose $B = A_1^{-1}$ to show that $A_2 A_1$ is **similar** to $A_1 A_2$. Assuming invertibility (no zero eigenvalues) this shows that $A_2 A_1$ and $A_1 A_2$ have the same eigenvalues.

38 This matrix has column $1 = 2$ (column 2) so $\boldsymbol{x}_1 = (1, -2, 0)$ is an eigenvector with $\lambda_1 = 0$. Also $A(1, 1, 1) = (1, 1, 1)$ and $\lambda_2 = 1$. Trace $=$ zero so $\lambda_3 = -1$. Then $1^{2020} = 1$ and $(-1)^{2020} = 1$ and $(0)^{2020} = 0$. So A^{2019} has the same eigenvalues and eigenvectors as A : $A^{2019} = A$ and $A^{2020} = A^2$. **TO COMPLETE FOR 2023**

Problem Set 6.3, page 238

1 (a) ASB stays symmetric like S when $B = A^{\mathrm{T}}$

(b) ASB is similar to S when $B = A^{-1}$

To have both (a) and (b) we need $B = A^{\mathrm{T}} = A^{-1}$ to be an **orthogonal matrix** Q. Then QSQ^{T} is similar to S and also symmetric like S.

2 $\lambda = 0, 4, -2$; unit vectors $\pm(0, 1, -1)/\sqrt{2}$ and $\pm(2, 1, 1)/\sqrt{6}$ and $\pm(1, -1, -1)/\sqrt{3}$. Those are for S. The eigenvalues of T are $\lambda = 0, \sqrt{5}, -\sqrt{5}$ in Λ (trace $= 0$). The eigenvectors of T are $\frac{1}{3}(2, 2, -1)$ and $(1 + \sqrt{5}, 1 - \sqrt{5}, 2)$ and $(1 - \sqrt{5}, 1 + \sqrt{5}, 4)$.

3 $S = \begin{bmatrix} 9 & 12 \\ 12 & 16 \end{bmatrix}$ has $\lambda = 0$ and 25 so the columns of Q are the two eigenvectors:

$Q = \begin{bmatrix} .8 & .6 \\ -.6 & .8 \end{bmatrix}$ or we can exchange columns or reverse the signs of any column.

4 (a) $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ has $\lambda = -1$ and 3 (b) The pivots $1, 1 - b^2$ have the same signs as the λ's

(c) The trace is $\lambda_1 + \lambda_2 = 2$, so S can't have two negative eigenvalues.

5 $(A^{\mathrm{T}}CA)^{\mathrm{T}} = A^{\mathrm{T}}C^{\mathrm{T}}(A^{\mathrm{T}})^{\mathrm{T}} = A^{\mathrm{T}}CA$. When A is 6 by 3, C will be 6 by 6 and the triple product $A^{\mathrm{T}}CA$ is 3 by 3.

6 $\lambda = 10$ and -5 in $\Lambda = \begin{bmatrix} 10 & 0 \\ 0 & -5 \end{bmatrix}$, $x = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ have to be normalized to

unit vectors in $Q = \dfrac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}$. Then $S = Q\Lambda Q^{\mathrm{T}}$.

If $A^3 = 0$ then all $\lambda^3 = 0$ so all $\lambda = 0$ as in $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. If A is *symmetric* then

$A^3 = Q\Lambda^3 Q^{\mathrm{T}} = 0$ requires $\Lambda = 0$. The only *symmetric* A is $Q \, 0 \, Q^{\mathrm{T}} = $ zero matrix.

7 $\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} = 2 \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} + 4 \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$; $\begin{bmatrix} 9 & 12 \\ 12 & 16 \end{bmatrix} = 0 \begin{bmatrix} .64 & -.48 \\ -.48 & .36 \end{bmatrix} + 25 \begin{bmatrix} .36 & .48 \\ .48 & .64 \end{bmatrix}$

8 $\begin{bmatrix} x_1 & x_2 \end{bmatrix}$ is an orthogonal matrix so $P_1 + P_2 = x_1 x_1^{\mathrm{T}} + x_2 x_2^{\mathrm{T}} =$

$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} x_1^{\mathrm{T}} \\ x_2^{\mathrm{T}} \end{bmatrix} = QQ^{\mathrm{T}} = I$; also $P_1 P_2 = x_1(x_1^{\mathrm{T}} x_2)x_2^{\mathrm{T}} =$ zero matrix.

Second proof: $P_1 P_2 = P_1(I - P_1) = P_1 - P_1 = 0$ since $P_1^2 = P_1$.

9 $A = \begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix}$ has $\lambda = ib$ and $-ib$. The block matrices $\begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix}$ and $\begin{bmatrix} 0 & A \\ A & 0 \end{bmatrix}$ are also skew-symmetric with $\lambda = ib$ (twice) and $\lambda = -ib$ (twice).

10 M is skew-symmetric and **orthogonal**; every λ is imaginary with $|\lambda| = 1$. So λ's must be $i, i, -i, -i$ to have trace zero.

11 $A = \begin{bmatrix} i & 1 \\ 1 & -i \end{bmatrix}$ has $\lambda = 0, 0$ and only one independent eigenvector $x = (i, 1)$. The good property for complex matrices is not $A^{\mathrm{T}} = A$ (symmetric) but $\overline{A}^{\mathrm{T}} = A$ (Hermitian with real eigenvalues and orthogonal eigenvectors).

12 S has $Q = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$; B has $X = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 2d \end{bmatrix}$. Perpendicular in Q Not perpendicular in X since $S^{\mathrm{T}} = S$ but $B^{\mathrm{T}} \neq B$

13 $S = \begin{bmatrix} 1 & 3+4i \\ 3-4i & 1 \end{bmatrix}$ is a *Hermitian matrix* $(\overline{S}^{\mathrm{T}} = S)$. Its eigenvalues 6 and -4 are *real*. Here is the proof that λ is always real when $\overline{S}^{\mathrm{T}} = S$:

$\qquad Sx = \lambda x$ leads to $\overline{S}\overline{x} = \overline{\lambda}\overline{x}$. Transpose to $\overline{x}^{\mathrm{T}} S = \overline{x}^{\mathrm{T}}\overline{\lambda}$ using $\overline{S}^{\mathrm{T}} = S$.

$\qquad$ Then $\overline{x}^{\mathrm{T}} Sx = \overline{x}^{\mathrm{T}}\lambda x$ and also $\overline{x}^{\mathrm{T}} Sx = \overline{x}^{\mathrm{T}}\overline{\lambda}x$. So $\lambda = \overline{\lambda}$ is real.

14 (a) False. $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ (b) True from $A^{\mathrm{T}} = Q\Lambda Q^{\mathrm{T}} = A$ (c) True from $S^{-1} = Q\Lambda^{-1}Q^{\mathrm{T}}$ (d) False!

(e) True. If x is a column of the identity matrix, then the energy $x^{\mathrm{T}} Sx$ is a diagonal entry of S. Since S is positive definite in this problem, each diagonal entry is a positive number $x^{\mathrm{T}} Sx$.

15 A and A^{T} have the same λ's but the *order* of the $\boldsymbol{x}$'s can change. $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$

has $\lambda_1 = i$ and $\lambda_2 = -i$ with $\boldsymbol{x}_1 = (1, i)$ first for A but $\boldsymbol{x}_1 = (1, -i)$ is first for A^{T}.

16 A is invertible, orthogonal, permutation, diagonalizable; B is projection, diagonalizable. A allows $QR, X\Lambda X^{-1}, Q\Lambda Q^{\mathrm{T}}$; B allows $X\Lambda X^{-1}$ and $Q\Lambda Q^{\mathrm{T}}$.

17 Symmetry gives $Q\Lambda Q^{\mathrm{T}}$ if $b = 1$; repeated λ and no X if $b = -1$; singular if $b = 0$.

18 Orthogonal and symmetric requires $|\lambda| = 1$ and λ real, so $\lambda = \pm 1$. Then $S = \pm I$ or

$$\pm S = Q\Lambda Q^{\mathrm{T}} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}.$$

19 Eigenvectors $(1, 0)$ and $(\mathbf{1}, \mathbf{1})$ give a $45°$ angle even with A^{T} very close to A.

20 a_{11} is $\begin{bmatrix} q_{11} & \ldots & q_{1n} \end{bmatrix} \begin{bmatrix} \lambda_1 \overline{q}_{11} & \ldots & \lambda_n \overline{q}_{1n} \end{bmatrix}^{\mathrm{T}} \le \lambda_{\max} \left(|q_{11}|^2 + \cdots + |q_{1n}|^2 \right) = \lambda_{\max}$.

21 (a) $\boldsymbol{x}^{\mathrm{T}}(A\boldsymbol{x}) = (A\boldsymbol{x})^{\mathrm{T}}\boldsymbol{x} = \boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}\boldsymbol{x} = -\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ so $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x} = 0$. (b) $\overline{\boldsymbol{z}}^{\mathrm{T}}A\boldsymbol{z}$ is pure imaginary, its real part is $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x} + \boldsymbol{y}^{\mathrm{T}}A\boldsymbol{y} = 0 + 0$ (c) $\det A = \lambda_1 \ldots \lambda_n \ge 0$: because pairs of λ's $= ib, -ib$ multiply to give $+b^2$.

22 Since S is diagonalizable with eigenvalue matrix $\Lambda = 2I$, the matrix S itself has to be $X\Lambda X^{-1} = X(2I)X^{-1} = 2I$. The unsymmetric matrix $[2\ 1\ ;\ 0\ 2]$ also has $\lambda = 2, 2$ but this matrix can't be diagonalized.

23 (a) $S^{\mathrm{T}} = S$ and $S^{\mathrm{T}}S = I$ lead to $S^2 = I$.

(b) The only possible eigenvalues of S are 1 and -1.

(c) $\Lambda = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}$ so $\boldsymbol{S} = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \Lambda \begin{bmatrix} Q_1^{\mathrm{T}} \\ Q_2^{\mathrm{T}} \end{bmatrix} = \boldsymbol{Q_1 Q_1^{\mathrm{T}}} - \boldsymbol{Q_2 Q_2^{\mathrm{T}}}$ with $Q_1^{\mathrm{T}}Q_2 = 0$.

24 Suppose $a > 0$ and $ac > b^2$ so that also $c > b^2/a > 0$.

(i) The eigenvalues have the *same sign* because $\lambda_1 \lambda_2 = \det = ac - b^2 > 0$.

(ii) That sign is *positive* because $\lambda_1 + \lambda_2 > 0$ (it equals the trace $a + c > 0$).

25 Only $S_4 = \begin{bmatrix} 1 & 10 \\ 10 & 101 \end{bmatrix}$ has two positive eigenvalues since $101 > 10^2$.

$\boldsymbol{x}^{\mathrm{T}}S_1\boldsymbol{x} = 5x_1^2 + 12x_1x_2 + 7x_2^2$ is negative for example when $x_1 = 4$ and $x_2 = -3$: A_1 is not positive definite as its determinant confirms; S_2 has trace c_0; S_3 has $\det = 0$.

26 Positive definite for $-3 < b < 3$

$$\begin{bmatrix} 1 & 0 \\ b & 1 \end{bmatrix} \begin{bmatrix} 1 & b \\ 0 & 9-b^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ b & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 9-b^2 \end{bmatrix} \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix} = LDL^{\mathrm{T}}$$

Positive definite for $c > 8$

$$\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 0 & c-8 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & c-8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = LDL^{\mathrm{T}}.$$

Positive definite for $c > |b|$

$$L = \begin{bmatrix} 1 & 0 \\ -b/c & 1 \end{bmatrix} \qquad D = \begin{bmatrix} c & 0 \\ 0 & c-b^2/c \end{bmatrix} \qquad S = LDL^{\mathrm{T}}.$$

27 $x^2 + 4xy + 3y^2 = (x+2y)^2 - y^2 = $ *difference of squares* is negative at $x = 2$, $y = -1$, where the first square is zero.

28 $S = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ produces $f(x,y) = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \mathbf{2xy}$. S has $\lambda = 1$ and $\lambda = -1$. Then S is an *indefinite matrix* and $f(x,y) = 2xy$ has a *saddle point*.

29 $A^{\mathrm{T}}A = \begin{bmatrix} 1 & 2 \\ 2 & 13 \end{bmatrix}$ and $A^{\mathrm{T}}A = \begin{bmatrix} 6 & 5 \\ 5 & 6 \end{bmatrix}$ are positive definite; $A^{\mathrm{T}}A = \begin{bmatrix} 2 & 3 & 3 \\ 3 & 5 & 4 \\ 3 & 4 & 5 \end{bmatrix}$ is singular (and positive semidefinite). The first two A's have independent columns. The 2 by 3 A cannot have full column rank 3, with only 2 rows; third $A^{\mathrm{T}}A$ is singular.

30 $S = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$ has pivots $2, \frac{3}{2}, \frac{4}{3}$; $T = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$ is singular; $T \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.

31 Corner determinants $|S_1| = 2$, $|S_2| = 6$, $|S_3| = 30$. The pivots are $2/1, 6/2, 30/6$.

32 S is positive definite for $c > 1$; determinants $c, c^2 - 1$, and $(c-1)^2(c+2) > 0$. T is *never* positive definite (determinants $d - 4$ and $-4d + 12$ are never both positive).

33 $S = \begin{bmatrix} 1 & 5 \\ 5 & 10 \end{bmatrix}$ is an example with $a + c > 2b$ but $ac < b^2$, so not positive definite.

34 The eigenvalues of S^{-1} are positive because they are $1/\lambda(S)$. Also the energy is $\boldsymbol{x}^{\mathrm{T}} S^{-1} \boldsymbol{x} = (S^{-1}\boldsymbol{x})^{\mathrm{T}} S (S^{-1}\boldsymbol{x}) > 0$ for all $\boldsymbol{x} \neq \mathbf{0}$.

35 $\boldsymbol{x}^{\mathrm{T}} S \boldsymbol{x}$ is zero when $(x_1, x_2, x_3) = (0, 1, 0)$ because of the zero on the diagonal. Actually $\boldsymbol{x}^{\mathrm{T}} S \boldsymbol{x}$ goes *negative* for $\boldsymbol{x} = (1, -10, 0)$ because the second pivot is *negative*.

36 If a_{jj} were smaller than all λ's, $S - a_{jj}I$ would have all eigenvalues > 0 (positive definite). But $S - a_{jj}I$ has a *zero* in the (j,j) position; impossible by Problem 35.

37 (a) The determinant is positive; all $\lambda > 0$ (b) All projection matrices except I are singular   (c) The diagonal entries of D are its eigenvalues

(d) $S = -I$ has $\det = +1$ when n is even, but this S is *negative* definite.

38 S is positive definite when $s > 8$; T is positive definite when $t > 5$ by determinants.

39 $A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{9} & \\ & \sqrt{1} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \dfrac{}{\sqrt{2}} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}; A = Q \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix} Q^{\mathrm{T}} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}.$

40 The ellipse $x^2 + xy + y^2 = 1$ comes from $S = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix}$ with $\lambda = \dfrac{1}{2}$ and $\dfrac{3}{2}$.

The axes have half-lengths $\sqrt{2}$ and $\sqrt{2/3}$.

41 $\begin{array}{c} S = C^{\mathrm{T}}C \\ \boldsymbol{S} \text{ not } \boldsymbol{A} \end{array} = \begin{bmatrix} 9 & 3 \\ 3 & 5 \end{bmatrix}; \begin{bmatrix} 4 & 8 \\ 8 & 25 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ and $C = \begin{bmatrix} 2 & 4 \\ 0 & 3 \end{bmatrix}$

42 The Cholesky factors $C = \left(L\sqrt{D} \right)^{\mathrm{T}} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix}$ and $C = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & \sqrt{5} \end{bmatrix}$ have

square roots of the pivots from D. Note again $C^{\mathrm{T}}C = LDL^{\mathrm{T}} = S$.

43 (a) $\det S = (1)(10)(1) = 10$;   (b) $\lambda = 2$ and 5;   (c) $\boldsymbol{x}_1 = (\cos\theta \sin\theta)$ and $\boldsymbol{x}_2 = (-\sin\theta, \cos\theta)$; (d) The λ's are positive, so S is positive definite.

44 $ax^2 + 2bxy + cy^2$ has a saddle point if $ac < b^2$. The matrix is *indefinite* ($\lambda < 0$ and $\lambda > 0$) because the determinant $ac - b^2$ is *negative*.

45 If $c > 9$ the graph of z is a bowl, if $c < 9$ the graph has a saddle point. When $c = 9$ the graph of $z = (2x + 3y)^2$ is a "trough" staying at zero along the line $2x + 3y = 0$.

46 A product ST of symmetric positive definite matrices comes into many applications. The "*generalized*" eigenvalue problem $K\boldsymbol{x} = \lambda M\boldsymbol{x}$ has $ST = M^{-1}K$. (Often we use **eig**(K, M) without actually inverting M.) All eigenvalues λ of ST are positive :

$ST\boldsymbol{x} = \lambda\boldsymbol{x}$ gives $(T\boldsymbol{x})^{\mathrm{T}}ST\boldsymbol{x} = (T\boldsymbol{x})^{\mathrm{T}}\lambda x$. Then $\lambda = \boldsymbol{x}^{\mathrm{T}}T^{\mathrm{T}}ST\boldsymbol{x}/\boldsymbol{x}^{\mathrm{T}}T\boldsymbol{x} > 0$.

47 Put parentheses in $\boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}CA\boldsymbol{x} = (A\boldsymbol{x})^{\mathrm{T}}C(A\boldsymbol{x})$. Since C is assumed positive definite, this energy can drop to zero only when $A\boldsymbol{x} = \boldsymbol{0}$. Sine A is assumed to have independent columns, $A\boldsymbol{x} = \boldsymbol{0}$ only happens when $\boldsymbol{x} = \boldsymbol{0}$. Thus $A^{\mathrm{T}}CA$ has positive energy and is positive definite.

My textbooks *Computational Science and Engineering* and *Introduction to Applied Mathematics* start with many examples of $A^{\mathrm{T}}CA$ in a wide range of applications. I believe positive definiteness of $A^{\mathrm{T}}CA$ is a unifying concept from linear algebra.

48 (a) The eigenvalues of $\lambda_1 I - S$ are $\lambda_1 - \lambda_1, \lambda_1 - \lambda_2, \ldots, \lambda_1 - \lambda_n$. Those are ≥ 0; $\lambda_1 I - S$ is semidefinite.

(b) Semidefinite matrices have energy $\boldsymbol{x}^{\mathrm{T}} \left(\lambda_1 I - S \right) \boldsymbol{x}_2 \geq 0$. Then $\lambda_1 \boldsymbol{x}^{\mathrm{T}} \boldsymbol{x} \geq \boldsymbol{x}^{\mathrm{T}} S \boldsymbol{x}$.

(c) Part (b) says $\boldsymbol{x}^{\mathrm{T}} S \boldsymbol{x} / \boldsymbol{x}^{\mathrm{T}} \boldsymbol{x} \leq \lambda_1$ for all $\boldsymbol{x}$. Equality at the eigenvector with $S\boldsymbol{x} = \lambda_1 \boldsymbol{x}$. So the maximum value of $\boldsymbol{x}^{\mathrm{T}} S \boldsymbol{x} / \boldsymbol{x}^{\mathrm{T}} \boldsymbol{x}$ is λ_1.

49 Energy $\boldsymbol{x}^{\mathrm{T}} S \boldsymbol{x} = a \left(x_1 + x_2 + x_3 \right)^2 + c \left(x_2 - x_3 \right)^2 \geq 0$ if $a \geq 0$ and $c \geq 0$: semidefinite.

S has rank ≤ 2 and determinant $= 0$; cannot be positive definite for any a and c.

Problem Set 6.4, page 269

1 $z = 1 - i$ leads to $\overline{z} = 1 + i$ and $r = \sqrt{2}$ and $\dfrac{1}{z} = \dfrac{1+i}{(1-i)(1+i)} = \dfrac{1}{2}(1+i)$

and $\theta = -\dfrac{\pi}{4} = -45°$.

2 $\det \begin{bmatrix} 1-\lambda & 1+i \\ 1-i & 2-\lambda \end{bmatrix} = \lambda^2 - 3\lambda + 2 - 2 = 0$ gives eigenvalues $\lambda = \mathbf{3}$ and $\mathbf{0}$.

3 If $Qx = \lambda x$ then $||Qx|| = |\lambda|\, ||x||$. Square both sides and use $\overline{Q}^{\mathrm{T}} Q = I$ to find $|\lambda|^2 = 1$. Therefore $|\lambda| = 1$ for unitary matrices Q.

4 $F_3 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & e^{2\pi 1/3} & e^{4\pi 1/3} \\ 1 & e^{4\pi 1/3} & e^{8\pi 1/3} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & \frac{1}{2}\left(-1+\sqrt{3}\,i\right) & \frac{1}{2}\left(-1-\sqrt{3}\,i\right) \\ 1 & \frac{1}{2}\left(-1-\sqrt{3}\,i\right) & \frac{1}{2}\left(-1+\sqrt{3}\,i\right) \end{bmatrix}$

5 $F_6 = 6$ by 6 matrix $= \begin{bmatrix} I & B \\ I & -B \end{bmatrix} \begin{bmatrix} F_3 & 0 \\ 0 & F_3 \end{bmatrix} \begin{bmatrix} \text{columns} \\ 0,2,4,1,3,5 \\ \text{of } I \text{ (6 by 6)} \end{bmatrix}$

The 3 by 3 matrix B is diagonal with entries $1, e^{2\pi i/6}, e^{4\pi i/6}$.

6 $CD = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 2 \\ 2 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 4 & 4 \\ 4 & 4 & 4 \\ 4 & 4 & 4 \end{bmatrix}$

$$\begin{array}{ccccc} & & 1 & 2 & 1 \\ & & 1 & 1 & 1 \\ \hline & & 1 & 2 & 1 \\ & 1 & 2 & 1 & \\ 1 & 2 & 1 & & \\ \hline \end{array}$$

convolution $c * d$ 1 3 4 3 1   reduces to 4 4 4 for cyclic convolution $c \circledast d$

7 Convolution Rule $F(c \circledast d) = (Fc) \mathbin{.*} (Fd)$. This is $F \begin{bmatrix} 4 \\ 4 \\ 4 \end{bmatrix} = F \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \mathbin{.*} F \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$

with the 3 by 3 Fourier matrix $F = F_3$: **Multiply components for .$*$.**

$$F \begin{bmatrix} 4 \\ 4 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & e^{2\pi i/3} & e^{4\pi i/3} \\ 1 & e^{4\pi i/3} & e^{8\pi i/3} \end{bmatrix} \begin{bmatrix} 4 \\ 4 \\ 4 \end{bmatrix} = \begin{bmatrix} 12 \\ 0 \\ 0 \end{bmatrix}$$

$$F \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} \qquad F \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ e^{2\pi i/3} \\ e^{4\pi i/3} \end{bmatrix} \text{ and } \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} .* \begin{bmatrix} 4 \\ e^{2\pi i/3} \\ e^{4\pi i/3} \end{bmatrix} \begin{bmatrix} 12 \\ 0 \\ 0 \end{bmatrix}$$

8 $\cos\theta + i\sin\theta = \left(1 - \dfrac{1}{2}\theta^2 + \cdots\right) + i\left(\theta - \dfrac{\theta^3}{6} + \cdots\right) = 1 + i\theta + \dfrac{1}{2}(i\theta)^2 + \dfrac{1}{6}(i\theta)^3 + \cdots$

9 $(e^{i\theta})(e^{i\theta}) = e^{2i\theta}$ is $(\cos\theta + i\sin\theta)^2 = \cos 2\theta + i\sin 2\theta$.

The left side is $\cos^2\theta + 2i\cos\theta\sin\theta + i^2\sin^2\theta$.

Matching the right side gives $\cos 2\theta = \cos^2\theta - \sin^2\theta$ and $\sin 2\theta = 2\cos\theta\sin\theta$

10 The eigenvalues of a circulant matrix C are $F\mathbf{c}$ in equation (10).

If C is invertible then all its eigenvalues must be nonzero.

In that case C^{-1} is also a circulant because its entries (from the formula for C^{-1}) are also constant down each (cyclic) diagonal. There are other proofs too.

11 This problem is looking for a solution !

12 An n by n circulant matrix has $\overline{C}^{\mathrm{T}} = C$ (Hermitian) if its diagonal entries have c_0 real, $\overline{c}_1 = c_{n-1}, \overline{c}_2 = c_{n-2}, \ldots$ The circulant has $\overline{C}^{\mathrm{T}}C = I$ (unitary) if $|c_0 + c_1 x + \cdots + c_{n-1}x^{n-1}|^2 = 1$.

13 Columns 0 and 2 of the Fourier matrix F_4 in equation (7) add to $(2, 0, 2, 0)$. Columns 1 and 3 add to $(2, 0, -2, 0)$.

14 $z = w^2 = e^{2\pi i/32}$ would be a 32nd root of $1 : z^{32} = 1$.

$z = \sqrt{w} = e^{2\pi i/128}$ would be a 128th root of 1.

15 The 4 eigenvalues $0, 2, 4, 2$ of C come from the eigenvalues $1, i, -1, -i$ of P_4.

$\lambda = 2 - 1 - 1 = 0 \qquad \lambda = 2 - i - i^3 = 2 \qquad \lambda = 2 - (-1) - (-1)^3 = 4 \qquad \lambda = 2 + i + i^3 = 2.$

Problem Set 6.5, page 280

1 Eigenvalues 4 and 1 with eigenvectors $(1,0)$ and $(1,-1)$ give solutions $\boldsymbol{u}_1 = e^{4t} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

and $\boldsymbol{u}_2 = e^t \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. If $\boldsymbol{u}(0) = \begin{bmatrix} 5 \\ -2 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 2 \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, then use those

coefficients 3 and 2: $\boldsymbol{u}(t) = 3e^{4t} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 2e^t \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

2 $z(t) = 2e^t$ solves $dz/dt = z$ with $z(0) = 2$. Then $dy/dt = 4y - 6e^t$ with $y(0) = 5$

gives $y(t) = 3e^{4t} + 2e^t$ as in Problem 1.

3 (a) If every column of A adds to zero, this means that the rows add to the zero row.

So the rows are dependent, and A is singular, and $\lambda = 0$ is an eigenvalue.

(b) The eigenvalues of $A = \begin{bmatrix} -2 & 3 \\ 2 & -3 \end{bmatrix}$ are $\lambda_1 = 0$ with eigenvector $\boldsymbol{x}_1 = (3,2)$ and

$\lambda_2 = -5$ (to give trace $= -5$) with $\boldsymbol{x}_2 = (1,-1)$. Then the usual 3 steps:

1. Write $u(0) = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$ as $\begin{bmatrix} 3 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \boldsymbol{x}_1 + \boldsymbol{x}_2 =$ combination of eigenvectors

2. The solutions follow those eigenvectors: $e^{0t}\boldsymbol{x}_1$ and $e^{-5t}\boldsymbol{x}_2$

3. The solution $\boldsymbol{u}(t) = \boldsymbol{x}_1 + e^{-5t}\boldsymbol{x}_2$ has steady state $\boldsymbol{x}_1 = (3,2)$ since $e^{-5t} \to 0$.

4 $d(v+w)/dt = (w-v) + (v-w) = 0$, so the total $v+w$ is constant.

$A = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$ has $\begin{matrix} \lambda_1 = 0 \\ \lambda_2 = -2 \end{matrix}$ with $\boldsymbol{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\boldsymbol{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

$\begin{bmatrix} v(0) \\ w(0) \end{bmatrix} = \begin{bmatrix} 30 \\ 10 \end{bmatrix} = 20 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 10 \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ leads to $\begin{matrix} v(1) = 20 + 10e^{-2} & v(\infty) = 20 \\ w(1) = 20 - 10e^{-2} & w(\infty) = 20 \end{matrix}$

5 $\dfrac{d}{dt} \begin{bmatrix} v \\ w \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ has $\lambda = 0$ and $\lambda = +\mathbf{2}$: $v(t) = \mathbf{20 + 10e^{2t}} \to -\infty$ as

$t \to \infty$.

6 $A = \begin{bmatrix} a & 1 \\ 1 & a \end{bmatrix}$ has real eigenvalues $a+1$ and $a-1$. These are both negative if $\boldsymbol{a < -1}$.

In this case the solutions of $d\boldsymbol{u}/dt = A\boldsymbol{u}$ approach zero.

$B = \begin{bmatrix} b & -1 \\ 1 & b \end{bmatrix}$ has complex eigenvalues $b+i$ and $b-i$. These have negative real parts

if $\boldsymbol{b < 0}$. In this case all solutions of $d\boldsymbol{v}/dt = B\boldsymbol{v}$ approach zero.

7 A projection matrix has eigenvalues $\lambda = 1$ and $\lambda = 0$. Eigenvectors $P\boldsymbol{x} = \boldsymbol{x}$ fill the

subspace that P projects onto: here $\boldsymbol{x} = (c, c)$. Eigenvectors with $P\boldsymbol{x} = \boldsymbol{0}$ fill the

perpendicular subspace: here $\boldsymbol{x} = (c, -c)$. For the solution to $d\boldsymbol{u}/dt = -P\boldsymbol{u}$,

$$\boldsymbol{u}(0) = \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix} \qquad \boldsymbol{u}(t) = e^{-t} \begin{bmatrix} 2 \\ 2 \end{bmatrix} + e^{0t} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \text{ approaches } \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

8 $\begin{bmatrix} 6 & -2 \\ 2 & 1 \end{bmatrix}$ has $\lambda_1 = 5$, $\boldsymbol{x}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $\lambda_2 = 2$, $\boldsymbol{x}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$; rabbits $r(t) = 20e^{5t} + 10e^{2t}$,

$w(t) = 10e^{5t} + 20e^{2t}$. The ratio of rabbits to wolves approaches $20/10$; (somewhat

against nature) e^{5t} dominates.

9 (a) $\begin{bmatrix} 4 \\ 0 \end{bmatrix} = 2\begin{bmatrix} 1 \\ i \end{bmatrix} + 2\begin{bmatrix} 1 \\ -i \end{bmatrix}$.   (b) Then $u(t) = 2e^{it}\begin{bmatrix} 1 \\ i \end{bmatrix} + 2e^{-it}\begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} 4\cos t \\ 4\sin t \end{bmatrix}$.

10 $\dfrac{d}{dt}\begin{bmatrix} y \\ y' \end{bmatrix} = \begin{bmatrix} y' \\ y'' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 4 & 5 \end{bmatrix}\begin{bmatrix} y \\ y' \end{bmatrix}$. This correctly gives $y' = y'$ and $y'' = 4y + 5y'$.

$A = \begin{bmatrix} 0 & 1 \\ 4 & 5 \end{bmatrix}$ has $\det(A - \lambda I) = \lambda^2 - 5\lambda - 4 = 0$. Directly substituting $y = e^{\lambda t}$ into

$y'' = 5y' + 4y$ also gives $\lambda^2 = 5\lambda + 4$ and the same two values of λ. Those values are

$\frac{1}{2}(5 \pm \sqrt{41})$ by the quadratic formula.

11 The series for e^{At} is $e^{At} = I + t \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \text{ zeros } = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$. Then

$$\begin{bmatrix} y(t) \\ y'(t) \end{bmatrix} = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y(0) \\ y'(0) \end{bmatrix} = \begin{bmatrix} y(0) + y'(0)t \\ y'(0) \end{bmatrix}. \text{ This } y(t) = y(0) + y'(0)t \text{ solves}$$

the equation—the factor t tells us that A had only one eigenvector: not diagonalizable.

12 $A = \begin{bmatrix} 0 & 1 \\ -9 & 6 \end{bmatrix}$ has trace 6, det 9, $\lambda = 3$ and 3 with *one* independent eigenvector

$(1, 3)$. Substitute $y = te^{3t}$ to show that this gives the needed second solution ($y = e^{3t}$

is the first solution).

13 (a) $y(t) = \cos 3t$ and $\sin 3t$ solve $y'' = -9y$. It is $\mathbf{3} \cos 3t$ that starts with $y(0) = 3$ and

$y'(0) = 0.$ (b) $A = \begin{bmatrix} 0 & 1 \\ -9 & 0 \end{bmatrix}$ has det $= 9$: $\lambda = 3i$ and $-3i$ with eigenvectors

$$x = \begin{bmatrix} 1 \\ 3i \end{bmatrix} \text{ and } \begin{bmatrix} 1 \\ -3i \end{bmatrix}. \text{ Then } \boldsymbol{u}(t) = \frac{3}{2}e^{3it} \begin{bmatrix} 1 \\ 3i \end{bmatrix} + \frac{3}{2}e^{-3it} \begin{bmatrix} 1 \\ -3i \end{bmatrix} = \begin{bmatrix} \mathbf{3} \cos \mathbf{3}t \\ -\mathbf{9} \sin \mathbf{3}t \end{bmatrix}.$$

14 When A is skew-symmetric, the derivative of $||u(t)||^2$ is *zero*. Then $||\boldsymbol{u}(t)|| = ||e^{At}\boldsymbol{u}(0)||$

stays at $||\boldsymbol{u}(0)||$. So the matrix e^{At} is *orthogonal* when A is skew-symmetric ($A^T = -A$).

15 $\boldsymbol{u}_p = 4$ and $\boldsymbol{u}(t) = ce^t + 4$. For the matrix equation, the particular solution $\boldsymbol{u}_p = A^{-1}\boldsymbol{b}$

is $\begin{bmatrix} 4 \\ 2 \end{bmatrix}$ and $\boldsymbol{u}(t) = c_1 e^t \begin{bmatrix} 1 \\ t \end{bmatrix} + c_2 e^t \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 4 \\ 2 \end{bmatrix}$.

16 $d/dt(e^{At}) = A + A^2 t + \frac{1}{2}A^3 t^2 + \frac{1}{6}A^4 t^3 + \cdots = A(I + At + \frac{1}{2}A^2 t^2 + \frac{1}{6}A^3 t^3 + \cdots)$.

This is exactly Ae^{At}, the derivative we expect from e^{At}.

17 $e^{Bt} = I + Bt$ (short series with $B^2 = 0$) $= \begin{bmatrix} \mathbf{1} & -\mathbf{4}t \\ \mathbf{0} & \mathbf{1} \end{bmatrix}$. Derivative $= \begin{bmatrix} 0 & -4 \\ 0 & 0 \end{bmatrix} =$

$Be^{Bt} = B$ in this example.

18 The solution at time $t + T$ is $e^{A(t+T)}\boldsymbol{u}(0)$. Thus e^{At} times e^{AT} equals $e^{A(t+T)}$.

19 $A^2 = A$ gives $e^{At} = I + At + \frac{1}{2}\boldsymbol{At^2} + \frac{1}{6}\boldsymbol{At^3} + \cdots = I + (e^t - 1)A$.

20 $e^A = \begin{bmatrix} e & 4(e-1) \\ 0 & 1 \end{bmatrix}$ from **21** and $e^B = \begin{bmatrix} 1 & -4 \\ 0 & 1 \end{bmatrix}$ from **19**. By direct multiplication

$e^A e^B \neq e^B e^A \neq e^{A+B} = \begin{bmatrix} e & 0 \\ 0 & 1 \end{bmatrix}$.

21 The matrix has $A^2 = \begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix}^2 = \begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix} = A$. *Then all* $A^n = A$. So $e^{At} =$

$I + (t + t^2/2! + \cdots)A = I + (e^t - 1)A = \begin{bmatrix} e^t & 3(e^t - 1) \\ 0 & 0 \end{bmatrix}$ as in Problem 19.

22 (a) The inverse of e^{At} is e^{-At} (b) If $A\boldsymbol{x} = \lambda\boldsymbol{x}$ then $e^{At}\boldsymbol{x} = e^{\lambda t}\boldsymbol{x}$ and $e^{\lambda t} \neq 0$.

To see $e^{At}\boldsymbol{x}$, write $(I + At + \frac{1}{2}A^2 t^2 + \cdots)\boldsymbol{x} = (1 + \lambda t + \frac{1}{2}\lambda^2 t^2 + \cdots)\boldsymbol{x} = e^{\lambda t}\boldsymbol{x}$.

23 Invert $\begin{bmatrix} 1 & 0 \\ \Delta t & 1 \end{bmatrix}$ to produce $\boldsymbol{U}_{n+1} = \begin{bmatrix} 1 & 0 \\ -\Delta t & 1 \end{bmatrix}\begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix}\boldsymbol{U}_n = \begin{bmatrix} 1 & \Delta t \\ -\Delta t & 1 - (\Delta t)^2 \end{bmatrix}\boldsymbol{U}_n$.

At $\Delta t = 1$, $\begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}$ has $\lambda = e^{i\pi/3}$ and $e^{-i\pi/3}$. Both eigenvalues have $\lambda^6 = 1$ so

$\boldsymbol{A^6 = I}$. Therefore $\boldsymbol{U}_6 = A^6 \boldsymbol{U}_0$ comes exactly back to $\boldsymbol{U}_0$.

24 First A has $\lambda = \pm i$ and $A^4 = I$.

Second A has $\lambda = -1, -1$ and $A^n = (-1)^n \begin{bmatrix} 1 - 2n & -2n \\ 2n & 2n + 1 \end{bmatrix}$ Linear growth.

25 With $a = \Delta t/2$ the trapezoidal step is $\boldsymbol{U}_{n+1} = \dfrac{1}{1 + a^2}\begin{bmatrix} 1 - a^2 & 2a \\ -2a & 1 - a^2 \end{bmatrix}\boldsymbol{U}_n$.

That matrix has orthonormal columns $\Rightarrow$ orthogonal matrix $\Rightarrow \|\boldsymbol{U}_{n+1}\| = \|\boldsymbol{U}_n\|$

26 For proof 2, square the start of the series to see $(I + A + \frac{1}{2}A^2 + \frac{1}{6}A^3)^2 = I + 2A + \frac{1}{2}(2A)^2 + \frac{1}{6}(2A)^3 + \cdots$. The diagonalizing proof is easiest when it works (but it needs a diagonalizable A).

Problem Set 7.1, page 295

1 $A^{\mathrm{T}}A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 64 \end{bmatrix}$ $AA^{\mathrm{T}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 64 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ give $\sigma_1 = 8$ and $\sigma_2 = 1$.

$v_1 = (0,0,1)$, $v_2 = (0,1,0)$, $u_1 = (0,1,0)$, $u_2 = (1,0,4)$. After removing

row 3 of A and column 3 of A^{T}, $\begin{bmatrix} 1 & 0 \\ 0 & 64 \end{bmatrix}$ still has $\sigma_1^2 = 64$ and $\sigma_2^2 = 1$.

2 $\det(B - \lambda I) = -\lambda^3 + \frac{1}{125} = 0$ gives $\lambda = \frac{1}{5}$ times 1 and $e^{2\pi i/3}$ and $e^{4\pi i/3}$.

The singular values are $\sigma = 8$ and 1 and $1/1000$. So λ changed by $1/5$ and σ only

changed by $1/1000$.

3 A^{T} has the same singular values as A, and the singular vectors change from $Av = \sigma u$

to $Au = \sigma v$.

4 $\begin{bmatrix} 0 & A \\ A^{\mathrm{T}} & 0 \end{bmatrix} \begin{bmatrix} u_k \\ v_k \end{bmatrix} = \begin{bmatrix} Av_k \\ A^{\mathrm{T}}u_k \end{bmatrix} = \sigma_k \begin{bmatrix} u_k \\ v_k \end{bmatrix}$ and $\begin{bmatrix} 0 & A \\ A^{\mathrm{T}} & 0 \end{bmatrix} \begin{bmatrix} -u_k \\ v_k \end{bmatrix} = \begin{bmatrix} Av_k \\ -A^{\mathrm{T}}u_k \end{bmatrix} = -\sigma_k \begin{bmatrix} -u_k \\ v_k \end{bmatrix}$

So this one symmetric matrix S reveals the u's and v's and σ's in the SVD of A.

5 $A^{\mathrm{T}}A$ is symmetric with $\lambda_1 = 25$ and $\lambda_2 = 0$ so A has $\sigma_1 = 5$. The eigenvectors of

$A^{\mathrm{T}}A$ are $v_1 = (2,1)$ and $v_2 = (-1,2)$: *orthogonal*. They are the v's in $A = U\Sigma V^{\mathrm{T}}$.

6 $A_1 A_1^{\mathrm{T}} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$ produces $\lambda^2 - 3\lambda + 1 = 0$ and

$\lambda = \frac{1}{2}\left(3 \pm \sqrt{5}\right)$. The singular values are the square roots $\sigma = \frac{1}{2}\left(\sqrt{5} \pm 1\right)$.

$A_2 A_2^{\mathrm{T}} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix}$ has $\lambda^2 - 6\lambda + 4 = 0$ and

$\lambda = \frac{1}{2}\left(6 \pm \sqrt{20}\right) = 3 \pm \sqrt{5}$. The singular values are the square roots $\sigma = \frac{\sqrt{2}}{2}\left(\sqrt{5} \pm 1\right)$.

For the singular vectors I recommend the SVD commands in MATLAB or Julia or

Mathematica.

7 There are 20 singular values because a random 20 by 40 matrix almost surely has rank 20.

8 (a) The singular values of $A + I$ are square roots of eigenvalues of $(A + I)^{\mathrm{T}}(A + I)$. They are **not** eigenvalues of $A^{\mathrm{T}}A + I$.

(b) **This formula $V\Sigma^{-1}U^{\mathrm{T}}$ is the best way to compute the pseudoinverse A^{+}.** We could check the four Penrose conditions on A^{+} from Section 4.5. For example

$$AA^{+}A = \left(\Sigma\, \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^{\mathrm{T}}\right)\left(\Sigma\, \boldsymbol{v}_j \boldsymbol{u}_j^{\mathrm{T}}/\sigma_j\right)\left(\Sigma\, \sigma_k \boldsymbol{u}_k \boldsymbol{v}_k^{\mathrm{T}}\right) = \Sigma\, \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^{\mathrm{T}}.$$

Notice also that $AA^{+} = \Sigma\, \boldsymbol{u}_i \boldsymbol{v}_i^{\mathrm{T}} \boldsymbol{v}_i \boldsymbol{u}_i = \Sigma\, \boldsymbol{u}_i^{\mathrm{T}} \boldsymbol{u}_i = U^{\mathrm{T}}U = $ **projection**.

9 The singular values of Q are the positive square roots of eigenvalues of $Q^{\mathrm{T}}Q$—and all those eigenvalues are 1 because $Q^{\mathrm{T}}Q = I$ when Q is orthogonal.

10 If the λ's are in descending order, the maximum of $R(\boldsymbol{x}) = (\lambda_1 c_1^2 + \cdots + \lambda_n c_n^2)/(c_1^2 + \cdots + c_n^2)$ is λ_1 (when $\boldsymbol{x} = \boldsymbol{v}_1$). Then $c_1, c_2, \ldots, c_n$ is $1, 0, \ldots, 0$. The minimum is $R(\boldsymbol{x}) = \lambda_n$ when $\boldsymbol{x} = \boldsymbol{v}_n$ and $\boldsymbol{c} = (0, 0, \ldots, 0, 1)$.

11 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{v}_1 = 0$ means that the coefficient is $\boldsymbol{c_1} = \boldsymbol{0}$ in $\boldsymbol{x} = c_1 \boldsymbol{v}_1 + \cdots + c_n \boldsymbol{v}_n$. Then
$$\max \frac{\lambda_2 c_2^2 + \cdots + \lambda_n c_n^2}{c_2^2 + \cdots + c_n^2} = \boldsymbol{\lambda_2}.$$

12 The first matrix has $A^{\mathrm{T}}A = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}$ with $\lambda = 8$ and $\lambda = 2$. The eigenvectors of $A^{\mathrm{T}}A = $ right singular vectors $\boldsymbol{v}_1, \boldsymbol{v}_2$ of A are $(1, 1)/\sqrt{2}$ and $(1, -1)/\sqrt{2}$. The left singular vectors are $\boldsymbol{u} = A\boldsymbol{v}/\sigma = (4, 0)/\sqrt{2}\sqrt{8} = (1, 0)$ and $(0, 2)/\sqrt{2}\sqrt{2} = (0, 1)$.

The second matrix has $A^{\mathrm{T}}A = \begin{bmatrix} 25 & 25 \\ 25 & 25 \end{bmatrix}$ so $\lambda = 50$ and $\lambda = 0$. The right singular vectors of A are again $\boldsymbol{v}_1 = (1, 1)/\sqrt{2}$ with $\sigma_1 = \sqrt{50}$ and $\boldsymbol{v}_2 = (1, -1)/\sqrt{2}$ with no σ_2 (or you could say $\sigma_2 = 0$ but our convention is no σ_2). Then $\boldsymbol{u}_1 = A\boldsymbol{v}_1/\sqrt{50} = (3, 4)/5$.

13 This matrix has $A^{\mathrm{T}}A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ with eigenvalues $\lambda = 3, 1, 0$ and $\sigma_1 = \sqrt{3}$ and $\sigma_2 = 1$ and no σ_3. The eigenvectors of $A^{\mathrm{T}}A$ are $\boldsymbol{v}_1 = (1, 2, 1)/\sqrt{6}$ and

$v_2 = (1, 0, -1)/\sqrt{2}$ and $v_3 = (1, -1, 1)/\sqrt{3}$. Then $Av = \sigma u$ gives $u = (1, 1)/\sqrt{2}$ and $u_2 = (1, -1)/\sqrt{2}$.

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 1 & 0 & -1 \end{bmatrix} \begin{matrix} /\sqrt{6} \\ /\sqrt{2} \end{matrix}$$

14 This small question is a key to everything. It is based on the associative law $(AA^{\mathrm{T}})A = A(A^{\mathrm{T}}A)$. Here we are applying both sides to an eigenvector v of $A^{\mathrm{T}}A$:

$$(AA^{\mathrm{T}})Av = A(A^{\mathrm{T}}A)v = A\lambda v = \lambda Av.$$

So Av **is an eigenvector of** AA^{T} with the same eigenvalue $\lambda = \sigma^2$.

15 $A = U\Sigma V^{\mathrm{T}} = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & \\ & 0 \end{bmatrix} \begin{bmatrix} v_1 & v_2 \end{bmatrix}^{\mathrm{T}} = \dfrac{\begin{bmatrix} 1 & 3 \\ 3 & -1 \end{bmatrix}}{\sqrt{10}} \begin{bmatrix} \sqrt{50} & 0 \\ 0 & 0 \end{bmatrix} \dfrac{\begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}}{\sqrt{5}}$

This $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$ is a 2 by 2 matrix of rank 1. Its row space has basis v_1, its nullspace has basis v_2, its column space has basis u_1, its left nullspace has basis u_2:

$$\text{Row space} \quad \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad \text{Nullspace} \quad \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

$$\text{Column space} \quad \frac{1}{\sqrt{10}} \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \quad \mathbf{N}(A^{\mathrm{T}}) \quad \frac{1}{\sqrt{10}} \begin{bmatrix} 3 \\ -1 \end{bmatrix}.$$

16 (a) The main diagonal of $A^{\mathrm{T}}A$ contains the squared lengths $||\text{row } 1||^2, \cdots, ||\text{row } m||^2$. So the trace of $A^{\mathrm{T}}A$ is the sum of all a_{ij}^2.

(b) If A has rank 1, then $A^{\mathrm{T}}A$ has rank 1. So the only singular value of A is $\sigma_1 = (\text{trace } A^{\mathrm{T}}A)^{1/2}$.

17 The number $\sigma_{\max}(A^{-1})\sigma_{\max}(A)$ is the same as $\sigma_{\max}(A)/\sigma_{\min}(A)$. This is ≥ 1. It equals 1 if all σ's are equal, and $A = U\Sigma V^{\mathrm{T}}$ is a multiple of an orthogonal matrix. The ratio $\sigma_{\max}/\sigma_{\min}$ is the important **condition number** of A.

18 The smallest change in A is to set its smallest singular value σ_2 to zero.

Problem Set 7.2, page 301

1 (a) Suppose the identity matrix I is N by N, and an N by N approximating matrix A has rank $r < N$. Then $I - A$ will have $N - r$ eigenvalues equal to 1, meaning that **the error norm $||I - A||$ is at least** 1, and I is impossible to compress by a lower rank matrix.

(b) A matrix with a horizontal-vertical cross looks like A :

$$\begin{bmatrix} \text{zeros} & \text{ones} & \text{zeros} \\ \text{ones} & \text{ones} & \text{ones} \\ \text{zeros} & \text{ones} & \text{zeros} \end{bmatrix} = \begin{bmatrix} \text{zeros} & \text{ones} & \text{zeros} \\ \text{zeros} & \text{ones} & \text{zeros} \\ \text{zeros} & \text{ones} & \text{zeros} \end{bmatrix} + \begin{bmatrix} \text{zeros} & \text{zeros} & \text{zeros} \\ \text{ones} & \text{zeros} & \text{ones} \\ \text{zeros} & \text{zeros} & \text{zeros} \end{bmatrix}$$

Those are both rank one matrices (all nonzero rows equal) so **A has rank 2**.

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

and the rank is 2.

$B = \begin{bmatrix} 1 & 2 & 2 \\ 1 & 3 & 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \end{bmatrix}$ also has rank 2.

3 $BB^{\mathrm{T}} = \begin{bmatrix} 1 & 2 & 2 \\ 1 & 3 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 9 & 13 \\ 13 & 19 \end{bmatrix}$ trace $= \mathbf{28}$ and $\det = \mathbf{2}$.

$B^{\mathrm{T}}B = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 \\ 1 & 3 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 5 & 5 \\ 5 & 13 & 13 \\ 5 & 13 & 13 \end{bmatrix}$ trace $= \mathbf{28}$ and $\det = \mathbf{0}$.
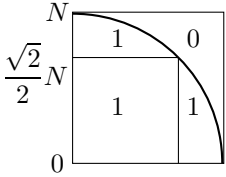
The 2 nonzero eigenvalues must be the same for both matrices. They are $\sigma_1, \ \sigma_2 = 14 \pm \sqrt{14^2 - 2}$. I would call B compressible when σ_2 is so much smaller than σ_1.

4 (computer question $\mathsf{svd}(A)$).

5 The Japanese flag has a circle filled by 1's, with diameter $= 2N$ 1's. Outside the circle are zeros. The rank is approximately CN. What is the number C? Alex Townsend

contributed this key idea: The circle contains a big square matrix filled by 1's. The rank of that all-ones matrix is only 1.

So we only have to count the rows above and below that square! Multiply by 2 to include the columns to the left and right of the square.

The picture shows $\left(1 - \dfrac{\sqrt{2}}{2}\right) N$ rows of 1's above the square—and repeated below the square. It also shows $\left(1 - \dfrac{\sqrt{2}}{2}\right) N$ columns of 1's to the right of the square—and repeated to the left.

Combined, those $(2 - \sqrt{2})N$ rows and columns (plus 1 for the big square) tell us the rank of this $2N$ by $2N$ Japanese flag containing the red circle.

6 The N by N matrix A is filled by the values $A_{ij} = F(i/N, j/N)$ of the two-variable function $F(x, y)$, by taking the points $(x, y) = (i/N, j/N)$ on a uniform square grid (x and y go from 0 to 1). Three choices of that function F:

1) $F = xy$ produces a symmetric **rank-1 matrix**. Its i, j entry is a multiple of the product i times j. All rows of F contain a multiple of the vector $(1, 2, \ldots, N)$.

2) $F_2 = x + y$ gives a sum of 2 rank-one matrices (**the rank is 2**). One matrix has constants along each row. The other has constants down each column.

3) $F_3 = (x, y) = x^2 + y^2$ will also produce a sum of constant rows (from x^2) and constant columns (from y^2). Again rank $= 2$.

7 Symmetric matrix S if $F(x, y) = F(y, x)$. Example $F = x + y$.

Antisymmetric matrix A if $F(x, y) = -F(y, x)$. Example $F = x - y$.

Matrix of rank 2 if $F(x, y) = F(x) + F(y)$ (and other possibilities too?)

Singular matrix M from a sum of less than n rank-one matrices (please expand this part of the answer).

Problem Set 7.3, page 307

1 The row averages of A_0 are 3 and 0. Therefore

$$A = \begin{bmatrix} 2 & 1 & 0 & -1 & -2 \\ -1 & 1 & 0 & 1 & -1 \end{bmatrix} \quad \text{and} \quad S = \frac{AA^{\text{T}}}{4} = \frac{1}{4}\begin{bmatrix} 10 & 0 \\ 0 & 4 \end{bmatrix}$$

The eigenvalues of S are $\lambda_1 = \dfrac{10}{4}$ and $\lambda_2 = \dfrac{4}{4} = 1$. The top eigenvector of S is

$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. I think this means that a **horizontal line** (the x axis) is closer to the five points

$(2,-1),\ldots,(-2,-1)$ in the columns of A than any other line through the origin $(0,0)$.

2 Now the row averages of A_0 are $\frac{1}{2}$ and 2. Therefore

$$A = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -1 & 0 & 1 & 1 & 0 & -1 \end{bmatrix} \quad \text{and} \quad S = \frac{AA^{\text{T}}}{5} = \frac{1}{5}\begin{bmatrix} \frac{3}{2} & 0 \\ 0 & 4 \end{bmatrix}.$$

Again the rows of A are accidentally orthogonal (because of the special patterns of those rows). This time the top eigenvector of S is $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$. So a **horizontal line** is closer to the six points $\left(\frac{1}{2},-1\right),\ldots,\left(-\frac{1}{2},-1\right)$ from the columns of A than any other line through the center point $(0,0)$.

3 $A_0 = \begin{bmatrix} 1 & 2 & 3 \\ 5 & 2 & 2 \end{bmatrix}$ has row averages 2 and 3 so $A = \begin{bmatrix} -1 & 0 & 1 \\ 2 & -1 & -1 \end{bmatrix}$.

Then $S = \dfrac{1}{2}AA^{\text{T}} = \dfrac{1}{2}\begin{bmatrix} 2 & -3 \\ -3 & 6 \end{bmatrix}$.

Then trace $(S) = \frac{1}{2}(8)$ and $\det(S) = \left(\frac{1}{2}\right)^2(3)$. The eigenvalues $\lambda(S)$ are $\frac{1}{2}$ times the roots of $\lambda^2 - 8\lambda + 3 = 0$. Those roots are $4 \pm \sqrt{16-3}$. Then the σ's are $\sqrt{\lambda_1}$ and $\sqrt{\lambda_2}$.

4 This matrix A with orthogonal rows has $S = \dfrac{AA^{\text{T}}}{n-1} = \dfrac{1}{3}\begin{bmatrix} 2 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 4 \end{bmatrix}$.

With λ's in descending order $\lambda_1 > \lambda_2 > \lambda_3$, the eigenvectors are $(0, 1, 0)$ and $(0, 0, 1)$ and $(1, 0, 0)$. The first eigenvector shows the $\boldsymbol{u}_1$ direction $= y$ axis. Combined with the second eigenvector $\boldsymbol{u}_2$ in the z direction, the best plane is the yz plane.

These problems are examples where the sample **correlation matrix** (rescaling S so all its diagonal entries are 1) would be the identity matrix. If we think the original scaling is not meaningful and the rows should have the same length, then there is no reason to choose $\boldsymbol{u}_1 = (0, 1, 0)$ from the 8 in row 2.

5 Recall that least squares measures vertical errors (squared distances up or down from data points to the closest line) while PCA measures perpendicular distances to the line. They are different problems. Ordinary least squares is different from PCA $=$ perpendicular least squares.

$A^{\mathrm{T}}A\widehat{\boldsymbol{x}} = A^{\mathrm{T}}b$ is $\begin{bmatrix} 3 & 0 \\ 0 & 14 \end{bmatrix} \widehat{\boldsymbol{x}} = \begin{bmatrix} 0 \\ 5 \end{bmatrix}$ leads to $\widehat{\boldsymbol{x}} = \begin{bmatrix} 0 \\ 5/14 \end{bmatrix}$. Best line is $y = \dfrac{5}{14}t$.

PCA finds the line through $(0, 0)$ whose perpendicular distances to the points $(-3, -1)$, $(1, 0), (2, 1)$ is smallest. The computation finds the top eigenvector of $A^{\mathrm{T}}A$, where A is now the 2 by 3 matrix of data points:

$$AA^{\mathrm{T}} = \begin{bmatrix} -3 & 1 & 2 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} -3 & -1 \\ 1 & 0 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 14 & 5 \\ 5 & 2 \end{bmatrix} \text{ has } \lambda^2 - 16\lambda + 3 = 0.$$

Then $\lambda = 8 \pm \sqrt{61}$ and the top eigenvector of AA^{T} is in the direction of $(5, \sqrt{61} - 6)$ $\approx (\mathbf{5}, \mathbf{1.8})$. That is the (approximate) direction of the line $y = \dfrac{1.8}{5}t$.

6 See **eigenfaces** on Wikipedia.

7 The closest matrix A_3 of rank 3 has the 3 top singular values $5, 4, 3$. Then $A - A_3$ has singular values $\mathbf{2}$ and $\mathbf{1}$.

8 If A has $\sigma_1 = 9$ and B has $\sigma_1 = 4$, then $A + B$ has $\sigma_1 \le 13$ because $||A + B|| \le ||A|| + ||B||$. Also $\sigma_1 \ge 5$ for $A + B$ because $||A + B|| + || - B|| \ge ||A||$.

Problem Set 8.1, page 315

1 With $w = 0$ linearity gives $T(v + 0) = T(v) + T(0)$. Thus $T(0) = 0$. With $c = -1$ linearity gives $T(-0) = -T(0)$. But $-0 = 0$. This is a second proof that $T(0) = 0$.

2 Combining $T(cv) = cT(v)$ and $T(dw) = dT(w)$ with addition gives $T(cv + dw) = cT(v) + dT(w)$. Then one more addition gives $cT(v) + dT(w) + eT(u)$.

3 (d) $T(v) = (0, 1) = $ constant and $(f)\, T(v) = v_1 v_2$ are not linear.

4 (a) $S(T(v)) = v$ (b) $S(T(v_1) + T(v_2)) = S(T(v_1)) + S(T(v_2))$: **linear**.

5 Choose $v = (1, 1)$ and $w = (-1, 0)$. Then $T(v) + T(w) = (v + w) = (0, 1)$ but $T(v + w) = T(2, 1)$ was defined as $(0, 0)$.

6 (a) $T(v) = v / \|v\|$ does not satisfy $T(v + w) = T(v) + T(w)$ or $T(cv) = cT(v)$

    (b) and (c) are linear   (d) satisfies $T(cv) = cT(v)$ only for $c \geq 0$

7 (a) $T(T(v)) = v$ (b) Nonlinear $v + (2, 2)$ (c) $T(T(v)) = -v$ (d) $T(T(v)) = T(v)$.

8 (a) The range of $T(v_1, v_2) = (v_1 - v_2, 0)$ is the line of vectors $(c, 0)$. The nullspace is the line of vectors (c, c).   (b) $T(v_1, v_2, v_3) = (v_1, v_2)$ has range $\mathbf{R}^2$, kernel $\{(0, 0, v_3)\}$ (c) $T(v) = 0$ has range $\{0\}$, kernel $\mathbf{R}^2$ (d) $T(v_1, v_2) = (v_1, v_1)$ has range = multiples of $(1, 1)$, kernel = multiples of $(1, -1)$.

9 If $T(v_1, v_2, v_3) = (v_2, v_3, v_1)$ then $T(T(v)) = (v_3, v_1, v_2)$; $T^3(v) = v$; $T^{100}(v) = T(v)$.

10 $T(v) = (4, 4); (2, 2); (2, 2)$; if $v = (a, b) = b(1, 1) + \frac{a-b}{2}(2, 0)$ then $T(v) = b(2, 2) + (0, 0)$.

11 (a) $T(1, 0) = 0$ (b) $(0, 0, 1)$ is not in the range   (c) $T(0, 1) = 0$.

12 For multiplication $T(v) = Av$: $V = \mathbf{R}^n$, $W = \mathbf{R}^m$; the outputs fill the column space; v is in the kernel if $Av = 0$.

13 The *distributive law* (page 69) gives $A(M_1 + M_2) = AM_1 + AM_2$. The *distributive law* over c's gives $A(cM) = c(AM)$.

14 Now $T(M) = AM$ with an invertible A. Multiply $AM = 0$ and $AM = B$ by A^{-1} to get $M = 0$ and $M = A^{-1}B$. The kernel contains only the zero matrix $M = 0$.

15 This A is *not* invertible. $AM = I$ is impossible. $A \begin{bmatrix} 2 & 2 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. The

range of T contains only matrices AM whose columns are multiples of $(1, 3)$.

16 No matrix A gives $A \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. To professors: Linear transformations on

matrix space come from **4 by 4 matrices**. Those in Problems 13–15 were special.

17 For $T(M) = M^{\mathrm{T}}$ (a) $T^2 = I$ is True   (b) True   (c) True   (d) False.

18 $T(I) = 0$ but $M = \begin{bmatrix} 0 & b \\ 0 & 0 \end{bmatrix} = T(M)$; these M's fill the range. Every $M = \begin{bmatrix} a & 0 \\ c & d \end{bmatrix}$

is in the kernel. Notice that $\dim$ (range) $+ \dim$ (kernel) $= 3 + 1 = \dim$ (input space

of 2 by 2 M's).

19 Linear transformations keep straight lines straight! And two parallel edges of a square

(edges differing by a fixed v) go to two parallel edges (edges differing by $T(v)$). So

the output is a parallelogram.

20 (a) Horizontal lines stay horizontal, vertical lines stay vertical   (b) House squashes

onto a line   (c) Vertical lines stay vertical because $T(1, 0) = (a_{11}, 0)$.

21 $D = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$ doubles the width of the house. $A = \begin{bmatrix} .7 & .7 \\ .3 & .3 \end{bmatrix}$ *projects* the house

(since $A^2 = A$ from trace $= 1$ and $\lambda = 0, 1$). The (non-orthogonal) projection is onto

the column space of $A =$ line through $(.7, .3)$. $U = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ will *shear* the house

horizontally: The point at (x, y) moves over to $(x + y, y)$.

22 (a) $A = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$ with $d > 0$ leaves the house AH sitting straight up

(b) $A = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$ rotates the house.

23 $T(v) = -v$ rotates the house by $180°$ around the origin. Then the affine transformation

$T(v) = -v + (1, 0)$ shifts the rotated house one unit to the right.

24 A code to add a chimney will be gratefully received!

25 This code needs a correction: add spaces between $-10 \quad 10 \quad -10 \quad 10$

26 $\begin{bmatrix} 1 & 0 \\ 0 & .1 \end{bmatrix}$ compresses vertical distances by 10 to 1. $\begin{bmatrix} .5 & .5 \\ .5 & .5 \end{bmatrix}$ projects onto the $45°$ line.

$\begin{bmatrix} .5 & .5 \\ -.5 & .5 \end{bmatrix}$ rotates by $45°$ clockwise and contracts by a factor of $\sqrt{2}$ (the columns have

length $1/\sqrt{2}$). $\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ has determinant -1 so the house is "flipped and sheared." One

way to see this is to factor the matrix as LDL^{T}:

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & \\ & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \text{(shear) (flip left-right) (shear)}.$$

27 Linear transformations of $\mathbf{R}^2$ take circles to ellipses (see figure in Section 6.7).

28 (a) $ad - bc = 0$ \qquad (b) $ad - bc > 0$ \qquad (c) $|ad - bc| = 1$. \qquad If vectors to two corners transform to themselves then by linearity $T = I$. (This is not always true if one corner is $(0, 0)$.)

Problem Set 8.2, page 324

1 For $Sv = d^2v/dx^2$

Basis $v_1, v_2, v_3, v_4 = 1, x, x^2, x^3$ The matrix for S is $B = \begin{bmatrix} 0 & 0 & \mathbf{2} & 0 \\ 0 & 0 & 0 & \mathbf{6} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

$Sv_1 = Sv_2 = \mathbf{0}$, $Sv_3 = 2v_1$, $Sv_4 = 6v_2$;

2 $Sv = d^2v/dx^2 = 0$ for linear functions $v(x) = a + bx$. All $(a, b, 0, 0)$ are in the nullspace of the second derivative matrix B.

3 (Matrix $A)^2 = B$ when transformation $T(T(v)) = S(v)$ and output basis = input basis.

4 The third derivative matrix has **6** in the $(1, 4)$ position; since the third derivative of x^3 is 6. This matrix also comes from AB. The fourth derivative of a cubic is zero, and B^2 is the zero matrix.

5 $T(v_1 + v_2 + v_3) = 2w_1 + w_2 + 2w_3$; A times $(1, 1, 1)$ gives $(2, 1, 2)$.

6 $v = c(v_2 - v_3)$ gives $T(v) = \mathbf{0}$; nullspace is $(0, c, -c)$; solutions $(1, 0, 0) + (0, c, -c)$.

7 $(1, 0, 0)$ is not in the column space of the matrix A, and w_1 is not in the range of the linear transformation T. Key point: *Column space* of matrix matches *range* of transformation. Nullspace matches normal.

8 We don't know $T(w)$ unless the w's are the same as the v's. In that case the matrix is A^2.

9 Rank of $A = 2 =$ dimension of the *range* of T. The outputs Av (column space) match the outputs $T(v)$ (the range of T). The "output space" W is like $\mathbf{R}^m$: it contains all outputs but may not be filled up by the column space.

10 The matrix for T is $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$. For the output $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ choose input $v = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} =$

$A^{-1} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. This means: For the output w_1 choose the input $v_1 - v_2$.

11 $A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$ so $T^{-1}(\boldsymbol{w}_1) = \boldsymbol{v}_1 - \boldsymbol{v}_2, T^{-1}(\boldsymbol{w}_2) = \boldsymbol{v}_2 - \boldsymbol{v}_3, T^{-1}(\boldsymbol{w}_3) =$

$\boldsymbol{v}_3$. The columns of A^{-1} describe T^{-1} from $\boldsymbol{W}$ back to $\boldsymbol{V}$. The only solution to $T(\boldsymbol{v}) = 0$ is $\boldsymbol{v} = 0$.

12 (c) $T^{-1}(T(\boldsymbol{w}_1)) = \boldsymbol{w}_1$ is wrong because $\boldsymbol{w}_1$ is not generally in the input space.

13 (a) $T(\boldsymbol{v}_1) = \boldsymbol{v}_2, T(\boldsymbol{v}_2) = \boldsymbol{v}_1$ is its own inverse   (b) $T(\boldsymbol{v}_1) = \boldsymbol{v}_1, T(\boldsymbol{v}_2) = 0$ has $T^2 = T$ (c) If $T^2 = I$ for part (a) and $T^2 = T$ for part (b), then T must be I.

14 (a) $\begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix} =$ inverse of (a)   (c) $A \begin{bmatrix} 2 \\ 6 \end{bmatrix}$ must be $2A \begin{bmatrix} 1 \\ 3 \end{bmatrix}$.

15 (a) $M = \begin{bmatrix} r & s \\ t & u \end{bmatrix}$ transforms $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ to $\begin{bmatrix} r \\ t \end{bmatrix}$ and $\begin{bmatrix} s \\ u \end{bmatrix}$; this is the "easy"

direction.   (b) $N = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1}$ transforms in the inverse direction, back to the standard basis vectors.   (c) $ad = bc$ will make the forward matrix singular and the inverse impossible.

16 $MW = \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix}^{-1} = \begin{bmatrix} 3 & -1 \\ -7 & 3 \end{bmatrix}$.

17 Reordering basis vectors is done by a *permutation matrix*. Changing lengths is done by a *positive diagonal matrix*.

18 $(a, b) = (\cos \theta, -\sin \theta)$. Minus sign from $Q^{-1} = Q^{\mathrm{T}}$.

19 $M = \begin{bmatrix} 1 & 1 \\ 4 & 5 \end{bmatrix}$; $\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 5 \\ -4 \end{bmatrix} =$ first column of $M^{-1} =$ coordinates of $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ in basis $\begin{bmatrix} 1 \\ 4 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \end{bmatrix}$ because $5 \begin{bmatrix} 1 \\ 4 \end{bmatrix} - 4 \begin{bmatrix} 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

20 $\boldsymbol{w}_2(x) = 1 - x^2$; $\boldsymbol{w}_3(x) = \frac{1}{2}(x^2 - x)$; $\boldsymbol{y} = 4\boldsymbol{w}_1 + 5\boldsymbol{w}_2 + 6\boldsymbol{w}_3$.

21 w's to v's: $\begin{bmatrix} 0 & 1 & 0 \\ .5 & 0 & -.5 \\ .5 & -1 & .5 \end{bmatrix}$. v's to w's: inverse matrix $= \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{bmatrix}$. *The key*

idea: The matrix multiplies the coordinates in the v basis to give the coordinates in the

w basis.

22 The 3 equations to match $4, 5, 6$ at $x = a, b, c$ are $\begin{bmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$. This

Vandermonde determinant equals $(b - a)(c - a)(c - b)$. So a, b, c must be distinct to

have $\det \neq 0$ and one solution A, B, C.

23 The matrix M with these nine entries must be invertible.

24 Start from $A = QR$. Column 2 is $a_2 = r_{12}q_1 + r_{22}q_2$. This gives a_2 as a combination

of the q's. So the change of basis matrix is R.

25 Start from $A = LU$. Row 2 of A is ℓ_{21}(row 1 of U) $+ \ell_{22}$ (row 2 of U). The change of

basis matrix is always *invertible*, because basis goes to basis.

26 The matrix for $T(v_i) = \lambda_i v_i$ is $\Lambda = \mathsf{diag}(\lambda_1, \lambda_2, \lambda_3)$.

27 If T is not invertible, $T(v_1), \ldots, T(v_n)$ is not a basis. We couldn't choose $w_i = T(v_i)$.

28 (a) $\begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix}$ gives $T(v_1) = \mathbf{0}$ and $T(v_2) = 3v_1$. (b) $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ gives $T(v_1) = v_1$

and $T(v_1 + v_2) = v_1$ (which combine into $T(v_2) = \mathbf{0}$ by *linearity*).

29 $T(x, y) = (x, -y)$ is reflection across the x-axis. Then reflect across the y-axis to get

$S(x, -y) = (-x, -y)$. Thus $ST = -I$.

30 S takes (x, y) to $(-x, y)$. $S(T(v)) = (-\mathbf{1}, \mathbf{2})$. $S(v) = (-2, 1)$ and $T(S(v)) = (\mathbf{1}, -\mathbf{2})$.

31 Multiply the two reflections to get $\begin{bmatrix} \cos 2(\theta - \alpha) & -\sin 2(\theta - \alpha) \\ \sin 2(\theta - \alpha) & \cos 2(\theta - \alpha) \end{bmatrix}$ which is *rotation*

by $2(\theta - \alpha)$. In words: $(1, 0)$ is reflected to have angle 2α, and that is reflected again

to angle $2\theta - 2\alpha$.

32 The matrix for T in this basis is $A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

33 The basis v_1, v_2, v_3, v_4 for the space of 2 by 2 matrices is meant to be

$$v_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \qquad v_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad v_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \qquad v_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Apologies for the reference to Worked Example 8.2A, which was omitted in final editing of the book. Question 33 asks for the 4 by 4 matrix A that represents $T(M) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} M$ in this basis of v's. As always, multiply each basis matrix by $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ to find $T(v_1), \ldots, T(v_4)$. Write each output as a conbination of the output basis (also v_1 to v_4). The coefficients of v_1 to v_4 tell you each column of A.

Multiplying by $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ gives $T(v_1) = A \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} a & 0 \\ c & 0 \end{bmatrix} = av_1 + cv_3$. Similarly $T(v_2) = av_2 + cv_4$ and $T(v_3) = bv_1 + dv_3$ and $T(v_4) = bv_2 + dv_4$. The matrix for T in this basis is $\begin{bmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ c & 0 & d & 0 \\ 0 & c & 0 & d \end{bmatrix}$

34 False: We will not know $T(v)$ for *every* v unless the n v's are linearly independent.

Problem Set 8.3, page 334

1 For this matrix J, the rank of $J - 3I$ is **3** so the dimension of the nullspace is only 1. There is only 1 independent eigenvector even though $\lambda = 3$ is a *double root* of $\det(J - \lambda I) = 0$: a repeated eigenvalue.

$$J = \begin{bmatrix} 2 & & & \\ & 2 & & \\ & & 3 & 1 \\ & & & 3 \end{bmatrix}.$$

2 $J = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ is similar to all other 2 by 2 matrices A that have **2** zero eigenvalues but only **1** independent eigenvector. Then $J = B_1^{-1} A_1 B_1$ is the same as $B_1 J = A_1 B_1$:

$$B_1 J = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 4 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix} = A_1 B_1$$

$$B_2 J = \begin{bmatrix} 4 & 1 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 4 & -8 \\ 2 & -4 \end{bmatrix} \begin{bmatrix} 4 & 1 \\ 2 & 0 \end{bmatrix} = A_2 B_2$$

3 Every matrix is similar to its transpose (same eigenvalues, same multiplicity, more than that the same Jordan form). In this example

$$BJ = \begin{bmatrix} & & 1 \\ & 1 & \\ 1 & & \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} & & 1 \\ & 1 & \\ 1 & & \end{bmatrix} = J^{\mathrm{T}} B.$$

4 Here J and K are *different* Jordan forms (block sizes $2, 2$ versus block sizes $3, 1$). Even though J and K have the same λ's (all zero) and same rank, J and K are *not similar*. If $BK = JB$ then B is *not invertible* :

$$BK = B \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & b_{11} & b_{12} & 0 \\ 0 & b_{21} & b_{22} & 0 \\ 0 & b_{31} & b_{32} & 0 \\ 0 & b_{41} & b_{42} & 0 \end{bmatrix}$$

$$JB = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} B = \begin{bmatrix} b_{21} & b_{22} & b_{23} & b_{24} \\ 0 & 0 & 0 & 0 \\ b_{41} & b_{42} & b_{43} & b_{44} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Those right hand sides agree only if $b_{21} = 0, b_{41} = 0, b_{24} = 0, b_{44} = 0, b_{22} = 0,$ $b_{42} = 0$. But then also $b_{11} = b_{22} = 0$ and $b_{31} = b_{42} = 0$. So the first column has $b_{11} = b_{21} = b_{31} = b_{41} = 0$ and B is not invertible.

5 If A^3 is the zero matrix then every eigenvalue of A is $\lambda = 0$ (because $Ax = \lambda x$ leads to $\boldsymbol{\theta} = A^3 x = \lambda^3 x$). The Jordan form J will also have $J^3 = 0$ because $J = B^{-1}AB$ has $J^3 = B^{-1}A^3B = 0$. The blocks of J must become zero blocks in J^3. So those blocks of J can be

$$\begin{bmatrix} 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{but not} \quad \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \left(\begin{matrix} \text{third power} \\ \text{is not zero} \end{matrix} \right)$$

The rank of J (and A) is largest if every block is 3 by 3 of rank 2. Then rank $\leq \dfrac{2}{3}n$.

If $A^n =$ zero matrix then A is *not invertible* and rank $(A) < n$.

6 This question substitutes $u_1 = te^{\lambda t}$ and $u_2 = e^{\lambda t}$ to show that u_1, u_2 solve the system $\boldsymbol{u}' = J\boldsymbol{u}$:

$$\begin{aligned} u_1' &= \lambda u_1 + u_2 & e^{\lambda t} + t\lambda e^{\lambda t} &= \lambda(te^{\lambda t}) + (e^{\lambda t}) \\ u_2' &= \qquad \lambda u_2 & \lambda e^{\lambda t} &= \qquad \lambda(e^{\lambda t}). \end{aligned}$$

Certainly $u_1 = 0$ and $u_2 = 1$ at $t = 0$, so we have the solution and it involves $te^{\lambda t}$ (the factor t appears because λ is a double eigenvalue of J).

7 The equation $u_{k+2} - 2\lambda u_{k+1} + \lambda^2 u_k$ is certainly solved by $u_k = \lambda^k$. But this is a second order equation and there must be another solution. In analogy with $te^{\lambda t}$ for the differential equation in 8.3.6, that second solution is $u_k = k\lambda^k$. Check:

$$(k+2)\lambda^{k+2} - 2\lambda(k+1)\lambda^{k+1} + \lambda^2(k)\lambda^k = \big[k + 2 - 2(k+1) + k\big]\lambda^{k+2} = 0.$$

8 $\lambda^3 = 1$ has 3 roots $\lambda = 1$ and $e^{2\pi i/3}$ and $e^{4\pi i/3}$. Those are $\mathbf{1}, \boldsymbol{\lambda}, \boldsymbol{\lambda^2}$ if we take $\lambda = e^{2\pi i/3}$. The Fourier matrix is

$$F_3 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & \lambda & \lambda^2 \\ 1 & \lambda^2 & \lambda^4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & e^{2\pi i/3} & e^{4\pi i/3} \\ 1 & e^{4\pi i/3} & e^{8\pi i/3} \end{bmatrix}.$$

9 A 3 by 3 circulant matrix has the form on page 425:

$$C = \begin{bmatrix} c_0 & c_1 & c_2 \\ c_2 & c_0 & c_1 \\ c_1 & c_2 & c_0 \end{bmatrix} \text{ with } C \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = (c_0 + c_1 + c_2) \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$C \begin{bmatrix} 1 \\ \lambda \\ \lambda^2 \end{bmatrix} = (c_0 + c_1\lambda + c_2\lambda^2) \begin{bmatrix} 1 \\ \lambda \\ \lambda^2 \end{bmatrix} \qquad C \begin{bmatrix} 1 \\ \lambda^2 \\ \lambda^4 \end{bmatrix} = (c_0 + c_1\lambda^2 + c_2\lambda^4) \begin{bmatrix} 1 \\ \lambda^2 \\ \lambda^4 \end{bmatrix}.$$

Those 3 eigenvalues of C are exactly the 3 components of $F\mathbf{c} = F \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix}$,

10 The Fourier cosine coefficient c_3 is in formula (7) with integrals from $-\pi$ to π. Because f drops to zero at $x = L$, the integral stops at L:

$$a_3 = \frac{\int f(x)\cos 3x \, dx}{\int (\cos 3x)^2 \, dx} = \frac{1}{\pi} \int_{-L}^{L} (1)(\cos 3x) \, dx = \frac{1}{3\pi} \Big[\sin 3x\Big]_{x=-L}^{x=L} = \frac{2\sin 3L}{3\pi}.$$

Note that we should have defined $f(x) = 0$ for $L < |x| < \pi$ (not 2π !).

Problem Set 9.1, page 345

1 Equations (1) and (2) give the first three terms in the Taylor Series for $F(x)$ and $F(\boldsymbol{x}) =$ one variable x or several variables $\boldsymbol{x} = (x_1, \ldots, x_n)$. The formulas are exact when F is a **quadratic function** (powers $1, x, x^2$ or in N dimensions $1, x_1$ to x_N and all products $x_i x_j$ from x_j^2 to $x_1 x_N$ to x_2^2 to $x_2 x_N$ and eventually $x_{N-1} x_N$ and x_N^2).

2 $f(x) = x^2 + 1, f'(x) = 2x, f'' = 2$. Then Newton's method computes x_{n+1} from x_n:

$$x_{n+1} = x_n - \frac{f(x_n)}{f^v(x_n)} = x_n - \frac{x_n^2 + 1}{2x_n} = \frac{1}{2}\left(x_n - \frac{1}{x_n}\right).$$

This iteration stays with real numbers so it can't converge to the solutions $x = i$ or $x = -i$. Instead the numbers x_n are **"chaotic"**.

The key insight is that the cotangent function $x_n = \cot 2^n \theta$ gives the correct x_{n+1} from x_n:

$$\frac{\cos 2\theta}{\sin 2\theta} = \frac{1}{2}\left(\frac{\cos\theta}{\sin\theta} - \frac{\sin\theta}{\cos\theta}\right) \quad \text{or} \quad \cot 2\theta = \frac{1}{2}\left(\cot\theta - \frac{1}{\cot\theta}\right).$$

In the left equation, the common denominator is $2\sin\theta\cos\theta = \sin 2\theta$. The numerator is $\cos^2\theta - \sin^2\theta = \cos 2\theta$. The identity says this about the iteration:

$$\text{If } x_0 = \cot\theta \text{ then } x_1 = \frac{1}{2}\left(x_0 - \frac{1}{x_0}\right) = \cot 2\theta.$$

Then $x_2 = \cot 4\theta$. Then $x_n = \cot 2^n \theta$. **This is the formula**.

Example 1 Start with $\theta = \pi/4$ (cotangent is $x_0 = 1$). The first step gives $\theta = \pi/2$ (cotangent equals 0). The next step is $\theta = \pi$ (iteration blows up because $\sin\pi = 0$).

Example 2 Start with $\theta = \pi/3$ (cotangent is $x_0 = 1/\sqrt{3}$). The first step gives $\theta = 2\pi/3$ (cotangent equals $-1/\sqrt{3}$). The next step is $\theta = 4\pi/3$ (which is the original θ plus π). The iteration cycles between $1/\sqrt{3}$ and $-1/\sqrt{3}$.

Example 3 Start with a small θ (a large cotangent). After the first step, $\cot 2\theta$ is approximately cut in half (use calculus). The cotangent decreases until the angle $2^n\theta$ passes $\pi/3$. Then the next step makes it larger.

The iteration eventually blows up if $\theta/\pi = \text{integer}/2^N$.

The iteration eventually cycles if $\theta/\pi = $ any other fraction p/q.

The iteration is not periodic (or convergent) if θ/π is irrational. **Chaos**.

We can't find $\sqrt{-1}$ by Newton's method which stays real.

3 The derivative of $-\log x_{11}$ with respect to x_{11} is $-1/x_{11}$. **The second derivative is** $+1/x_{11}^2 > 0$. So $-\log x$ is a convex function—its slope $-1/x$ is increasing.

4 For the function $F = -\log \det A = -\log(ac - b^2)$ the first derivatives are

$$\frac{\partial F}{\partial a} = \frac{-c}{ac - b^2} \qquad \frac{\partial F}{\partial b} = \frac{+2b}{ac - b^2} \qquad \frac{\partial F}{\partial c} = \frac{-a}{ac - b^2}$$

Then all second derivatives involve $D = ac - b^2 =$

$$\frac{\partial^2 F}{\partial a^2} = \frac{c^2}{D^2} \qquad \frac{\partial}{\partial b}\left(\frac{\partial F}{\partial a}\right) = \frac{+2bc}{D^2} \qquad \frac{\partial}{\partial c}\left(\frac{\partial F}{\partial a}\right) = \frac{-1}{D} + \frac{ac}{D^2}.$$

This already shows that the matrix of second derivatives is complicated and we need a new idea.

5 The gradient of $F(\boldsymbol{x}) = \dfrac{1}{2}\left(\boldsymbol{x}^{\mathrm{T}} A^{\mathrm{T}} A \boldsymbol{x} - 2\boldsymbol{b}^{\mathrm{T}} A \boldsymbol{x} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{b}\right)$ is $\boldsymbol{\nabla} F = A^{\mathrm{T}} A \boldsymbol{x} - A^{\mathrm{T}}\boldsymbol{b}$. So the minimizing vector $\overline{\boldsymbol{x}}$ solves $A^{\mathrm{T}} A \overline{\boldsymbol{x}} = A^{\mathrm{T}}\boldsymbol{b}$ as we know from Chapter 4. That gives $\overline{\boldsymbol{x}}$ in one step, where gradient descent takes a sequence of simpler steps of sizes s_k to approach $\overline{\boldsymbol{x}}$:

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - s_k \boldsymbol{\nabla} F(\boldsymbol{x}_k) = \boldsymbol{x}_k - s_k(A^{\mathrm{T}} A \boldsymbol{x}_k - A^{\mathrm{T}}\boldsymbol{b}).$$

6 $F = \dfrac{1}{2}\left(x^2 + \dfrac{1}{4}y^2\right)$ has gradient $\left(\dfrac{\partial F}{\partial x}, \dfrac{\partial F}{\partial y}\right) = \left(x, \dfrac{y}{4}\right)$. One step from $(x_0, y_0) = \left(\dfrac{1}{4}, 1\right)$ goes to $(x_1, y_1) = \left(\dfrac{1}{4}, 1\right) - s\left(\dfrac{1}{4}, \dfrac{1}{4}\right)$.

7

8 Certainly x^2 is minimized at $x = 0$ and e^{-y} is minimized at $y = \infty$. Then $F_{\min} = 0 + 0$. At $(x_0, y_0) = (1, 1)$ the function is $F = \dfrac{1}{2} + \dfrac{1}{e}$ and the gradient is $\nabla F = (2x, -e^{-y})$ and descent goes to $(x_1, y_1) = (1, 1) = -s\nabla F = (1, 1) - s(2, -1/e)$.

Problem Set 9.2, page 353

1 The two equations have rows $a_1 = \begin{bmatrix} 1 & 2 \end{bmatrix}$ and $a_2 = \begin{bmatrix} 2 & 3 \end{bmatrix}$. The right hand sides of $Ax = b$ are $b_1 = 3$ and $b_3 = 5$. The classical Kaczmarz method updates x to solve $a_1 x = b_1$ and then $a_2 x = b_2$ (and repeat indefinitely). Starting from $x_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ here is the first cycle from equation (6): x_1 solves the first equation $\begin{bmatrix} 1 & 2 \end{bmatrix} x_1 = 3$

$$x_1 = x_0 + \frac{3 - \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix}}{1^2 + 2^2} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \frac{3}{5} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3/5 \\ 6/5 \end{bmatrix}$$

Note that x_1 solves the first equation $\begin{bmatrix} 1 & 2 \end{bmatrix} x_1 = 3$ but not $a_2 x_1 = 5$.

Then x_2 modifies x_1 to solve that equation but now the first equation fails:

$$x = \begin{bmatrix} 3/5 \\ 6/5 \end{bmatrix} + \frac{5 - \begin{bmatrix} 2 & 3 \end{bmatrix} \begin{bmatrix} 3/5 \\ 6/5 \end{bmatrix}}{2^2 + 3^2} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 3/5 \\ 6/5 \end{bmatrix} + \frac{1/5}{13} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 3 + 2/13 \\ 6 + 3/13 \end{bmatrix}.$$

This vector x_2 solves the second equation $\begin{bmatrix} 2 & 3 \end{bmatrix} x_2 = 5$ but not the first equation.

Repeating the Kaczmarz double step brings us closer to the true solution $x_\infty = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

With a computer you can take more Kaczmarz steps to $x_2, x_3, \ldots$ and compare the convergence rate with **"random Kaczmarz"**–when the order of the equations and updates jumps randomly between equations 1 and 2.

2 The problem (least squares) is to minimize $\ell(x) = (1-F(x))^2$ where $F(x) = F_2(F_1(x))$.

Backpropagation computes $\dfrac{\partial \ell}{\partial F}$ and $\dfrac{\partial F}{\partial \boldsymbol{x}}$:

$$\frac{\partial \ell}{\partial F} = 2(1-F(x)) \qquad \frac{\partial F}{\partial x} = \frac{\partial F_2}{\partial F_1}\frac{\partial F_1}{\partial x} = \left(\frac{\partial F_2}{\partial F_1}\right)(-\sin(\sin x))\cos x = 0 \ \text{ at } \ x = 0$$

3 (a) The limits of $\tanh(x)$ at $x = 0, \infty, -\infty$ are $0, 1, -1$.

(b) The graph has an inflection point at $x = 0$, where the second derivative of $\tanh(x)$

is zero.

(c) The derivative of $\tanh(x) = (e^x - e^{-x})/(e^x + e^{-x})$ by the quotient rule is

$$\frac{(e^x + e^{-x})^2 - (e^x - e^{-x})^2}{(e^x + e^{-x})^2} = \frac{4}{(e^x + e^{-x})^2} > 0 \quad \text{so } \tanh \text{ is increasing}$$

4 Component i of $\tanh(A\boldsymbol{x} + \boldsymbol{b})$ is $y = \tanh(\boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{x} + b_i)$. The scalar function $\tanh(x)$

has the derivative $4/(e^x + e^{-x})^2$ from Problem 3. Then the chain rule gives

$$\frac{\partial}{\partial b_i}(\tanh(A\boldsymbol{x} + \boldsymbol{b})) = \frac{4}{(e^x + e^{-x})^2} \quad \text{with} \ \ x = \boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{x} + b_i.$$

5 The partial derivatives of $F(x,y) = F_2(x, F_1(y))$ are $\dfrac{\partial F}{\partial x} = \dfrac{\partial F_2}{\partial x}(x_1, F_1(y))$ and $\dfrac{\partial F}{\partial y} = \dfrac{\partial F_2}{\partial F_1}\dfrac{\partial F_1}{\partial y}$.

Problem Set 9.3, page 363

1 To minimize with a constraint $3x + 4y = 1$, introduce a Lagrange multiplier in $L(x, y, \lambda) = |x|^p + |y|^p - \lambda(3x + 4y - 1)$.

$$\partial L/\partial x = px^{p-1} - 3\lambda = 0 \qquad\qquad x^{p-1} = 3\lambda/p$$

$$\partial L/\partial y = py^{p-1} - 4\lambda = 0 \qquad\qquad y^{p-1} = 4\lambda/p$$

$$-\partial L/\partial\lambda = 3x + 4y - 1 = 0 \qquad \mathbf{3(3\lambda/p)^q + 4(4\lambda/p)^q = \ 1}$$

Here $q = 1/(p-1)$ and $q+1 = p/(p-1)$. The bold equation is $(3^{q+1} + 4^{q+1})\lambda^q = p^q$. The solution is $\lambda = p/E$ with $E = (3^{q+1} + 4^{q+1})^{1/q}$. Knowing λ we can find x and y.

2 Suppose v_1 and v_2 are in the unit ball : their norms are ≤ 1. We want to show that any point $w = cv_1 + (1 - c)v_2$ between v_1 and v_2 (which means $0 < c < 1$) is also in the ball (which means $||w|| \leq 1$). Use the triangle inequality :

$$||w|| \leq ||cv_1|| + ||(1 - c)v_2|| \leq |c| + |1 - c| = 1 \ \text{ because } \ 0 < c < 1.$$

3 (a) $L(\mathbf{X}, \lambda) = \dfrac{1}{2}x_1^2 + 2x_2^2 - \lambda(x_1 + 3x_2 - b)$.

(b) $\partial L/\partial x_1 = \qquad x_1 - \lambda = 0$

$\quad\partial L/\partial x_2 = \qquad 4x_2 - 3\lambda = 0$

$\quad\partial L/\partial\lambda = x_1 + 3x_2 - b = 0 = (\lambda) + 3(3\lambda/4) - b$

Solve the last equation : $(4/4 + 9/4)\lambda = b$ or $\lambda = \mathbf{4b/13}$. Then $x_1 = 4b/13$ and $x_2 = 3b/13$.

(d) The minimum of $F = \dfrac{1}{2}x_1^2 + 2x_2^2$ is $\dfrac{1}{2}\left(\dfrac{4b}{13}\right)^2 + 2\left(\dfrac{3b}{13}\right)^2 = (8 + 18)b^2/13^2 = 2b^2/13$. The derivative of that $F_{\min}$ is $\partial F_{\min}/\partial b = 4b/13$. This is exactly λ !

4 $L = \dfrac{1}{2}(x_1^2 + 4x_2^2) - \lambda(2x_1 + x_2 - 5)$.

$\quad\partial L/\partial x_1 = \qquad x_1 - 2\lambda = 0 \qquad\qquad x_1 = 2\lambda$

$\quad\partial L/\partial x_2 = \qquad 4x_2 - \lambda = 0 \qquad\qquad x_2 = \lambda/4$

$\quad-\partial L/\partial\lambda = 2x_1 + x_2 - 5 = 0 \qquad \left(4 + \dfrac{1}{4}\right)\lambda = 5 \ \text{ or } \ \lambda = 20/17$

Then $x_1 = 40/17$ and $x_2 = 5/17$ and $F = \dfrac{1}{2}\left(x_1^2 + 4x_2^2\right) = \dfrac{1}{2}(1700)/17^2 = \mathbf{50/17}$.

5 K has 2 positive pivots and 1 negative pivot. Then it must have 2 positive eigenvalues and 1 negative eigenvalue.

6 Line 4 of this problem proves the Law of Inertia for the number of positive pivots of S (= number of positive eigenvalues, because K is symmetric). Then the problem asks for a straightforward and basic proof when S is 2 by 2 : There are 2 steps.

> **1** Determinants of $S = LDL^{\mathrm{T}} = Q\Lambda Q^{\mathrm{T}}$ are $(\det D)(\det L)^2$ and $(\det \Lambda)(\det Q)^2$. So $\det D$ has the same sign as $\det \Lambda$. If those diagonal matrices D and Λ are 2 by 2, then negative determinants mean 1 positive pivot and 1 positive eigenvalue.

> **2** If D and Λ both have positive determinants, then the two pivots have the same sign and the two eigenvalues have the same sign. Then positive eigenvalues means positive trace (sum of eigenvalues = sum of diagonal entries).

7 By symmetry the minimum of $F = \frac{1}{2}\left(x_1^2 + x_2^2 + x_3^2\right)$ with $x_1 + x_2 + x_3 = 3$ is $\frac{1}{2}(1 + 1 + 1) = 1.5$. With the added constraint $x_1 + 2x_2 + 3x_3 = 12$, we have two multipliers λ and σ :

Lagrange function $L = F - \lambda(x_1 + x_2 + x_3 - 3) - \sigma(x_1 + 2x_2 + 3x_3 - 12)$.

$$
\begin{array}{llll}
\partial L/\partial x_1 = & x_1 - \lambda - \sigma = 0 & & x_1 = -2 \\
\partial L/\partial x_2 = & x_2 - \lambda - 2\sigma = 0 & & x_2 = 1 \\
\partial L/\partial x_3 = & x_3 - \lambda - 3\sigma = 0 & & x_3 = 4 \\
-\partial L/\partial \lambda = & x_1 + x_2 + x_3 - 3 = 0 \;\;\rightarrow\;\; 3\lambda + 6\sigma = 3 \;\;\rightarrow\;\; \lambda = -5 \\
-\partial L/\partial \sigma = x_1 + 2x_2 + 3x_3 - 12 = 0 \;\;\rightarrow\;\; 6\lambda + 14\sigma = 12 \;\;\rightarrow\;\; \sigma = 3
\end{array}
$$

The second minimum is $F = \dfrac{1}{2}(4 + 1 + 16) = 10.5 = $ larger than 1.5 from extra constraint !

8 The constraints are $x_1 \le 1, -x_1 \le 1, x_2 \le 1, -x_2 \le 1$. This is $A\boldsymbol{x} \le \boldsymbol{b}$ for

$$
A = \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & -1 \end{bmatrix}
\qquad
\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}
\qquad
\boldsymbol{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}
$$

Problem Set 9.4, page 369

1 Yes, $x \geq 0$ is needed for this step: If $A^\mathrm{T} y \leq c$ then $(A^\mathrm{T} y)^\mathrm{T} x \leq c^\mathrm{T} x$.

$3 \leq 4$ leads to $3x \leq 4x$ if $x \geq 0$. But $x = -1$ would give a wrong output $-3 < -4$.

2 The corners (x_1, x_2, x_3) with $x_1 + 2x_2 + 2x_3 = 4$ are $(4, 0, 0)$ and $(0, 2, 0)$ and $(0, 0, 2)$. When the cost is $c^\mathrm{T} x = 5x_1 + 3x_3 + 8x_3$, the corner $(0, 2, 0)$ gives the minimum cost $c^\mathrm{T} x = \mathbf{6}$. The other corners $(4, 0, 0)$ and $(0, 0, 2)$ cost 20 and 16.

3 The dual problem maximizes $y^\mathrm{T} b = 4y_1$ subject to $y_1 \geq 5, 2y_1 \geq 3, 2y_1 \geq 8$. The best y_1 is $y_1 = 3/2$ leading to $y^\mathrm{T} b = 4(3/2) = \mathbf{6}$. Notice strong duality: min of $y^\mathrm{T} b$ equals max of $c^\mathrm{T} x$ in Problem $2 = \mathbf{6}$.

4 With 2 constraints on x_1 to x_4, we can set $4 - 2 = 2$ of the x's to zero and see if the other 2 x's are ≥ 0 as required. The problem asks about $(x_1, x_2, 0, 0) = (4, 2, 0, 0)$ which satisfies the constraints. The cost at the corner depends on the cost function $c^\mathrm{T} x$ which the problem statement forgot to include.

5 First payoff matrix: R chooses row 1 every time and C chooses column 2 every time. The payoff to C is 2 every time.

Second matrix: If R chooses rows 1 and 2 with probabilities x and $1 - x$, the payoffs to C are $x + 8(1 - x)$ for column 1 and $4x + 2(1 - x)$ for column 2. Those payoffs are equal if $8 - 7x = 2 + 2x$ or $6 = 9x$ or $x = 2/3$: payoff $= \mathbf{10/3}$.

If C chooses columns 1 and 2 with probabilities y and $1 - y$, the payoffs to C are $y + 4(1 - y) = 4 - 3y$ when R chooses row 1 and $8y + 2(1 - y) = 2 + 6y$ when R chooses row 2. Those are equal when $4 - 3y = 2 + 6y$ or $y = 2/9$ and the payoff to C is again $4 - 2/3 = \mathbf{10/3}$. Duality holds and the game is worth $10/3$ to player C.

6 If $A^{\mathrm{T}} = -A$ (square antisymmetric matrix) then any choice $\boldsymbol{x}$ of the row frequencies by the row player R can be matched by the same probabilities $\boldsymbol{x}$ for the column player C (and vice versa). But $\boldsymbol{x}^{\mathrm{T}}(A\boldsymbol{x}) = (A\boldsymbol{x})^{\mathrm{T}}\boldsymbol{x} = -\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ is always zero! So both players can avoid any loss and the value of this antisymmetric game is zero.

Example: $A = \begin{bmatrix} 0 & -\mathbf{1} \\ \mathbf{1} & 0 \end{bmatrix}$ has optimal strategies $\boldsymbol{x} = (1,0)$ for row player R and $y = (1,0)$ for column player C and all payoffs $= 0$.

7 Player R chooses row 1 with probability $\sigma_n/(\sigma_1 + \sigma_n)$ and row n with probability $\sigma_1/(\sigma_1 + \sigma_n)$. Player C chooses column 1 and column n with those same probabilities and expects the same average payments. (This is effectively a 2 by 2 diagonal matrix with the same mixed strategy for both players.) The average payment from R to C will be $\sigma_1\sigma_n/(\sigma_1 + \sigma_n)$.

8 $||(x_1, x_2, x_3)||_1 \le 2$ is equivalent to $\pm x_1 \pm x_2 \pm x_3 \le 2$. Those 3 choices of plus or minus sign give 8 linear inequalities.

9 We have not explained semidefinite programming well enough to make this a fair question.

10 If $A\boldsymbol{x} \le \boldsymbol{b}$ and $\boldsymbol{x} \ge \mathbf{0}$ and $A\boldsymbol{X} \le \boldsymbol{b}$ and $\boldsymbol{X} \ge \mathbf{0}$, then $\frac{1}{2}A(\boldsymbol{x} + \boldsymbol{X}) \le \boldsymbol{b}$ and $\frac{1}{2}(x + X) \ge 0$. Convexity is a crucial property in the theory of optimization.

Problem Set 10.1, page 372

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

Problem Set 10.2, page 381

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

Problem Set 10.3, page 386

1 If all sample values are $x = 20$, then the sample mean is $\mu = 20$ and the sample variance is $S^2 = 0$. If $x = 20$ or 21 (12 samples each) then $\mu = 20.5$ and

$$S^2 = \frac{1}{23}(24)(0.5)^2 = \frac{6}{23}.$$

2 The sample mean and the expected value both increase by 7. The variance does not change.

3 Fraction of integers divisible by 3 or 7 or both $= \frac{1}{3} + \frac{1}{7} - \frac{1}{21} = \frac{9}{21} = \frac{3}{7}$.

Since $\frac{1}{21}$ of the integers are divisible by **both 3 and 7**, those were double counted in $\frac{1}{3}$ and $\frac{1}{7}$. Then $\frac{1}{21}$ must be subtracted off.

4 The probabilities p_0 to p_9 that a random number from 1 to 10 (and 1 to 100 and 1 to 1000) ends in $0, 1, , \ldots, 9$ are $\frac{1}{10}$ for each case. The expected mean of that digit is 4.5. The expected variance is $\sigma^2 = \frac{2}{10}(0.5^2 + 1.5^2 + 2.5^2 + 3.5^2 + 4.5^2)$.

5 By Problem 4, the last digit is $0, 1, \ldots, 9$ with equal probability $p = \frac{1}{10}$. The squares of 0 to 9 end in $0, 1, 4, 9, 6, 5, 6, 9, 4, 1$. So the probabilities p_0 to p_9 for those squares to end in $0, 1, 4, 9, 6, 5$ are $\frac{1}{10}, \frac{2}{10}, \frac{2}{10}, \frac{2}{10}, \frac{2}{10}, \frac{1}{10}$. The mean of those ten ending digits is

$$0\left(\frac{1}{10}\right) + (1 + 4 + 9 + 6)\left(\frac{2}{10}\right) + 5\left(\frac{1}{10}\right) = \frac{45}{10} = 4.5.$$

The variance σ^2 is $\frac{1}{10}\left[4.5^2 + 2\left(3.5^2 + 0.5^2 + 4.5^2 + 1.5^2\right) + 0.5^2\right]$.

Crazy question.

6 The **first digit** of the numbers from 1 to 1000 is $\mathbf{1}$(112 times). It is 2 to 9   111 times each. Reason : The first digit is 2 to 9 in $1 + 10 + 100$ cases. First digit $= 1$ in 1 extra case. Total count is $8(111) + 1(112) = 1000$.

Mean : $m = \frac{1}{1000}(112 + 111(2 + 3 + \cdots + 9)) = \frac{1}{1000}[1 + 111(45)]$

Variance : $\sigma^2 = $ prize for this computation !

7 The squares of $157, 312, 696, 602$ start with $2, 9, 4, 3$. The sample mean μ (the average) is $18/4$. The sample variance is

$$S^2 = \frac{1}{3}\left[\left(2 - \frac{18}{4}\right)^2 + \left(9 - \frac{18}{4}\right)^2 + \left(4 - \frac{18}{4}\right)^2 + \left(3 - \frac{18}{4}\right)^2\right].$$

8 $\sigma^2 = \sum p_i(x_i - m)^2 = \sum p_i x_i^2 - 2\sum p_i x_i m + \sum p_i m^2 = \sum p_i x_i^2 - 2m^2 + m^2 = \sum p_i x_1^2 - m^2$

9 How long did the experiment take ? On what device?

10 The key formula is $\mathrm{E}[(x - m)]^2 = \mathrm{E}[x^2] - (\mathrm{E}[x])^2$ as in Problem 8. The only difference is that expected values E are given by integrals (not sums). The useful identity $\mathrm{E}[(x - m)]^2 = \mathrm{E}[x^2] - m^2$ is still true with $m = \mathrm{E}[x]$.

11 To integrate over the x-y plane, the problem statement shows the correct change of variables from $dxdy$ to $rdrd\theta$. Then $-\infty < x,\ y < \infty$ becomes $0 \le r \le \infty$ and $0 \le \theta \le 2\pi$. The integral of $d\theta$ gives 2π and $\int e^{-r^2/2}\, rdr = \left[-e^{-r^2/2}\right]_0^\infty = 1.$