Problem Set 7.1, page 393

1 Suppose your pulse is measured at \( b_1 = 70 \) beats per minute, then \( b_2 = 120 \), then \( b_3 = 80 \). The least squares solution to three equations \( v = b_1, v = b_2, v = b_3 \) with \( A^T = [1 \ 1 \ 1] \) is \( \hat{v} = (A^T A)^{-1} A^T b = \ldots \). Use calculus and projections:

(a) Minimize \( E = (v - 70)^2 + (v - 120)^2 + (v - 80)^2 \) by solving \( dE/dv = 0 \).

Solution (a) \( \frac{dE}{dv} = 2(v - 70) + 2(v - 120) + 2(v - 80) = 0 \) at the minimizing \( \hat{v} \).

Cancel the 2's: \( 3v = 70 + 120 + 80 = 270 \) so \( \hat{v} = v_{\text{average}} = 90 \)

(b) Project \( b = (70, 120, 80) \) onto \( a = (1, 1, 1) \) to find \( \hat{v} = a^T b / a^T a \).

Solution (b) The projection of \( b \) onto the line through \( a \) is \( p = a \hat{v} : \)

\[
\begin{bmatrix} 70 \\ 120 \\ 80 \end{bmatrix} \quad a = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \hat{v} = \frac{a^T b}{a^T a} = \frac{270}{3} = 90.
\]

2 Suppose \( Av = b \) has \( m \) equations \( a_i v = b_i \) in one unknown \( v \). For the sum of squares \( E = (a_1 v - b_1)^2 + \cdots + (a_m v - b_m)^2 \), find the minimizing \( \hat{v} \) by calculus. Then form \( A^T A v = A^T b \) with one column in \( A \), and reach the same \( \hat{v} \).

Solution To minimize \( E \) we solve \( dE/dv = 0 \). For \( m = 3 \) equations \( a_i v = b_i \),

\[
\frac{dE}{dv} = 2a_1(a_1 v - b_1) + 2a_2(a_2 v - b_2) + 2a_3(a_3 v - b_3) = 0 \quad \text{is zero when}
\]

\[
v = \hat{v} = \frac{a_1 b_1 + a_2 b_2 + a_3 b_3}{a_1^2 + a_2^2 + a_3^2} = \frac{a^T b}{a^T a}.
\]

When \( A \) has one column, \( A^T A \hat{v} = A^T b \) is the same as \( (a^T a) \hat{v} = (a^T b) \).

3 With \( b = (4, 1, 0, 1) \) at the points \( x = (0, 1, 2, 3) \) set up and solve the normal equation for the coefficients \( \hat{v} = (C, D) \) in the nearest line \( C + Dx \). Start with the four equations \( Av = b \) that would be solvable if the points fell on a line.

Solution The unsolvable equation has \( m = 4 \) points on a line: only \( n = 2 \) unknowns.

\[
Av = b \quad \Rightarrow \quad \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \end{bmatrix} \quad \text{leading to} \quad A^T A \hat{v} = A^T b:
\]

\[
\begin{bmatrix} 4 & 6 \\ 6 & 14 \end{bmatrix} \begin{bmatrix} \hat{C} \\ \hat{D} \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} \hat{C} \\ \hat{D} \end{bmatrix} = \frac{1}{20} \begin{bmatrix} 14 & -6 & -6 & 4 \end{bmatrix} = \frac{1}{20} \begin{bmatrix} 60 \\ -60 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}
\]

The closest line to the four points is \( b = 3 - x \).

4 In Problem 3, find the projection \( p = Av \). Check that those four values lie on the line \( C + Dx \). Compute the error \( e = b - p \) and verify that \( A^T e = 0 \).

Solution The projection \( p = A \hat{v} \) is
Find the projection $p = A\hat{v}$ in Problem 7. This gives the three heights of the closest line. Show that the error vector is $e = (2, -6, 4)$.

Solution $p = A\hat{v} = (5, 13, 17)$ gives the heights of the closest line. The error is $b - p = (2, -6, 4)$. This error $e$ has $Pe = Pb - Pp = p - p = 0$. 

5 (Problem 3 by calculus) Write down $E = |b - Av|^2$ as a sum of four squares: the last one is $(1 - C - 3D)^2$. Find the derivative equations $\partial E / \partial C = \partial E / \partial D = 0$. Divide by 2 to obtain $A^T A \hat{v} = A^T b$.

Solution Minimize $E = (4 - C)^2 + (1 - C - D)^2 + (-C - 2D)^2 + (1 - C - 3D)^2$. The partial derivatives are $\partial E / \partial C = 0$ and $\partial E / \partial D = 0$ at the minimum:

$-2(4 - C) - 2(1 - C - D) - 2(-C - 2D) - 2(1 - C - 3D) = 0$

$-2(1 - C - D) - 4(-C - 2D) - 6(1 - C - 3D) = 0$

Factoring out $-2$ and collecting terms this is the same equation $A^T A \hat{v} = A^T b!$

$6 - 4C - 6D = 0$ or $\left[ \begin{array}{c} 4 \\ 6 \\ 14 \end{array} \right] \hat{C} \left[ \begin{array}{c} 6 \\ 4 \end{array} \right] = 0$.

6 For the closest parabola $C + Dt + Et^2$ to the same four points, write down 4 unsolvable equations $Av = b$ for $v = (C, D, E)$. Set up the normal equations for $\hat{v}$. If you fit the best cubic $C + Dt + Et^2 + Ft^3$ to those four points (thought experiment), what is the error vector $e$?

Solution The parabola $C + Dt + Et^2$ fits the 4 points exactly if $Av = b$:

$t = 0 \quad C + 0D + 0E = 4$

$t = 1 \quad C + 1D + 1E = 1$

$t = 2 \quad C + 2D + 4E = 0$

$t = 3 \quad C + 3D + 9E = 1$

$A^T A = \left[ \begin{array}{ccc} 4 & 6 & 14 \\ 6 & 14 & 36 \\ 14 & 36 & 98 \end{array} \right]. \phi A^T b = \left[ \begin{array}{c} 1 + 4 + 0 + 1 \\ 0 + 1 + 0 + 3 \\ 0 + 1 + 0 + 9 \end{array} \right] = \left[ \begin{array}{c} 6 \\ 4 \\ 10 \end{array} \right]$.

The cubic $C + Dt + Et^2 + Ft^3$ can fit 4 points exactly, with error = zero vector.

7 Write down three equations for the line $b = C + Dt$ to go through $b = 7$ at $t = -1, b = 7$ at $t = 1$, and $b = 21$ at $t = 2$. Find the least squares solution $\hat{v} = (C, D)$ and draw the closest line.

Solution $\left[ \begin{array}{cc} 1 & -1 \\ 1 & 1 \\ 1 & 2 \end{array} \right] \left[ \begin{array}{c} C \\ D \end{array} \right] = \left[ \begin{array}{c} 7 \\ 7 \\ 21 \end{array} \right]$. The solution $\hat{x} = \left[ \begin{array}{c} 9 \\ 4 \end{array} \right]$ comes from $\left[ \begin{array}{cc} 3 & 2 \\ 2 & 6 \end{array} \right] \left[ \begin{array}{c} C \\ D \end{array} \right] = \left[ \begin{array}{c} 35 \\ 42 \end{array} \right]$.

Chapter 7. Applied Mathematics and $A^T A$

$P = \left[ \begin{array}{ccc} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{array} \right] \left[ \begin{array}{c} 3 \\ 2 \\ 1 \\ 0 \end{array} \right]$ with error $e = b - p = \left[ \begin{array}{c} -1 \\ -1 \\ 1 \end{array} \right]$

The best line $C + Dx = 3 - x$ does produce $p = (3, 2, 1, 0)$ at the four points $x = 0, 1, 2, 3$. Multiply this $e$ by $A^T$ to get $A^T e = \left[ \begin{array}{c} 0 \\ 0 \end{array} \right]$ as expected.
9 Suppose the measurements at \( t = -1, 1, 2 \) are the errors \( 2, -6, 4 \) in Problem 8. Compute \( \hat{v} \) and the closest line to these new measurements. Explain the answer: \( \mathbf{b} = (2, -6, 4) \) is perpendicular to \( \mathbb{R} \) so the projection is \( \mathbf{p} = \mathbf{0} \).

**Solution** If \( \mathbf{b} \) is previous error \( \mathbf{e} \) then \( \mathbf{b} \) is perpendicular to the column space of \( A \). Projection of \( \mathbf{b} \) is \( \mathbf{p} = \mathbf{0} \).

10 Suppose the measurements at \( t = -1, 1, 2 \) are \( \mathbf{b} = (5, 13, 17) \). Compute \( \hat{v} \) and the closest line \( \mathbf{e} \). The error is \( \mathbf{e} = \mathbf{0} \) because this \( \mathbf{b} \) is \( \mathbb{R} \).

**Solution** If \( \mathbf{b} = A\hat{x} = (5, 13, 17) \) then \( \hat{x} = (9, 4) \) and \( \mathbf{e} = \mathbf{0} \) since \( \mathbf{b} \) is in the column space of \( A \).

11 Find the best line \( C + Dt \) to fit \( \mathbf{b} = 4, 2, -1, 0, 0 \) at times \( t = -2, -1, 0, 1, 2 \).

**Solution** The least squares equation is

\[
\begin{bmatrix}
5 & 0 \\
0 & 10 \\
\end{bmatrix}
\begin{bmatrix}
C \\
D \\
\end{bmatrix}
= 
\begin{bmatrix}
5 \\
-10 \\
\end{bmatrix}.
\]

Solution: \( C = 1, D = -1 \). Line \( 1 - t \). Symmetric \( t \)'s \( \Rightarrow \) diagonal \( A^TA \).

12 Find the plane that gives the best fit to the 4 values \( \mathbf{b} = (0, 1, 3, 4) \) at the corners \( (1, 0) \) and \( (0, 1) \) and \( (-1, 0) \) and \( (0, -1) \) of a square. At those 4 points, the equations \( C + Dx + Ey = b \) are \( A\mathbf{v} = \mathbf{b} \) with 3 unknowns \( \mathbf{v} = (C, D, E) \).

**Solution**

\[
\begin{bmatrix}
1 & 1 & 0 \\
1 & -1 & 0 \\
1 & 0 & -1 \\
\end{bmatrix}
\begin{bmatrix}
C \\
D \\
E \\
\end{bmatrix}
= 
\begin{bmatrix}
0 \\
1 \\
3 \\
4 \\
\end{bmatrix}
\]

has \( A^TA = \begin{bmatrix}
4 & 0 & 0 \\
0 & 2 & 0 \\
0 & 0 & 2 \\
\end{bmatrix} \) and \( A^TA = \begin{bmatrix}
8 \\
2 \\
-2 \\
\end{bmatrix} \\
and \begin{bmatrix}
-3 \\
-3 \\
\end{bmatrix} \\
\]

The solution \( (C, D, E) = (2, -1, 2) \) gives the best plane \( 2 - x - \frac{3}{2}y \).

13 With \( \mathbf{b} = 0, 8, 8, 20 \) at \( t = 0, 1, 3, 4 \) set up and solve the normal equations \( A^TA\mathbf{v} = \mathbf{A}^T\mathbf{b} \). For the best straight line \( C + Dt \), find its four heights \( p_i \) and four errors \( e_i \). What is the minimum value \( E = e_1^2 + e_2^2 + e_3^2 + e_4^2 \)?

**Solution**

\[
\begin{bmatrix}
1 & 0 \\
1 & 1 \\
1 & 3 \\
1 & 4 \\
\end{bmatrix}
\begin{bmatrix}
C \\
D \\
\end{bmatrix}
= 
\begin{bmatrix}
0 \\
8 \\
8 \\
20 \\
\end{bmatrix}
\]

give \( A^TA = \begin{bmatrix}
4 & 8 \\
8 & 26 \\
\end{bmatrix} \) and \( A^TA = \begin{bmatrix}
36 \\
112 \\
\end{bmatrix} \).

\[
A^TA\hat{x} = A^T\mathbf{b}
\]

\[
E = ||e||^2 = 44
\]

\[
\hat{x} = \begin{bmatrix}
1 \\
4 \\
\end{bmatrix}
\]

and \( \mathbf{p} = A\hat{x} = \begin{bmatrix}
1 \\
5 \\
13 \\
17 \\
\end{bmatrix} \) and \( e = \mathbf{b} - \mathbf{p} = \begin{bmatrix}
-1 \\
3 \\
-3 \\
\end{bmatrix} \).

14 (By calculus) Write down \( E = ||\mathbf{b} - A\hat{v}||^2 \) as a sum of four squares—the last one is \( (C + 4D - 20)^2 \). Find the derivative equations \( \partial E/\partial C = 0 \) and \( \partial E/\partial D = 0 \). Divide by \( 2 \) to obtain the normal equations \( A^TA\hat{v} = A^T\mathbf{b} \).

**Solution**

\[
E = (C + 0D)^2 + (C + 1D - 8)^2 + (C + 3D - 8)^2 + (C + 4D - 20)^2.
\]

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

15 Which of the four subspaces contains the error vector \( \mathbf{e} \)? Which contains \( \mathbf{p} \)? Which contains \( \hat{\mathbf{v}} \)?
Solution The error $e$ is contained in the nullspace $N(A^T)$, since $A^T e = 0$. The projection $p$ is contained in the column space $C(A)$. The vector $\hat{v}$ of coefficients can be any vector in $\mathbb{R}^n$.

16 Find the height $C$ of the best horizontal line to fit $b = (0, 8, 8, 20)$. An exact fit would solve the four unsolvable equations $C = 0, C = 8, C = 8, C = 20$. Find the 4 by 1 matrix $A$ in these equations and solve $A^T A \hat{v} = A^T b$.

Solution $E = (C - 0)^2 + (C - 8)^2 + (C - 8)^2 + (C - 20)^2$ and $A^T = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$. $A^T A = [4]$. $A^T b = [36]$ and $(A^T A)^{-1} A^T b = 9$ = best $C$. $e = (-9, -1, -1, 11)$.

17 Write down three equations for the line $b = C + Dt$ to go through $b = 7$ at $t = -1, b = 7$ at $t = 1$, and $b = 21$ at $t = 2$. Find the least squares solution $\hat{v} = (C, D)$ and draw the closest line.

Solution $\begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} [C] = \begin{bmatrix} 7 \\ 21 \end{bmatrix}$. The solution $\hat{v} = \begin{bmatrix} 9 \\ 4 \end{bmatrix}$ comes from $\begin{bmatrix} 3 & 2 & 2 & 6 \end{bmatrix} [C] = \begin{bmatrix} 35 \\ 42 \end{bmatrix}$.

18 Find the projection $p = A\hat{v}$ in Problem 17. This gives the three heights of the closest line. Show that the error vector is $e = (2, -6, 4)$. Why is $P e = 0$?

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

19 Suppose the measurements at $t = -1, 1, 2$ are the errors $2, -6, 4$ in Problem 18. Compute $\hat{v}$ and the closest line to these new measurements. Explain the answer: $b = (2, -6, 4)$ is perpendicular to $\hat{v}$, so the projection is $p = 0$.

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

20 Suppose the measurements at $t = -1, 1, 2$ are $b = (5, 13, 17)$. Compute $\hat{v}$ and the closest line and $e$. The error is $e = 0$ because this $b$ is $\text{in the column space of } A$.

Solution If $b = A\hat{v} = (5, 13, 17)$ then $\hat{v} = (9, 4)$ and $e = 0$ since $b$ is $\text{in the column space of } A$.

Questions 21–26 ask for projections onto lines. Also errors $e = b - p$ and matrices $P$.

21 Project the vector $b$ onto the line through $a$. Check that $e$ is perpendicular to $a$:

(a) $b = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $a = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

(b) $b = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$ and $a = \begin{bmatrix} -1 \\ -3 \\ 1 \end{bmatrix}$.

Solution (a) The projection $p$ is

$$p = a \frac{a^T b}{a^T a} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \frac{6}{3} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}$$

$e = b - p = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ perpendicular to $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Solution (b) In this case the projection is

$$p = a \frac{a^T b}{a^T a} = \begin{bmatrix} -1 \\ -3 \\ -1 \end{bmatrix} \frac{-11}{11} = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$$

and $e = b - p = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$. 
7.1. Least Squares and Projections

22 Draw the projection of $b$ onto $a$ and also compute it from $p = \hat{v}a$:

(a) $b = \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix}$ and $a = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$  
(b) $b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $a = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

**Solution** (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 $a^T b = 0$.

23 In Problem 22 find the projection matrix $P = aa^T / a^Ta$ onto each vector $a$. Verify in both cases that $P^2 = P$. Multiply $Pb$ in each case to find the projection $p$.

**Solution** $P_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $p = P_1 b = \begin{bmatrix} \cos \theta \\ 0 \end{bmatrix}$. $P_2 = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ and $p = P_2 b = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

24 Construct the projection matrices $P_1$ and $P_2$ onto the lines through the $a$'s in Problem 22. Is it true that $(P_1 + P_2)^2 = P_1 + P_2$? This would be true if $P_1P_2 = 0$.

**Solution** The projection matrices $P_1$ and $P_2$ (note correction $P_2$ not $P - 2$) are

$P_1 = \frac{aa^T}{a^Ta} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ \hspace{1cm} $P_2 = \frac{aa^T}{a^Ta} - \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$.

It is not true that $(P_1 + P_2)^2 = P_1 + P_2$. The sum of projection matrices is not usually a projection matrix.

25 Compute the projection matrices $aa^T / a^Ta$ onto the lines through $a_1 = (-1, 2, 2)$ and $a_2 = (2, -1, 2)$. Multiply those two matrices $P_1P_2$ and explain the answer.

**Solution** $P_1 = \frac{1}{9} \begin{bmatrix} 1 & -2 & -2 \\ -2 & 4 & 4 \\ -2 & 4 & 4 \end{bmatrix}$, $P_2 = \frac{1}{9} \begin{bmatrix} 4 & 4 & -2 \\ 4 & 4 & -2 \\ -2 & -2 & 1 \end{bmatrix}$.

$P_1P_2$ is zero matrix because $a_1$ is perpendicular to $a_2$.

26 Continuing Problem 25, find the projection matrix $P_3$ onto $a_3 = (2, -1, 2)$. Verify that $P_1 + P_2 + P_3 = I$. The basis $a_1, a_2, a_3$ is orthogonal!

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

We can add projections onto orthogonal vectors. This is important.

27 Project the vector $b = (1, 1)$ onto the lines through $a_1 = (1, 0)$ and $a_2 = (1, 2)$. Draw the projections $p_1$ and $p_2$ and add $p_1 + p_2$. The projections do not add to $b$ because the $a$'s are not orthogonal.

**Solution** The projections of $(1, 1)$ onto the lines through $(1, 0)$ and $(1, 2)$ are $p_1 = (1, 0)$ and $p_2 = (3/5, 6/5) = (0.6, 1.2)$. Then $p_1 + p_2 \neq b$.

28 (Quick and recommended) Suppose $A$ is the 4 by 4 identity matrix with its last column removed. $A$ is 4 by 3. Project $b = (1, 2, 3, 4)$ onto the column space of $A$. What shape is the projection matrix $P$ and what is $P$?

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


29 If \( A \) is doubled, then \( P = 2A(4A^TA)^{-1}2A^T \). This is the same as \( A(A^TA)^{-1}A^T \). The column space of \( 2A \) is the same as _____ . Is \( \tilde{b} \) the same for \( A \) and \( 2A \) ?

\[ Solution \] 2\( A \) has the same column space as \( A \). Same \( p \). But \( \tilde{x} \) for \( 2A \) is half of \( \tilde{x} \) for \( A \).

30 What linear combination of \( (1, 2, -1) \) and \( (1, 0, 1) \) is closest to \( b = (2, 1, 1) \)?

\[ Solution \] \( \frac{1}{2}(1, 2, -1) + \frac{3}{2}(1, 0, 1) = (2, 1, 1) \). So \( b \) is in the plane: no error \( e \). Projection shows \( Pb = b \).

31 (Important) If \( P^2 = P \) show that \( (I - P)^2 = I - P \). When \( P \) projects onto the column space of \( A \), \( I - P \) projects onto which fundamental subspace ?

\[ Solution \] 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.

32 If \( P \) is the 3 by 3 projection matrix onto the line through \( (1, 1, 1) \), then \( I - P \) is the projection matrix onto _____.

\[ Solution \] \( I - P \) is the projection onto the plane \( x_1 + x_2 + x_3 = 0 \), perpendicular to the direction \( (1, 1, 1) \):

\[ I - P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}. \]

33 Multiply the matrix \( P = A(A^TA)^{-1}A^T \) by itself. Cancel to prove that \( P^2 = P \). Explain why \( P(Pb) \) always equals \( Pb \). The vector \( Pb \) is in the column space so its projection is _____.

\[ Solution \] \( (A(A^TA)^{-1}A^T)^2 = A(A^TA)^{-1}(A^TA)(A^TA)^{-1}A^T = A(A^TA)^{-1}A^T \). So \( P^2 = P \). Geometric reason: \( Pb \) is in the column space (where \( P \) projects). Then its projection \( P(Pb) \) is \( Pb \) for every \( b \). So \( P^2 = P \).

34 If \( A \) is square and invertible, the warning against splitting \( (A^TA)^{-1} \) does not apply. Then \( AA^{-1}(A^TA)^{-1}A^T = I \) is true. When \( A \) is invertible, why is \( P = I \) and \( e = 0 \)?

\[ Solution \] If \( A \) is invertible then its column space is all of \( \mathbb{R}^n \). So \( P = I \) and \( e = 0 \).

35 An important fact about \( A^TA \) is this: If \( A^TAx = 0 \) then \( Ax = 0 \). New proof: The vector \( Ax \) is in the nullspace of _____. \( Ax \) is always in the column space of _____. To be in both of those perpendicular spaces, \( Ax \) must be zero.

\[ Solution \] If \( A^TAx = 0 \) then \( Ax \) is in the nullspace of \( A^T \). But \( Ax \) is always in the column space of \( A \). To be in both of those perpendicular spaces, \( Ax \) must be zero. So \( A \) and \( A^TA \) have the same nullspace.

Notes on mean and variance and test grades

If all grades on a test are 90, the mean is \( m = 90 \) and the variance is \( \sigma^2 = 0 \). Suppose the expected grades are \( g_1, \ldots, g_N \). Then \( \sigma^2 \) comes from squaring distances to the mean:

\[ \text{Mean} \ m = \frac{g_1 + \cdots + g_N}{N} \quad \text{Variance} \ \sigma^2 = \frac{(g_1 - m)^2 + \cdots + (g_N - m)^2}{N} \]

After every test my class wants to know \( m \) and \( \sigma \). My expectations are usually way off.
36 Show that \( \sigma^2 \) also equals \( \frac{1}{N}(g_1^2 + \cdots + g_N^2) - m^2 \).

Solution Each term \((g_i - m)^2\) equals \(g_i^2 - 2gm + m^2\), so

\[
\sigma^2 = \frac{(\text{sum of } g_i^2)}{N} - 2m \left( \frac{\text{sum of } g_i}{N} \right) + \frac{Nm^2}{N} = \frac{(\text{sum of } g_i^2)}{N} - \frac{2Nm^2}{N} + \frac{Nm^2}{N} = \frac{(\text{sum of } g_i^2)}{N} - m^2.
\]

37 If you flip a fair coin \( N \) times (1 for heads, 0 for tails) what is the expected number \( m \) of heads? What is the variance \( \sigma^2 \)?

Solution For a fair coin you expect \( \frac{N}{2} \) heads in \( N \) flips. The variance \( \sigma^2 \) turns out to be \( \frac{N}{4} \).

Problem Set 7.2, page 402

1 For a 2 by 2 matrix, suppose the 1 by 1 and 2 by 2 determinants \( a \) and \( ac - b^2 \) are positive. Then \( c > b^2/a \) is also positive.

(i) \( \lambda_1 \) and \( \lambda_2 \) have the same sign because their product \( \lambda_1 \lambda_2 \) equals \( \_\_\_\_\_\_ \).

(i) That sign is positive because \( \lambda_1 + \lambda_2 = \_\_\_\_\_\_ \).

Conclusion: The tests \( a > 0 \), \( ac - b^2 > 0 \) guarantee positive eigenvalues \( \lambda_1, \lambda_2 \).

Solution 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 \( a + c > 0 \)).

2 Which of \( S_1, S_2, S_3, S_4 \) has two positive eigenvalues? Use \( a \) and \( ac - b^2 \), don’t compute the \( \lambda \)'s. Find an \( x \) with \( x^T S_1 x < 0 \), confirming that \( A_1 \) fails the test.

\[
S_1 = \begin{bmatrix} 5 & 6 \\ 6 & 7 \end{bmatrix} \quad S_2 = \begin{bmatrix} -1 & -2 \\ -2 & -5 \end{bmatrix} \quad S_3 = \begin{bmatrix} 1 & 10 \\ 10 & 100 \end{bmatrix} \quad S_4 = \begin{bmatrix} 1 & 10 \\ 10 & 101 \end{bmatrix}.
\]

Solution Only \( S_4 \) has two positive eigenvalues since \( 101 > 10^2 \).

\( x^T S_1 x = 5x_1^2 + 12x_1x_2 + 7x_2^2 \) 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 \( 0 \); \( S_3 \) has \( \det = 0 \).
3 For which numbers \( b \) and \( c \) are these matrices positive definite?

\[
S = \begin{bmatrix} 1 & b \\ b & 9 \end{bmatrix} \quad S = \begin{bmatrix} 2 & 4 \\ 4 & c \end{bmatrix} \quad S = \begin{bmatrix} c & b \\ b & c \end{bmatrix}.
\]

**Solution**

Positive definite for 

\[-3 < b < 3\]

\[
\begin{bmatrix} 1 & b \\ b & 9 - b^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ b & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 9 - b^2 \end{bmatrix} \begin{bmatrix} 1 & b \\ b & 1 \end{bmatrix} = LDL^T
\]

Positive definite for \( c > 8 \)

\[
\begin{bmatrix} 2 & 4 \\ 4 & c - 8 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & c - 8 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix} = LDL^T.
\]

Positive definite for \( c > b \)

\[
L = \begin{bmatrix} 1 & 1 \\ -b/c & 0 \end{bmatrix} \quad D = \begin{bmatrix} c & 0 \\ 0 & c - b/c \end{bmatrix} \quad S = LDL^T.
\]

4 What is the energy \( q = ax^2 + 2bxy + cy^2 = x^T S x \) for each of these matrices? Complete the square to write \( q \) as a sum of squares \( d_1(\quad)^2 + d_2(\quad)^2 \).

\[
S = \begin{bmatrix} 1 & 2 \\ 2 & 9 \end{bmatrix} \quad \text{and} \quad S = \begin{bmatrix} 1 & 3 \\ 3 & 9 \end{bmatrix}.
\]

**Solution**

\( f(x, y) = x^2 + 4xy + 9y^2 = (x + 2y)^2 + 5y^2; \quad x^2 + 6xy + 9y^2 = (x + 3y)^2 \).

5 \( x^T S x \) certainly has a saddle point and not a minimum at \((0, 0)\). What symmetric matrix \( S \) produces this energy? What are its eigenvalues?

**Solution**

\( x^T S x = 2x_1x_2 \) comes from \( S = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \) which has eigenvalues 1 and \(-1\): \( S \) is **indefinite**.

6 Test to see if \( A^T A \) is positive definite in each case:

\[
A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix}.
\]

**Solution**

The first and second matrices have independent columns in \( A \), so \( A^T A \) is positive definite. The third matrix has dependent columns so \( A^T A \) is only **positive semidefinite**.

7 Which 3 by 3 symmetric matrices \( S \) and \( T \) produce these quadratic energies?

\( x^T S x = 2(x_1^2 + x_2^2 + x_3^2 - x_1x_2 - x_2x_3) \). Why is \( S \) positive definite?

\( x^T T x = 2(x_1^2 + x_2^2 + x_3^2 - x_1x_2 - x_1x_3 - x_2x_3) \). Why is \( T \) semidefinite?

**Solution**

\[
S = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \quad \text{is positive definite—its}
\]

determinants are \( D_1 = 2, D_2 = 3, D_3 = 4 \).
7.2. Positive Definite Matrices and the SVD

\[ T = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \]

is positive semidefinite with determinants \( D_1 = 2, D_2 = 3, D_3 = 0 \).

The energy \( x^T T x = 0 \) when \( x = (1, 1, 1) \).

8 Compute the three upper left determinants of \( S \) to establish positive definiteness. (The first is 2.) Verify that their ratios give the second and third pivots.

\[
\text{Pivots} = \text{ratios of determinants} \quad S = \begin{bmatrix} 2 & 2 & 0 \\ 2 & 5 & 3 \\ 0 & 3 & 8 \end{bmatrix}.
\]

Solution The upper left determinants of \( S \) are 2, 6, 30. The pivots are 2, 3, 5 (ratios of determinants). Notice that the product of pivots is 30.

9 For what numbers \( c \) and \( d \) are \( S \) and \( T \) positive definite? Test the 3 determinants:

\[
S = \begin{bmatrix} c & 1 & 1 \\ 1 & c & 1 \\ 1 & 1 & c \end{bmatrix} \quad \text{and} \quad T = \begin{bmatrix} 1 & 2 & 3 \\ 2 & d & 4 \\ 3 & 4 & 5 \end{bmatrix}.
\]

Solution For \( c = 1 \), the matrix \( S \) has eigenvalues 3, 0, 0. For any \( c \), the eigenvalues all add \( c - 1 \). So \( S \) is positive definite for \( c > 1 \). (Same answer using determinants.) For \( T \) the determinants are \( 1, d - 4, -4d + 12 \). If \( d > 4 \) then \( -4d + 12 \) is negative! So \( T \) is never positive definite for any \( d \).

10 If \( S \) is positive definite then \( S^{-1} \) is positive definite. Best proof: The eigenvalues of \( S^{-1} \) are positive because ______. Second proof (only for 2 by 2):

The entries of \( S^{-1} = \frac{1}{ac - b^2} \begin{bmatrix} c & -b \\ -b & a \end{bmatrix} \) pass the determinant tests ______.

Solution Positive definite \( \Rightarrow \) all eigenvalues \( \lambda > 0 \) \( \Rightarrow \) all eigenvalues \( 1/\lambda \) of \( S^{-1} \) are positive. Also for \( 2 \times 2 \): the determinant tests are passed.

11 If \( S \) and \( T \) are positive definite, their sum \( S + T \) is positive definite. Pivots and eigenvalues are not convenient for \( S + T \). Better to prove \( x^T (S + T) x > 0 \).

Solution Energy \( x^T (S + T) x = x^T S x + x^T T x > 0 + 0 \)

12 A positive definite matrix cannot have a zero (or even worse, a negative number) on its diagonal. Show that this matrix fails to have \( x^T S x > 0 \):

\[
\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 4 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \] is not positive when \( (x_1, x_2, x_3) = ( , , ) \).

Solution \( x^T S x \) is zero when \( x = (0, 1, 0) \).
Chapter 7. Applied Mathematics and $A^TA$

13 A diagonal entry $a_{jj}$ of a symmetric matrix cannot be smaller than all the $\lambda$'s. If it were, then $A - a_{jj}I$ would have ____ eigenvalues and would be positive definite. But $A - a_{jj}I$ has a ____ on the main diagonal.

**Solution** If $a_{jj}$ is smaller than all eigenvalues, then $A - a_{jj}I$ would have positive eigenvalues. But this matrix has a zero on the diagonal. But Problem 13, it can’t be positive definite. So $a_{jj}$ can’t be smaller than all eigenvalues!

14 Show that if all $\lambda > 0$ then $x^TSx > 0$. We must do this for every nonzero $x$, not just the eigenvectors. So write $x$ as a combination of the eigenvectors and explain why all “cross terms” are $x_i^T x_j = 0$. Then $x^T S x = (c_1 x_1 + \cdots + c_n x_n)^T (c_1 \lambda_1 x_1 + \cdots + c_n \lambda_n x_n) = c_1^2 \lambda_1 x_1^T x_1 + \cdots + c_n^2 \lambda_n x_n^T x_n > 0$.

**Solution** The “cross terms” have the form $(c_i x_i)^T (c_j \lambda_j x_j)$. This is zero because symmetric matrices $S$ have orthogonal eigenvectors.

15 Give a quick reason why each of these statements is true:

(a) Every positive definite matrix is invertible.
(b) The only positive definite projection matrix is $P = I$.
(c) A diagonal matrix with positive diagonal entries is positive definite.
(d) A symmetric matrix with a positive determinant might not be positive definite!

**Solution**

(a) All $\lambda_i > 0$ so zero is not an eigenvalue and $S$ is invertible
(b) All projection matrices except $P = I$ are singular
(c) The energy for a positive diagonal matrix is $x^T Dx = d_1 x_1^2 + \cdots + d_n x_n^2 > 0$ when $x \neq 0$
(d) $S = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$ has $\det S = 1$ but $S$ is **negative** definite

16 With positive pivots in $D$, the factorization $S = LDL^T$ becomes $L\sqrt{D}\sqrt{D}^T$. (Square roots of the pivots give $D = \sqrt{D}\sqrt{D}$.) Then $A = \sqrt{D}L^T$ yields the **Cholesky factorization** $S = A^TA$ which is “symmetrized $LU$”:

From $A = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$ find $S$. From $S = \begin{bmatrix} 4 & 8 \\ 8 & 25 \end{bmatrix}$ find $A = \text{chol}(S)$.

**Solution** If $A = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$ then $A^TA = \begin{bmatrix} 9 & 3 \\ 3 & 5 \end{bmatrix}$ = positive definite $S$.

$S = \begin{bmatrix} 4 & 8 \\ 8 & 25 \end{bmatrix} = LU = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 8 \\ 0 & 9 \end{bmatrix} = LDL^T = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 2 \\ 0 & 1 \end{bmatrix}$

so $A = \sqrt{D}L^T = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 0 & 3 \end{bmatrix}$. 
7.2. Positive Definite Matrices and the SVD

17 Without multiplying \( S = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \), find

(a) the determinant of \( S \)  
(b) the eigenvalues of \( S \)
(c) the eigenvectors of \( S \)  
(d) a reason why \( S \) is symmetric positive definite.

Solution \( \det S = 10 \), \( \lambda(S) = 2 \) and 5, eigenvectors \((\cos \theta, \sin \theta)\) and \((-\sin \theta, \cos \theta)\), \( S \) has positive eigenvalues.

18 For \( F_1(x, y) = \frac{1}{4}x^4 + x^2y + y^2 \) and \( F_2(x, y) = x^3 + xy - x \) find the second derivative matrices \( H_1 \) and \( H_2 \):

\[ H_1 = \begin{bmatrix} 3x^2 + 2y & 2x \\ 2x & 2 \end{bmatrix}, \quad H_2 = x^3 + xy - x \]

Test for minimum \( H = \begin{bmatrix} \partial^2 F / \partial x^2 & \partial^2 F / \partial x \partial y \\ \partial^2 F / \partial y \partial x & \partial^2 F / \partial y^2 \end{bmatrix} \) is positive definite

\( H_1 \) is positive definite so \( F_1 \) is concave up (= convex). Find the minimum point of \( F_1 \) and the saddle point of \( F_2 \) (look only where first derivatives are zero).

Solution \( F_1 = \frac{1}{4}x^4 + x^2y + y^2 \) has \( \partial F_1 / \partial x = x^3 + 2xy \) and \( \partial F_1 / \partial y = x^2 + 2y \). Then the 2nd derivatives are

\[ H_1 = \begin{bmatrix} 3x^2 + 2y & 2x \\ 2x & 2 \end{bmatrix}, \quad F_2 = x^3 + xy - x \]

19 The graph of \( z = x^2 + y^2 \) is a bowl opening upward. The graph of \( z = x^2 - y^2 \) is a saddle. The graph of \( z = -x^2 - y^2 \) is a bowl opening downward. What is a test on \( a, b, c \) for \( z = ax^2 + 2bxy + cy^2 \) to have a saddle point at \((0, 0)\)?

Solution \( ax^2 + 2bxy + cy^2 \) has a saddle point \((0, 0)\) if \( \partial z / \partial x = \partial z / \partial y = 0 \) (which is true) and if \( H = 2 \begin{bmatrix} a & b \\ c & d \end{bmatrix} \) is positive definite.

20 Which values of \( c \) give a bowl and which \( c \) give a saddle point for the graph of \( z = 4x^2 + 12xy + cy^2 \)? Describe this graph at the borderline value of \( c \).

Solution The matrix for this problem is \( S = \begin{bmatrix} 4 & 6 \\ 6 & c \end{bmatrix} \) and this has a saddle for \( c < 9 \). Then \( \lambda_1 > 0 > \lambda_2 \) because the determinants are \( 4 > 0 \) and \( 4c - 3b < 0 \).

21 When \( S \) and \( T \) are symmetric positive definite, \( ST \) might not even be symmetric. But its eigenvalues are still positive. Start from \( STx = \lambda x \) and take dot products with \( Tx \).

Then prove \( \lambda > 0 \).

Solution If \( STx = \lambda x \) then \( (Tx)^T STx = \lambda(Tx)^T x \). Left side > 0 because \( S \) is positive definite, right side has \( x^T Tx > 0 \) because \( T \) is positive definite. Therefore \( \lambda > 0 \).

22 Suppose \( C \) is positive definite (so \( y^T Cy > 0 \) whenever \( y \neq 0 \)) and \( A \) has independent columns (so \( Ax \neq 0 \) whenever \( x \neq 0 \)). Apply the energy test to \( x^T A^T CAx \) to show that \( A^T CA \) is positive definite: the crucial matrix in engineering.

Solution \( x^T A^T CAx = y^T Cy > 0 \) because \( y = Ax \) is only zero when \( x \) is zero (\( A \) has independent columns).
23 Find the eigenvalues and unit eigenvectors \( v_1, v_2 \) of \( A^T A \). Then find \( u_1 = A v_1 / \sigma_1 \):

\[
A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \quad \text{and} \quad A^T A = \begin{bmatrix} 10 & 20 \\ 20 & 40 \end{bmatrix} \quad \text{and} \quad A A^T = \begin{bmatrix} 5 & 15 \\ 15 & 45 \end{bmatrix}.
\]

Verify that \( u_1 \) is a unit eigenvector of \( A A^T \). Complete the matrices \( U, \Sigma, V \).

**Solution** \( A^T A \) has eigenvalues 50 and 0. Its eigenvectors are \( v_1 = (1, 2)/\sqrt{5} \) and \( v_2 = (-2, 1)/\sqrt{5} \). Then \( u_1 = A v_1 / \sqrt{50} = (50, 100)/\sqrt{250} \).

The SVD is

\[
\begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 & v_2 \end{bmatrix}^T.
\]

24 Write down orthonormal bases for the four fundamental subspaces of this \( A \).

**Solution** \( A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \) has bases \( \begin{bmatrix} 1 & 3 & 0 \\ 2 & 6 & 0 \end{bmatrix}/\sqrt{10} \) for \( \mathcal{C}(A) \), \( \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}/\sqrt{5} \) for row space \( \mathcal{C}(A^T) \), \( \begin{bmatrix} 2 & 0 \\ -1 & 1 \end{bmatrix}/\sqrt{5} \) for \( \mathcal{N}(A) \), \( \begin{bmatrix} 3 & 0 \\ -1 & 1 \end{bmatrix}/\sqrt{10} \) for \( \mathcal{N}(A^T) \).

25 (a) Why is the trace of \( A^T A \) equal to the sum of all \( a_{ij}^2 \)?

(b) For every rank-one matrix, why is \( \sigma_1^2 = \text{sum of all } a_{ij}^2 \)?

**Solution** The diagonal entries of \( A^T A \) are \( ||\text{column } 1||^2 \) to \( ||\text{column } n||^2 \). The sum of those is the sum of all \( a_{ij}^2 \). The trace of \( A^T A \) is always the sum of all \( \sigma_i^2 \) and for a rank one matrix that sum is only \( \sigma_1^2 \).

26 Find the eigenvalues and unit eigenvectors of \( A^T A \) and \( A A^T \). Keep each \( A v = \sigma u \): 

**Fibonacci matrix** \quad \( A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \)

Construct the singular value decomposition and verify that \( A = U \Sigma V^T \).

**Solution** \( A \) is symmetric with \( A^T A = A^2 = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \) with eigenvalues \( x \) from \( x^2 - 3x + 1 = 0 \) and \( x = \frac{1}{2} \left( 3 \pm \sqrt{5} \right) \). Then \( \sigma = \sqrt{x} = \frac{1}{2} \left( \sqrt{5} \pm 1 \right) \).

27 Compute \( A^T A \) and \( A A^T \) and their eigenvalues and unit eigenvectors for \( V \) and \( U \).

**Rectangular matrix** \quad \( A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \).

Check \( A V = U \Sigma \) (this will decide \( \pm \) signs in \( U \)). \( \Sigma \) has the same shape as \( A \).
My favorite example of the SVD is when

\[
\begin{bmatrix}
2 & 1 \\
1 & 2
\end{bmatrix}
\]

has eigenvalues 3 and 1, so \( A \) has singular values \( \sqrt{3} \) and 1. The unit eigenvectors are \((1,1)/\sqrt{2}\) and \((1,-1)/\sqrt{2}\). \(AA^T = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix} \)

has eigenvalues 3 and 1 and 0 and eigenvectors

\[
\begin{bmatrix}
1 \\
2 \\
1
\end{bmatrix}, \begin{bmatrix}
1 \\
1 \\
-1
\end{bmatrix}
\]

divided by \( \sqrt{6}, \sqrt{2}, \sqrt{3} \).

Construct the matrix with rank one that has \( Av = 12\mathbf{u} \) for \( v = \frac{1}{2}(1,1,1,1) \) and \( \mathbf{u} = \frac{1}{3}(2,2,1) \). Its only singular value is \( \sigma_1 = \boxed{12} \).

Suppose \( A \) is invertible (with \( \sigma_1 > \sigma_2 > 0 \)). Change \( A \) by as small a matrix as possible to produce a singular matrix \( A_0 \). Hint: \( U \) and \( V \) do not change.

\[
A = \begin{bmatrix}
\mathbf{u}_1 & \mathbf{u}_2 \\
\sigma_1 & \sigma_2
\end{bmatrix}
\begin{bmatrix}
\mathbf{v}_1 & \mathbf{v}_2
\end{bmatrix}^T
\]

find the nearest \( A_0 \).

The nearest singular matrix is \( A_0 = U \begin{bmatrix} \sigma_1 & 0 \\ 0 & 0 \end{bmatrix} V^T \). Since \( U \) and \( V \) are orthogonal matrices, the size of \( A - A_0 \) is only \( \sigma_2 \). In other words, \( u_1 \sigma_1 v_1^T \) is the closest rank 1 matrix to \( A \).

The SVD for \( A + I \) doesn’t use \( \Sigma + I \). Why is \( \sigma(A + I) \) not just \( \sigma(A) + I \)?

Multiply \( A^T A \) by \( A \). Put in parentheses to show that \( A \) is an eigenvector of \( A A^T \). Those are not the eigenvectors of \( A^T A \) (or \( A^T + I \)).

\[
A^T Av = \sigma^2 v \]

by \( A \). Put in parentheses to show that \( A \) is an eigenvector of \( AA^T \). We divide by its length \( ||Av|| = \sigma \) to get the unit eigenvector \( u \).

\[
A \times A^T Av = \sigma^2 v \]

is \( (AA^T)Av = \sigma^2 (Av) \). So \( Av \) is an eigenvector of \( AA^T \).

My favorite example of the SVD is when \( Av(x) = dv/dx \), with the endpoint conditions \( v(0) = 0 \) and \( v(1) = 0 \). We are looking for orthogonal functions \( v(x) \) so that their derivatives \( Av = dv/dx \) are also orthogonal. The perfect choice is \( v_1 = \sin \pi x \) and \( v_2 = \sin 2\pi x \) and \( v_k = \sin k\pi x \). Then each \( u_k \) is a cosine.

The derivative of \( v_1 \) is \( Av_1 = \pi \cos \pi x = \pi u_1 \). The singular values are \( \sigma_1 = \pi \) and \( \sigma_k = k\pi \). Orthogonality of the sines (and orthogonality of the cosines) is the foundation for Fourier series.

You may object to \( AV = U\Sigma \). The derivative \( A = d\mathbf{v}/dx \) is not a matrix ! The orthogonal factor \( V \) has functions \( \sin k\pi x \) in its columns, not vectors. The matrix \( U \) has cosine functions \( \cos k\pi x \). Since when is this allowed? One answer is to refer you to the chebfun package on the web. This extends linear algebra to matrices whose columns are functions—not vectors.
Another answer is to replace \(d/dx\) by a first difference matrix \(A\). Its shape will be \(N+1\) by \(N\). \(A\) has 1’s down the diagonal and \(-1\)’s on the diagonal below. Then \(AV = U\Sigma\) has discrete sines in \(V\) and discrete cosines in \(U\). For \(N = 2\) those will be sines and cosines of \(30^\circ\) and \(60^\circ\) in \(v_1\) and \(u_1\).

** Can you construct the difference matrix \(A\) (3 by 2) and \(A^T A\) (2 by 2)? The discrete sines are \(v_1 = (\sqrt{3}/2, \sqrt{3}/2)\) and \(v_2 = (\sqrt{3}/2, -\sqrt{3}/2)\). Test that \(Av_1\) is orthogonal to \(Av_2\). What are the singular values \(\sigma_1\) and \(\sigma_2\) in \(\Sigma\)?

Solution The sines and cosines are perfect examples of the \(v\)’s and \(u\)’s for the operator (infinite-dimensional matrix) \(A = \text{derivative } d/dx\). The sines \(v_k = \sin \pi kx\) are orthogonal, the cosines \(u_k = \cos \pi kx\) are orthogonal, and \(Av_k = \sigma_k u_k\). (The derivative of a sine is a cosine with \(\sigma_k = \pi k\).) For differences instead of derivatives, we can try the matrix \(A = \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{bmatrix}\).

** Problem Set 7.3, page 413

1 Transpose the derivative with integration by parts: \((dy/dx, g) = -(y, dg/dx)\). \(Ay\) is \(dy/dx\) with boundary conditions \(y(0) = 0\) and \(y(1) = 0\). Why is \(\int y'gdx\) equal to \(-\int yg'dx\)? Then \(A^T\) (which is normally written as \(A^*\)) is \(A^T g = -dg/dx\) with no boundary conditions on \(g\). \(A^T A\) is \(-y''\) with \(y(0) = 0\) and \(y(1) = 0\).

Solution Integration by parts for 0 ≤ \(x\) ≤ 1 produces boundary terms at \(x = 0\) and 1:

\[
\int_0^1 \frac{dy}{dx} g(x) \, dx = -\left. \int_0^1 y(x) \frac{dg}{dx} \, dx + y(x) g(x) \right|^{x=1}_{x=0}
\]

The boundary terms are zero if \(y(0) = y(1) = 0\). Then the adjoint (or transpose) of \(d/dx\) is \(-d/dx\), with no boundary condition on \(g\) when there are 2 boundary conditions on \(y\) (fixed-fixed).

Problems 2-6 have boundary conditions at \(x = 0\) and \(x = 1\): no initial conditions.

2 Solve this boundary value problem in two steps. Find the complete solution \(y_p + y_n\) with two constants in \(y_n\), and find those constants from the boundary conditions:

Solve \(-y'' = 12x^2\) with \(y(0) = 0\) and \(y(1) = 0\) and \(y_p = -x^4\).

Solution \(y_p = -x^4\) solves \(-y_p'' = 12x^2\). It has \(y_p(0) = 0\) and \(y_p = -1\). We need to add the solution to \(-Y'' = 0\) with \(Y(0) = 0\) and \(Y(1) = 1\). Then \(Y = A + Bx\) has \(A = 0\) and \(B = 1\). The complete solution is \(y = -x^4 + x\).

3 Solve the same equation \(-y'' = 12x^2\) with \(y(0) = 0\) and \(y'(1) = 0\) (zero slope).

Solution Changing \(y(1) = 0\) to \(y'(1) = 0\) will change the solution to \(y = -x^4 + Bx\) with \(y' = -4x^3 + B\). For \(y'(1) = 0\) we need \(B = 4\).
4. Solve the same equation \(-y'' = 12x^2\) with \(y'(0) = 0\) and \(y(1) = 0\). Then try for both slopes \(y'(0) = 0\) and \(y'(1) = 0\): this has no solution \(y = -x^4 + Ax + B\).

**Solution** With \(y'(0) = 0\) the solution we want is \(y = -x^4 + A\). The constant \(A\) is determined by \(y(1) = -1 + A = 0\). We cannot have \(y'(1) = 0\) because \(y' = -4x^3\).

5. Solve \(-y'' = 6x\) with \(y(0) = 2\) and \(y(1) = 4\). Boundary values need not be zero.

**Solution** \(-y'' = 6x\) leads to \(y = -x^3 + A + Bx\). The boundary conditions are \(y(0) = A = 2\) and \(y(1) = -1 + 2 + B = 4\). Then \(B = 3\) and \(y = -x^3 + 2 + 3x\).

6. Solve \(-y'' = e^x\) with \(y(0) = 5\) and \(y(1) = 0\), starting from \(y = y_p + y_n\).

**Solution** \(-y'' = e^x\) leads to \(y = -e^x + A + Bx\). The first boundary condition is \(y(0) = -1 + A = 5\) so that \(A = 6\). Then \(y(1) = -e + 6 + B = 0\) and \(B = e - 6\).

Problems 7-11 are about the LU factors and the inverses of second difference matrices.

7. The matrix \(T\) with \(T_{11} = 1\) factors perfectly into \(LU = A^T A\) (all its pivots are 1).

\[
T = \begin{bmatrix} 1 & -1 & 2 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = LU.
\]

Each elimination step adds the pivot row to the next row (and \(L\) subtracts to recover \(T\) from \(U\)). The inverses of those difference matrices \(L\) and \(U\) are sum matrices. Then the inverse of \(T = LU\) is \(U^{-1}L^{-1}\):

\[
T^{-1} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = U^{-1}L^{-1}.
\]

Compute \(T^{-1}\) for \(N = 4\) (as shown) and for any \(N\).

**Solution** \(T^{-1}\) is fixed-free second difference matrix.

8. The matrix equation \(TY = (0, 1, 0, 0) = \text{delta vector}\) is like the differential equation \(-y'' = \delta(x - a)\) with \(a = 2\Delta x = \frac{2}{3}\). The boundary conditions are \(y'(0) = 0\) and \(y(1) = 0\). Solve for \(y(x)\) and graph it from 0 to 1. Also graph \(Y = \text{second column of } T^{-1}\) at the points \(x = \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}\). The two graphs are ramp functions.

**Solution** Two integrations of the delta function \(\delta(x)\) will produce the unit ramp \(R(x) = 0\) for \(x \leq 0\), \(R(x) = x\) for \(x \geq 0\). Shifting \(\delta(x)\) to \(\delta\left(x - \frac{2}{3}\right)\) will shift the solution to \(y = -R\left(x - \frac{2}{3}\right) + A + Bx\). Then \(y'(0) = -1 + B\) gives \(B = 1\), and \(y(1) = 0\) gives \(-\frac{3}{5} + A + 1 = 0\) and \(A = -\frac{2}{5}\).
The matrix $B$ has $B_{11} = 1$ (like $T_{11} = 1$) and also $B_{NN} = 1$ (where $T_{NN} = 2$). Why does $B$ have the same pivots 1, 1, . . . as $T$, except for zero in the last pivot position? The early pivots don’t know $B_{NN} = 1.$

Then $B$ is not invertible: $-y'' = \delta(x-a)$ has no solution with $y'(0) = y'(1) = 0.$

**Solution** $B$ starts with the pivots 1, 1, 1, . . . (as $T$ did) but reducing the $N, N$ entry by 1 will reduce the last pivot by 1. So we have last pivot = zero and $B$ is not invertible. The analog for differential equations is $y' = 0$ at both endpoints: No ramp function except $y = 0$ can meet those boundary conditions.

When you compute $K^{-1}$, multiply by $\det K = N + 1$ to get nice numbers:

**Column 2 of $5K^{-1}$** solves the equation $Kv = 5\delta$ when the delta vector is $\delta = \begin{bmatrix} 1 & 1 & 1 & \ldots & 1 \end{bmatrix}$. We know from $KK^{-1} = I$ that $K$ times each column of $K^{-1}$ is a delta vector.

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

**Solution** Column 2 of $5K^{-1}$ is like the solution to $-y'' = 5\delta (x-\frac{2}{5})$. The column of $5K^{-1}$ has a max in row 2 and the solution $y(x)$ has a max at $x = \frac{2}{5}$.

$K$ comes with two boundary conditions. $T$ only has $y(1) = 0$. $B$ has no boundary conditions on $y$. Verify that $K = A^T A$. Then remove the first row of $A$ to get $T = A_1^T A_1$. Then remove the last row to get dependent rows: $B = A_0^T A_0$.

The backward first difference $A = \begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$ gives $K = A^T A$.

**Solution** $A$ is the matrix in Problem 7 with 1’s on the main diagonal and $-1$’s on the diagonal above. $A^T A$ is the symmetric second difference matrix with three nonzero diagonals. Those diagonals contain $-1$’s and 2’s and $-1$’s. Then removing the top row of $A$ gives a rectangular $A_1$ with $A_1^T A_1 = T$ as in Problem 7 ($T_{11} = 1$ not 2). Removing the last row gives $A_2$ with $A_2^T A_2 = B$ and $B_{NN} = 1$ not 2.

**Multiply $K_3$ by its eigenvector $y_n = (\sin n\pi h, \sin 2n\pi h, \sin 3n\pi h)$ to verify that the eigenvalues $\lambda_1, \lambda_2, \lambda_3$ are $\lambda_n = 2 - 2\cos \frac{n\pi}{2}$ in $K y_n = \lambda_n y_n$. This uses the trigonometric identity $\sin(A + B) + \sin(A - B) = 2 \sin A \cos B$.**

**Solution** The eigenvectors of $K$ are “sine vectors” just as the eigenfunctions of $-y'' = \lambda y$ with $y(0) = 0 = y(1)$ are sine functions.

Those eigenvalues of $K_3$ are $2 - \sqrt{2}$ and 2 and $2 + \sqrt{2}$. Those add to 7, which is the trace of $K_3$. Multiply those eigenvalues to get the determinant of $K_3$.

**Solution** Multiplying $2 - \sqrt{2}$ times $2 + \sqrt{2}$ gives 4 - 2 = 2. Then multiplying by 2 gives 4. This is the determinant (and $2 - \sqrt{2}, 2 + \sqrt{2}$ are the eigenvalues) of 3 by 3 matrix $K_3$. 

14 The slope of a ramp function is a step function. The slope of a step function is a delta function. Suppose the ramp function is \( r(x) = -x \) for \( x \leq 0 \) and \( r(x) = x \) for \( x \geq 0 \) (so \( r(x) = |x| \)). Find \( dr/dx \) and \( d^2r/dx^2 \).

Solution For the down-up ramp function \( r(x) = |x| \) = absolute value of \( x \), the derivatives are \( dr/dx = -1 \) then +1 and \( d^2r/dx^2 = 2\delta(x) \) because \( dr/dx \) jumps by 2 at \( x = 0 \).

15 Find the second differences \( y_{n+1} - 2y_n + y_{n-1} \) of these infinitely long vectors \( y \):

- **Constant** \((\ldots, 1, 1, 1, 1, \ldots)\)
- **Linear** \((\ldots,-1, 0, 1, 2, 3, \ldots)\)
- **Quadratic** \((\ldots,1, 0, 1, 4, 9, \ldots)\)
- **Cubic** \((\ldots,-1, 0, 1, 8, 27, \ldots)\)
- **Ramp** \((\ldots, 0, 0, 0, 1, 2, \ldots)\)
- **Exponential** \((\ldots,e^{-i\omega}, e^0, e^{i\omega}, e^{2i\omega}, \ldots)\).

It is amazing how closely those second differences follow second derivatives for \( y(x) = 1, x, x^2, x^3, \max(x, 0) \), and \( e^{i\omega x} \). From \( e^{i\omega x} \) we also get \( \cos \omega x \) and \( \sin \omega x \).

Solution The six second differences are: zero vector, zero vector, constant vector of 2’s, 6 times the linear vector, (for ramp: delta vector with \( \delta_0 = 1 \)), \( e^{i\omega} - 2 + e^{-i\omega} = 2 \cos \omega - 2 \) times the exponential vector. **Like 2nd derivatives** of 1, \( x \), \( x^2 \), \( x^3 \), ramp, \( e^{i\omega x} \).

**Problem Set 7.4, page 422**

1 What solution to Laplace’s equation completes “degree 3” in the table of pairs of solutions? We have one solution \( u = x^3 - 3xy^2 \), and we need another solution.

Solution Start with \( s = -y^3 \). Then \( s_{yy} = -6y \), and therefore we need \( s_{xx} = 6y \). Integrating twice with respect to \( x \) gives \( 3y^2x \). Therefore the second function is \( s(x, y) = -y^3 + 3x^2y \).

2 What are the two solutions of degree 4, the real and imaginary parts of \((x + iy)^4\)? Check \( u_{xx} + u_{yy} = 0 \) for both solutions.

Solution Expanding \((x + iy)^4\) gives

\[
(x + iy)^4 = x^4 - 6x^2y^2 + y^4 + (4x^3y - 4xy^3)i
\]

Therefore the two solutions would be:

\[
u(x, y) = x^4 - 6x^2y^2 + y^4 \quad \text{and} \quad s(x, y) = 4x^3y - 4xy^3
\]

Checking the first solution:

\[
\frac{\partial^2(x^4 - 6x^2y^2 + y^4)}{\partial x^2} + \frac{\partial^2(x^4 - 6x^2y^2 + y^4)}{\partial y^2} = (12x^2 - 12y^2) + (-12x^2 + 12y^2) = 0
\]

Checking the second solution:

\[
\frac{\partial^2(4x^3y - 4xy^3)}{\partial x^2} + \frac{\partial^2(4x^3y - 4xy^3)}{\partial y^2} = (24xy - 0) + (0 - 24xy) = 0
\]
What is the second \( x \)-derivative of \((x + iy)^n\)? What is the second \( y \)-derivative? Those cancel in \(u_{xx} + u_{yy}\) because \(i^2 = -1\).

**Solution** The second \( x \)-derivative of \((x + iy)^n\) is:

\[
\frac{\partial^2 (x + iy)^n}{\partial x^2} = n(n - 1)(x + iy)^{n-2}
\]

The second \( y \)-derivative of \((x + iy)^n\) cancels that because

\[
\frac{\partial^2 (x + iy)^n}{\partial y^2} = i \cdot i \cdot n(n - 1)(x + iy)^{n-2} = -n(n - 1)(x + iy)^{n-2}
\]

For the solved \(2 \times 2\) example inside a \(4 \times 4\) square grid, write the four equations (9) at the four interior nodes. Move the known boundary values 0 and 4 to the right hand sides of the equations. You should see \(K2D\) on the left side multiplying the correct solution \(U = (U_{11}, U_{12}, U_{21}, U_{22}) = (1, 2, 2, 3)\).

**Solution** The equations at the interior node would be:

\[
\begin{align*}
4U_{1,1} - U_{2,1} - U_{0,1} - U_{1,2} - U_{1,0} &= 0 \\
4U_{1,2} - U_{2,2} - U_{0,2} - U_{1,1} &= 0 \\
4U_{2,1} - U_{3,1} - U_{1,1} - U_{2,2} - U_{2,0} &= 0 \\
4U_{2,2} - U_{3,2} - U_{1,2} - U_{2,3} - U_{2,1} &= 0 \\
\end{align*}
\]

Substituting the known boundary values leaves:

\[
\begin{align*}
4U_{1,1} - U_{2,1} - U_{1,2} &= 4 \\
4U_{1,2} - U_{2,2} - U_{1,1} &= 8 \\
4U_{2,1} - U_{1,1} - U_{2,2} &= 0 \\
4U_{2,2} - U_{1,2} - U_{2,1} &= 4 \\
\end{align*}
\]

Writing this in matrix form gives:

\[
\begin{bmatrix}
4 & -1 & 0 & -1 \\
-1 & 4 & -1 & 0 \\
0 & -1 & 4 & -1 \\
-1 & 0 & -1 & 4 \\
\end{bmatrix}
\begin{bmatrix}
U_{1,1} \\
U_{1,2} \\
U_{2,1} \\
U_{2,2} \\
\end{bmatrix}
\begin{bmatrix}
4 \\
8 \\
0 \\
4 \\
\end{bmatrix}
\text{ and}
\begin{bmatrix}
U_{1,1} \\
U_{1,2} \\
U_{2,1} \\
U_{2,2} \\
\end{bmatrix}
\begin{bmatrix}
2 \\
3 \\
1 \\
2 \\
\end{bmatrix}
\]

Suppose the boundary values on the \(4 \times 4\) grid change to \(U = 0\) on three sides and \(U = 8\) on the fourth side. Find the four inside values so that each one is the average of its neighbors.

**Solution** The values at the 16 nodes will be

\[
\begin{array}{cccc}
0 & 0 & 0 & 0 \\
0 & \frac{1}{2} & \frac{1}{2} & 0 \\
0 & \frac{3}{2} & \frac{3}{2} & 0 \\
0/4 & 4 & 4 & 0/4 \\
\end{array}
\]

Notice that the corner boundary values **do not enter** the 5-point equations around interior points. Every interior value must be the average of its four neighbors. By symmetry the two middle columns must be the same.
6 (MATLAB) Find the inverse \((K2D)^{-1}\) of the 4 by 4 matrix displayed for the square grid.

*Solution* The circulant matrix \(K2D\) on page 422 has a circulant inverse:

\[
(K2D)^{-1} = \frac{1}{24} \begin{bmatrix} 7 & 2 & 1 & 2 \\ 2 & 7 & 2 & 1 \\ 1 & 2 & 7 & 2 \\ 2 & 1 & 2 & 7 \end{bmatrix}.
\]

7 Solve this Poisson finite difference equation (right side \(\neq 0\)) for the inside values \(U_{11}, U_{12}, U_{21}, U_{22}\). All boundary values like \(U_{10}\) and \(U_{13}\) are zero. The boundary has \(i\) or \(j\) equal to 0 or 3, the interior has \(i\) and \(j\) equal to 1 or 2:

\[4U_{ij} - U_{i-1,j} - U_{i+1,j} - U_{i,j-1} - U_{i,j+1} = 1\]

at four inside points.

*Solution* The interior solution to the Poisson equation (on this small grid) is

\[
\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}
\]

On a larger grid \(U_{ij}\) will not be constant in the interior.

8 A 5 \times 5 grid has a 3 by 3 interior grid: 9 unknown values \(U_{11}\) to \(U_{33}\). Create the 9 \times 9 difference matrix \(K2D\).

*Solution* Order the points by rows to get \(U_{11}, U_{12}, U_{13}, U_{21}, U_{22}, U_{23}, U_{31}, U_{32}, U_{33}\). Then \(K2D\) is symmetric with 3 by 3 blocks:

\[
K2D = \begin{bmatrix} A & -I & 0 \\ -I & A & -I \\ 0 & -I & A \end{bmatrix}
\]

\[
A = \begin{bmatrix} 4 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 4 \end{bmatrix}
\]

9 Use eig(\(K2D\)) to find the nine eigenvalues of \(K2D\) in Problem 8. Those eigenvalues will be positive! The matrix \(K2D\) is symmetric positive definite.

*Solution* eig(\(K2D\)) in Problem 8 produces 9 eigenvalues between 0 and 4:

\[
1.1716 \quad 2.5828 \quad 2.5828 \quad 4.0 \quad 4.0 \quad 4.0 \quad 5.4142 \quad 5.4142 \quad 6.8284
\]

10 If \(u(x)\) solves \(u_{xx} = 0\) and \(v(y)\) solves \(v_{yy} = 0\), verify that \(u(x)v(y)\) solves Laplace’s equation. Why is this only a 4-dimensional space of solutions? Separation of variables does not give all solutions—only the solutions with separable boundary conditions.

*Solution* If \(\frac{\partial^2 u}{\partial x^2} = 0\) and \(\frac{\partial^2 v}{\partial y^2} = 0\) then

\[
\frac{\partial^2 u(x)v(y)}{\partial x^2} + \frac{\partial^2 u(x)v(y)}{\partial y^2} = v(y)\frac{\partial^2 u(x)}{\partial x^2} + u(x)\frac{\partial^2 v(y)}{\partial y^2}
= v \cdot 0 + u \cdot 0 = 0
\]

Therefore \(u(x)v(y)\) solves Laplace’s equation. But the only solutions found this way are \(u(x)v(y) = (A + Bx)(C + Dy)\).
Problem Set 7.5, page 428

Problems 1 – 5 are about complete graphs. Every pair of nodes has an edge.

1 With \( n = 5 \) nodes and all edges, find the diagonal entries of \( A^T A \) (the degrees of the nodes). All the off-diagonal entries of \( A^T A \) are \(-1\). Show the reduced matrix \( R \) without row 5 and column 5. Node 5 is “grounded” and \( v_5 = 0 \).

Solution The complete graph (all edges included) has no zeros in \( A^T A \):

\[
A^T A = \begin{bmatrix}
4 & -1 & -1 & -1 & -1 \\
-1 & 4 & -1 & -1 & -1 \\
-1 & -1 & 4 & -1 & -1 \\
-1 & -1 & -1 & 4 & -1 \\
-1 & -1 & -1 & -1 & 4
\end{bmatrix}
\]

The grounded matrix would be

\[
(A^T A)_{\text{grounded}} = \begin{bmatrix}
4 & -1 & -1 & -1 & -1 \\
-1 & 4 & -1 & -1 & -1 \\
-1 & -1 & 4 & -1 & -1 \\
-1 & -1 & -1 & 4 & -1 \\
-1 & -1 & -1 & -1 & 4
\end{bmatrix}
\]

2 Show that the trace of \( A^T A \) (sum down the diagonal = sum of eigenvalues) is \( n^2 - n \). What is the trace of the reduced (and invertible) matrix \( R \) of size \( n - 1 \)?

Solution \( A^T A \) is \( n \) by \( n \) and each diagonal entry is \( n - 1 \). Therefore the trace is \( n(n - 1) = n^2 - n \). The reduced matrix \( R \) has \( n - 1 \) diagonal entries, each still equal to \( n - 1 \). Therefore the trace is \((n - 1)(n - 1) = n^2 - 2n + 1 \).

3 For \( n = 4 \), write the 3 by 3 matrix \( R = (A_{\text{reduced}})^T (A_{\text{reduced}}) \). Show that \( RR^{-1} = I \) when \( R^{-1} \) has all entries \( \frac{1}{2} \) off the diagonal and \( \frac{4}{3} \) on the diagonal.

Solution Reduced matrix \( R = \begin{bmatrix}
3 & -1 & -1 \\
-1 & 3 & -1 \\
-1 & -1 & 3
\end{bmatrix} \)

\( R \) by its proposed inverse gives

\[
\begin{bmatrix}
3 & -1 & -1 \\
-1 & 3 & -1 \\
-1 & -1 & 3
\end{bmatrix}
\]

4 For every \( n \), the reduced matrix \( R \) of size \( n - 1 \) is invertible. Show that \( RR^{-1} = I \) when \( R^{-1} \) has all entries \( \frac{1}{n} \) off the diagonal and \( \frac{2}{n} \) on the diagonal.

Solution

\[
\frac{1}{n} \begin{bmatrix}
2 & 1 & 1 \\
1 & 2 & 1 \\
1 & 1 & 2
\end{bmatrix} = \frac{1}{4} \begin{bmatrix}
6 - 1 - 1 & 3 - 2 - 1 & 3 - 1 - 2 \\
-2 + 3 - 1 & -1 + 6 - 1 & -1 + 3 - 2 \\
-2 - 1 + 3 & -1 - 2 + 3 & -1 - 1 + 6
\end{bmatrix} = I.
\]

5 Write the 6 by 3 matrix \( M = A_{\text{reduced}} \) when \( n = 4 \). The equation \( Mv = b \) is to be solved by least squares. The vector \( b \) is like scores in 6 games between 4 teams (team 4 always scores zero; it is grounded). Knowing the inverse of \( R = M^T M \), what is the least squares ranking \( v_1 \) for team 1 from solving \( M^T M \tilde{v} = M^T b \) ?

Solution Remove column 4 of \( A \) when node 4 is grounded (\( x_4 = 0 \)).
7.5. Networks and the Graph Laplacian

\[ M = \begin{bmatrix}
-1 & 1 & 0 \\
-1 & 0 & 1 \\
0 & -1 & 1 \\
-1 & 0 & 0 \\
0 & -1 & 0 \\
0 & 0 & -1 \\
\end{bmatrix} \]

has independent columns

The least squares solution \( \hat{v} \) to \( Mv = b \) comes from \( M^TM\hat{v} = M^Tb \). This \( \hat{v} \) gives the predicted point spreads when all teams play all other teams. The first component \( \hat{v}_1 \) would come from the first row of \( (M^TM)^{-1} \) multiplying by \( M^Tb \). Note that

\[ M^TM = \begin{bmatrix}
3 & -1 & -1 \\
-1 & 3 & -1 \\
-1 & -1 & 3 \\
\end{bmatrix} \quad \text{and} \quad (M^TM)^{-1} = \frac{1}{4} \begin{bmatrix}
2 & 1 & 1 \\
1 & 2 & 1 \\
1 & 1 & 2 \\
\end{bmatrix} \]

6 For the tree graph with 4 nodes, \( A^T A \) is in equation (1). What is the 3 by 3 matrix \( R = (A^T A)_{\text{reduced}} \)? How do we know it is positive definite?

Solution The reduced form of \( A^T A \) removes row 4 and column 4:

\[ A^T A = \begin{bmatrix}
1 & -1 & 0 \\
-1 & 2 & 0 \\
0 & -1 & 2 \\
0 & 0 & -1 \\
\end{bmatrix} \text{ reduces to invertible } \begin{bmatrix}
1 & -1 & 0 \\
-1 & 2 & -1 \\
0 & -1 & 2 \\
\end{bmatrix} \]

The first is positive semidefinite (A has dependent columns). The second is positive definite (the reduced A has 3 independent columns).

7 (a) If you are given the matrix \( A \), how could you reconstruct the graph?

Solution Each row of \( A \) tells you an edge in the graph.

(b) If you are given \( L = A^T A \), how could you reconstruct the graph (no arrows)?

Solution Each nonzero off the main diagonal of \( A^T A \) tells you an edge.

(c) If you are given \( K = A^T CA \), how could you reconstruct the weighted graph?

Solution Each nonzero off the main diagonal tells you the weight of that edge.

8 Find \( K = A^T CA \) for a line of 3 resistors with conductances \( c_1 = 1, c_2 = 4, c_3 = 9 \). Write \( K_{\text{reduced}} \) and show that this matrix is positive definite.

Solution A circle of three resistors has 3 edges and 3 nodes:

\[ A^T CA = \begin{bmatrix}
-1 & 1 & 0 \\
0 & -1 & 1 \\
1 & 0 & -1 \\
\end{bmatrix} \begin{bmatrix}
4 \\
9 \\
\end{bmatrix} = \begin{bmatrix}
-1 & 0 & 1 \\
1 & -1 & 0 \\
0 & 1 & -1 \\
\end{bmatrix} \begin{bmatrix}
5 & -4 & -1 \\
-4 & 13 & -9 \\
-1 & -9 & 10 \\
\end{bmatrix} \]

\( (A^T CA)_{\text{reduced}} = \begin{bmatrix}
-1 & 1 & 0 \\
0 & -1 & 1 \\
\end{bmatrix} \begin{bmatrix}
4 \\
9 \\
\end{bmatrix} = \begin{bmatrix}
-1 & 0 \\
1 & -1 \\
0 & 1 \\
\end{bmatrix} \begin{bmatrix}
5 & -4 \\
-4 & 13 \\
\end{bmatrix} \]

The determinant tests \( 5 > 0 \) and \( (5)(13) > 4^2 \) are passed.
9 A 3 by 3 square grid has \( n = 9 \) nodes and \( m = 12 \) edges. Number nodes by rows.
(a) How many nonzeros among the 81 entries of \( L = A^T A \)?

**Solution** The 9 nodes ordered by rows have 2, 3, 2, 3, 4, 3, 2, 3, 2 neighbors around them. Those add to 24 nonzeros off the diagonal. The 9 diagonal entries make 33 nonzeros out of \( 9^2 = 81 \) entries in \( L = A^T A \).
(b) Write down the 9 diagonal entries in the degree matrix \( D \): they are not all 4.

**Solution** Those 9 numbers are the degrees of the 9 nodes (= diagonal entries in \( A^T A \)).
(c) Why does the middle row of \( L = D - W \) have four −1’s? Notice \( L = K 2D \)!

**Solution** The middle node in the grid has 4 neighbors.

10 Suppose all conductances in equation (5) are equal to \( c \). Solve equation (6) for the voltages \( v_2 \) and \( v_3 \) and find the current \( I \) flowing out of node 1 (and into the ground at node 4). What is the “system conductance” \( I/V \) from node 1 to node 4?

This overall conductance \( I/V \) should be larger than the individual conductances \( c \).

**Solution** The reduced equation (6) with conductances \( c \) is

\[
\begin{bmatrix}
3c & -c \\
-c & 2c
\end{bmatrix}
\begin{bmatrix}
v_2 \\
v_3
\end{bmatrix}
= \begin{bmatrix}
cV \\
cV
\end{bmatrix}
\text{and}
\begin{bmatrix}
v_2 \\
v_3
\end{bmatrix}
= \begin{bmatrix}
0.6V \\
0.8V
\end{bmatrix}.
\]

Then the flows on the five edges in Figure 7.6 use \( A \) in equation (2). Remember the minus sign:

\[
-c A v = \begin{bmatrix}
-1 & 1 & 0 & 0 \\
-1 & 0 & 1 & 0 \\
0 & -1 & 1 & 0 \\
-1 & 0 & 0 & 1 \\
0 & -1 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
V \\
0.6V \\
0.8V \\
0
\end{bmatrix} = cV
\begin{bmatrix}
0.4 \\
0.2 \\
-0.2 \\
1.0 \\
0.6
\end{bmatrix}.
\]

The total flow (on edges 1+2+4 out of node 1, or on edges 3+4 into the grounded node 4, is \( I = 1.6eV \). The overall system conductance is 1.6\( c \), greater than the individual conductance \( c \) on each edge.

11 The multiplication \( A^T A \) can be columns of \( A^T \) times rows of \( A \). For the tree with \( m = 3 \) edges and \( n = 4 \) nodes, each (column times row) is \((4 \times 1)(1 \times 4) = 4 \times 4\). Write down those three column-times-row matrices and add to get \( L = A^T A \).

**Solution** Suppose the 3 tree edges go out of node 1 to nodes 2, 3, 4. (The problem allows to choose other trees, including a line of 4 nodes.) Then

\[
A = \begin{bmatrix}
-1 & 1 & 0 & 0 \\
-1 & 0 & 1 & 0 \\
-1 & 0 & 0 & 1
\end{bmatrix}
A^T A = \begin{bmatrix}
3 & -1 & -1 & -1 \\
-1 & 1 & 0 & 0 \\
-1 & 0 & 1 & 0 \\
-1 & 0 & 0 & 1
\end{bmatrix}
= \text{sum of (columns of } A^T \text{)}(\text{rows of } A)
\]

\[
= \begin{bmatrix}
-1 \\
1 \\
0 \\
0
\end{bmatrix}
\begin{bmatrix}
-1 & 1 & 0 & 0
\end{bmatrix} + \begin{bmatrix}
-1 \\
0 \\
1 \\
0
\end{bmatrix}
\begin{bmatrix}
-1 & 0 & 1 & 0
\end{bmatrix} + \begin{bmatrix}
-1 \\
0 \\
0 \\
1
\end{bmatrix}
\begin{bmatrix}
-1 & 0 & 0 & 1
\end{bmatrix}.
\]
12 A graph with two separate 3-node trees is not connected. Write its 6 by 4 incidence matrix \( A \). Find two solutions to \( Av = 0 \), not just one solution \( v = (1, 1, 1, 1, 1, 1) \). To reduce \( A^T A \) we must ground two nodes and remove two rows and columns.

**Solution** The incidence matrix for two 3-node trees is

\[
A = \begin{bmatrix}
A_{\text{tree}} & 0 \\
0 & A_{\text{tree}}
\end{bmatrix}
\] with \( A_{\text{tree}} = \begin{bmatrix} 1 & 1 & 0 \\
-1 & 0 & 1
\end{bmatrix} \) (for example)

The columns of \( A_{\text{tree}} \) add to zero so we have 2 independent solutions to \( Av = 0 \):

\[
v = \begin{bmatrix} 1 \\
1 \\
0 \\
0
\end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 \\
0 \\
1 \\
1
\end{bmatrix}
\]

come from \( A_{\text{tree}} \begin{bmatrix} 1 \\
1 \\
1 \\
1
\end{bmatrix} = \begin{bmatrix} 0 \\
0 \\
0 \\
0
\end{bmatrix} \).

13 “Element matrices” from column times row appear in the finite element method. Include the numbers \( c_1, c_2, c_3 \) in the element matrices \( K_1, K_1, K_3 \).

\[ K_i = (\text{row } i \text{ of } A^T \ (c_i)) \quad (\text{row } i \text{ of } A) \quad K = A^T CA = K_1 + K_2 + K_3. \]

Write the element matrices that add to \( A^T A \) in (1) for the 4-node line graph.

\[
A^T A = \begin{bmatrix}
K_1 \\
K_2 \\
K_3
\end{bmatrix}
\]

as indicated by the non-zero entries of \( K_1 + K_2 + K_3 \) from edges 1, 2, and 3.

**Solution** The three “element matrices” for the three edges come from multiplying the three columns of \( A^T \) by the three rows of \( A \). Then \( A^T A \) equals

\[
= \begin{bmatrix}
-1 & 1 \\
1 & 0 \\
0 & 0
\end{bmatrix} \begin{bmatrix} 0 \\
1 \\
0
\end{bmatrix} + \begin{bmatrix} 1 \\
-1 \\
0
\end{bmatrix} \begin{bmatrix} 0 & -1 & 1 \\
0 & 1 & 0
\end{bmatrix} + \begin{bmatrix} 0 \\
0 \\
1
\end{bmatrix} \begin{bmatrix} 0 & 0 & -1 \\
1 & 0 & 0
\end{bmatrix}.
\]

When the diagonal matrix \( C \) is included, those are multiplied by \( c_1, c_2, \) and \( c_3 \). Those products produce 2 by 2 blocks of nonzeros in \( 4 \times 4 \) matrices:

\[
K_1 = c_1 \begin{bmatrix}
1 & -1 \\
-1 & 1
\end{bmatrix} \quad K_2 = c_2 \begin{bmatrix} 1 \\
-1
\end{bmatrix} \quad K_3 = c_3 \begin{bmatrix} 1 & -1 \\
-1 & 1
\end{bmatrix}
\]

Then \( A^T CA = K_1 + K_2 + K_3 \). This “assembly” of the element stiffness matrices just requires placing the nonzeros correctly into the final matrix \( A^T CA \).

14 An \( n \) by \( n \) grid has \( n^2 \) nodes. How many edges in this graph? How many interior nodes? How many nonzeros in \( A \) and in \( L = A^T A \)? There are no zeros in \( L^{-1} \)!

**Solution** An \( n \) by \( n \) grid has \( n \) horizontal rows (\( n-1 \) edges on each row) and \( n \) vertical columns (\( n-1 \) edges down each column). Altogether \( 2n(n-1) \) edges. There are
Every edge produces 2 nonzeros (−1 and +1) in \( A \). Then \( A \) has \( 4n(n - 1) \) nonzeros. The matrix \( A^T A \) has size \( n^2 \) with \( n^2 \) diagonal nonzeros—and off the diagonal of \( A^T A \) there are two −1’s for each edge: altogether \( n^2 + 4n(n - 1) = 5n^2 - 4n \) nonzeros out of \( n^2 \) entries. For \( n = 2 \), this means 12 nonzeros in a 4 by 4 matrix.

15 When only \( e = C^{-1} w \) is eliminated from the 3-step framework, equation (??) shows

\[
\begin{bmatrix}
C^{-1} & A \\
A^T & 0
\end{bmatrix}
\begin{bmatrix}
w \\
v
\end{bmatrix} =
\begin{bmatrix}
b \\
f
\end{bmatrix}.
\]

Multiply the first block row by \( A^T C \) and subtract from the second block row:

\[
\text{After block elimination}
\begin{bmatrix}
C^{-1} & A \\
0 & -A^T CA
\end{bmatrix}
\begin{bmatrix}
w \\
v
\end{bmatrix} =
\begin{bmatrix}
b \\
f - A^T C b
\end{bmatrix}.
\]

After \( m \) positive pivots from \( C^{-1} \), why does this matrix have negative pivots? The two-field problem for \( w \) and \( v \) is finding a saddle point, not a minimum.

Solution The three equations \( e = b - A v \) and \( w = C e \) and \( A^T w = f \) reduce to two equations when \( e \) is replaced by \( C^{-1} w \):

\[
\begin{align*}
C^{-1} w &= b - A v \\
A^T w &= f
\end{align*}
\]

become

\[
\begin{bmatrix}
C^{-1} & A \\
A^T & 0
\end{bmatrix}
\begin{bmatrix}
v \\
w
\end{bmatrix} =
\begin{bmatrix}
b \\
f
\end{bmatrix}.
\]

Multiply the first equation by \( A^T C \) to get \( A^T w = A^T C b - A^T C A v \). Subtract from the second equation \( A^T w = f \), to eliminate \( w \):

\[
A^T C b - A^T C A v = f.
\]

This gives the second row of the block matrix after elimination:

\[
\begin{bmatrix}
C^{-1} & A \\
0 & -A^T CA
\end{bmatrix}
\begin{bmatrix}
v \\
w
\end{bmatrix} =
\begin{bmatrix}
b \\
f - A^T C b
\end{bmatrix}.
\]

The pivots of that matrix on the left side start with \( 1/c_1, 1/c_2, \ldots, 1/c_m \). Then we get the \( n \) pivots of \( -A^T C A \) which are \textbf{negative}, because this matrix is negative definite.

Altogether we are finding a saddle point \((v, w)\) of the energy (quadratic function). The derivative of that quadratic gives our linear equations. The block matrix in those equations has \( m \) positive eigenvalues and \( n \) negative eigenvalues.

16 The least squares equation \( A^T A v = A^T b \) comes from the projection equation \( A^T e = 0 \) for the error \( e = b - A v \). Write those two equations in the symmetric saddle point form of Problem 7 (with \( f = 0 \)).

In this case \( w = e \) because the weighting matrix is \( C = I \).

Solution Ordinary least squares for \( A v = b \) separates the data vector \( b \) in two perpendicular parts:

\[
b = (A\hat{v}) + (b - A\hat{v}) = \text{projection of } b + \text{error in } b.
\]

The error \( e = b - A v \) satisfies \( A^T e = A^T b - A^T A v = 0 \) (which means that \( A^T A v = A^T b \), the key equation). That equation \( d^T e = 0 \) is Kirchhoff’s Current Law for flows in
7.5. Networks and the Graph Laplacian

It is a candidate for the “most important equation in applied mathematics”—the conservation equation or continuity equation “flow in = flow out.”

In the form of Problem 15 (with \( C = I \)) the equations are

\[
\begin{bmatrix}
I & A \\
A^T & 0
\end{bmatrix}
\begin{bmatrix}
e \\
v
\end{bmatrix}
=
\begin{bmatrix}
b \\
0
\end{bmatrix}
or
\begin{bmatrix}
e + Av = b \\
A^Te = 0
\end{bmatrix}.
\]

17 Find the three eigenvalues and three pivots and the determinant of this saddle point matrix with \( C = I \). One eigenvalue is negative because \( A \) has one column:

\[
m = 2, n = 1 \quad \begin{bmatrix}
C^{-1} & A \\
A^T & 0
\end{bmatrix}
= \begin{bmatrix}
1 & 0 & -1 \\
0 & 1 & 1 \\
-1 & 1 & 0
\end{bmatrix}.
\]

**Solution** The eigenvalues come from \( \det(M - \lambda I) = 0 \):

\[
\begin{bmatrix}
1 - \lambda & 0 & -1 \\
0 & 1 - \lambda & 1 \\
-1 & 1 & -\lambda
\end{bmatrix} = -\lambda(1 - \lambda)^2 - 2(1 - \lambda) = 0.
\]

Then \((1 - \lambda)(\lambda^2 - \lambda - 2) = 0\) and \((1 - \lambda)(\lambda - 2)(\lambda + 1) = 0\) and the eigenvalues are \( \lambda = 1, 2, -1 \). Check the sum \( 1 + 2 - 1 = 2 \) equal to the trace (sum down the main diagonal \( 1 + 1 + 0 = 2 \)).

The determinant is the product \( \lambda_1 \lambda_2 \lambda_3 = (1)(2)(-1) = -2 \). Notice \( m = 2 \) positive \( \lambda \)'s and \( n = 1 \) negative eigenvalue.

Elimination finds the three pivots (which also multiply to give \( \det M = -2 \)):

\[
\begin{bmatrix}
1 & 0 & -1 \\
0 & 1 & 1 \\
-1 & 1 & 0
\end{bmatrix}
\rightarrow
\begin{bmatrix}
1 & 0 & -1 \\
0 & 1 & 1 \\
0 & 1 & -1
\end{bmatrix}
\rightarrow
\begin{bmatrix}
1 & 0 & -1 \\
0 & 1 & 1 \\
0 & 0 & -2
\end{bmatrix}.
\]