1. Find all least-squares solutions $\hat{x}$ of each of the following systems of equations $Ax = b$, and compute the projection b_V of b onto $V = \text{Col}(A)$ and the minimum value of $\|A\hat{x} - b\|$.

a) $\begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 2 \end{pmatrix} x = \begin{pmatrix} 1 \\ 4 \\ 3 \end{pmatrix}$
   b) $\begin{pmatrix} 1 & 2 & 1 \\ -1 & 1 & 0 \\ 2 & 2 & -1 \\ 4 & 3 & 0 \end{pmatrix} x = \begin{pmatrix} -1 \\ -1 \\ -1 \\ 7 \end{pmatrix}$

c) $\begin{pmatrix} 2 & 2 & -1 \\ -4 & -5 & 5 \\ 6 & 1 & 12 \end{pmatrix} x = \begin{pmatrix} -6 \\ -24 \\ -3 \end{pmatrix}$
   d) $\begin{pmatrix} 3 & 0 \\ 1 & -2 \\ 3 & 1 \end{pmatrix} x = \begin{pmatrix} 9 \\ 7 \\ 7 \end{pmatrix}$

2. Consider the data points

$$p_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \qquad p_2 = \begin{pmatrix} 2 \\ -1 \end{pmatrix} \qquad p_3 = \begin{pmatrix} 3 \\ 0 \end{pmatrix} \qquad p_4 = \begin{pmatrix} 4 \\ 5 \end{pmatrix}.$$

a) Find the best-fit line $y = Cx + D$ through these four points, and draw it on the grid below.
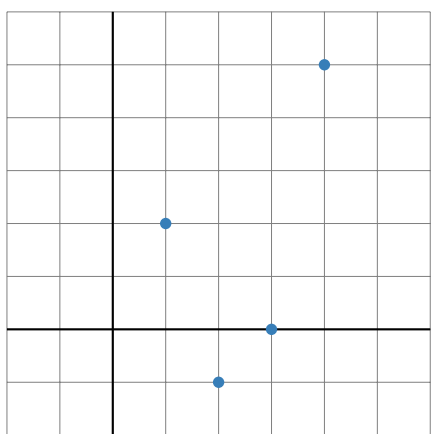
b) For each data point $p_i = (a_i, b_i)$, draw the error bar from $(a_i, y(a_i))$ to (a_i, b_i).

c) What is the minimum value of $\sum_{i=1}^{4}(b_i - y(a_i))^2$? How do you know?

d) Verify that the vector

$$\big(2 - y(1),\ -1 - y(2),\ 0 - y(3),\ 5 - y(4)\big)$$

is orthogonal to $(1, 2, 3, 4)$ and $(1, 1, 1, 1)$, and explain why this is necessary.

e) Find the best-fit *horizontal* line $y = D$ through these four points. Verify that D is the average of the y-values of the data points p_1, p_2, p_3, p_4.

3. Consider the data points

$$p_1 = \begin{pmatrix} -2 \\ 3 \end{pmatrix} \qquad p_2 = \begin{pmatrix} -1 \\ -1 \end{pmatrix} \qquad p_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad p_4 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

a) Find the best-fit parabola $y = Cx^2 + Dx + E$ through these four points, and draw it on the grid below.
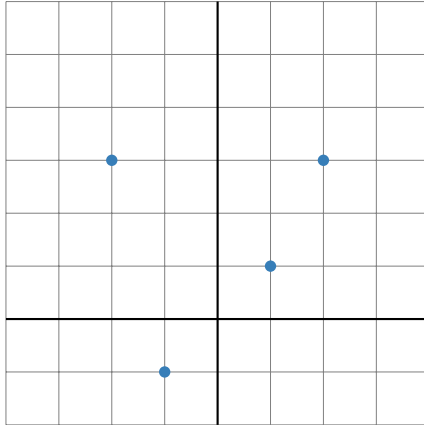
b) For each data point $p_i = (a_i, b_i)$, draw the error bar from $(a_i, y(a_i))$ to (a_i, b_i).

c) What is the minimum value of $\sum_{i=1}^4 (b_i - y(a_i))^2$? How do you know?

4. Consider the following data points:

$$p_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \qquad p_2 = \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix} \qquad p_3 = \begin{pmatrix} -1 \\ 1 \\ 3 \end{pmatrix} \qquad p_4 = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}.$$

a) Find the best-fit plane $z = Cx + Dy + E$ through these four points.

b) Interpret the minimized quantity in the situation of this problem.

5. Consider the data points $p_1, \ldots, p_8$:

$$\begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ 1.5 \end{pmatrix}, \begin{pmatrix} 2.5 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -.5 \\ -1 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \end{pmatrix}, \begin{pmatrix} -2.5 \\ 2 \end{pmatrix}, \begin{pmatrix} -1.5 \\ 3.5 \end{pmatrix}.$$

a) Find the best-fit *ellipse*

$$x^2 + By^2 + Cxy + Dx + Ey + F = 0$$

through these data points.

b) Interpret the minimized quantity in the situation of this problem.

[**Hint:** you can't see it on the graph above, but you can see it on this demo.]

In this problem, I recommend using SymPy (in the Sage cell on the course webpage) or another computer algebra system to do the computations. To solve a normal equation $A^T A x = A^T (1, 2, 3)$, you would use something like

```
(A.T*A).solve(A.T*Matrix([1,2,3]))
```

Remark: Carl Friedrich Gauss (1777–1865), arguably the greatest mathematician since antiquity, kept food on the table by doing astronomical calculations. He invented much of the linear algebra you are learning in order to compute the trajectories of celestial bodies. Essentially performing the calculations in this problem, he correctly predicted the (elliptical) orbit of the asteroid Ceres as it passed behind the sun in 1801.

6. Suppose that $\hat{x}$ is a vector such that $A\hat{x} = (1, 1, -1, -1)$. Explain why $\hat{x}$ is not a least-squares solution of $Ax = (1, 1, 1, 1)$.

7. Decide if each statement is true or false, and explain why.

a) A least-squares solution $\hat{x}$ of $Ax = b$ is a solution of $A\hat{x} = b_V$ for $V = \text{Col}(A)$.

b) Any solution of $A^T A \hat{x} = A^T b$ is a least-squares solution of $Ax = b$.

c) If A has full column rank, then $Ax = b$ has exactly one least-squares solution for every b.

d) If $Ax = b$ has at least one least-squares solution for every b, then A has full row rank.

8. For each set of vectors, decide if they are orthogonal, orthonormal, or neither; then compute $Q^T Q$, where Q is the matrix with the vectors as columns.

a) $\left\{ \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \right\}$

b) $\left\{ \dfrac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \dfrac{1}{2} \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, \dfrac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} \right\}$

c) $\left\{ \dfrac{1}{3} \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \dfrac{1}{3} \begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix} \right\}$

d) $\left\{ \dfrac{1}{3} \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \dfrac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \dfrac{1}{\sqrt{18}} \begin{pmatrix} -4 \\ 1 \\ 1 \end{pmatrix} \right\}$

9. The following subspaces V are given as the span of an *orthogonal* set of vectors. For each subspace V and vector b, compute the orthogonal projection b_V using the *projection formula*, and compute the projection matrix P_V using the *outer product formula*.

a) $V = \text{Span}\left\{ \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} \right\}$ $\qquad\qquad b = \begin{pmatrix} 3 \\ 8 \\ 1 \end{pmatrix}$

b) $V = \text{Span}\left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix} \right\}$ $\qquad\qquad b = \begin{pmatrix} 0 \\ 3 \\ 6 \end{pmatrix}$

c) $V = \text{Span}\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} \right\}$ $\qquad b = \begin{pmatrix} 4 \\ 2 \\ -4 \\ 2 \end{pmatrix}$

d) $V = \text{Span}\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$ $\qquad\qquad b = \begin{pmatrix} 7 \\ -3 \\ 2 \end{pmatrix}$

e) $V = \text{Span}\left\{ \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -4 \\ 1 \\ 1 \end{pmatrix} \right\}$ $\qquad b = \begin{pmatrix} 9 \\ -2 \\ 3 \end{pmatrix}$

10. For each subspace V of Problem 9, scale the spanning vectors to find an *orthonormal* basis of V, and (re)compute the projection matrix P_V using the formula $P_V = QQ^T$. (Your answers should be exact, in terms of square roots.)

11. Suppose that $\{u_1, u_2, \ldots, u_n\}$ is an orthonormal basis of $\mathbf{R}^n$. Use the outer product formula to explain why
$$I_n = u_1 u_1^T + u_2 u_2^T + \cdots + u_n u_n^T.$$

12. If Q has orthonormal columns, what is the least-squares solution of $Qx = b$?

13. Give an example of each of the following, or explain why no such example exists.

a) A matrix Q with orthonormal columns, but $QQ^T \neq I_n$.

b) Two nonzero orthogonal vectors that are linearly dependent.

c) An orthonormal basis for the plane $x + y + z = 0$.

14. Use the Gram–Schmidt process to find orthogonal bases of the following subspaces.

a) Span $\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \right\}$　　**b)** Span $\left\{ \begin{pmatrix} 1 \\ -1 \\ 2 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ -3 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 4 \\ -1 \end{pmatrix} \right\}$

c) Span $\left\{ \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ -5 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 5 \\ 12 \end{pmatrix} \right\}$　　**d)** Nul $\begin{pmatrix} 1 & 2 & -1 & 4 \\ 3 & 6 & -3 & 12 \end{pmatrix}$

15. Consider the plane
$$V = \text{Span} \left\{ \begin{pmatrix} 1 \\ -1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ -1 \\ 1 \end{pmatrix} \right\}$$
and the vector
$$v = \begin{pmatrix} 4 \\ -1 \\ 3 \\ 3 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ -1 \\ 2 \\ 1 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \\ -1 \\ 1 \end{pmatrix} \in V.$$
Find all vectors **contained in** V that are orthogonal to v.

[**Hint:** apply Gram–Schmidt to a set containing v.]

16. Consider the subspace $V = \text{Col}(A)$, where
$$A = \begin{pmatrix} 1 & 1 & 2 \\ -1 & -3 & 2 \\ 2 & 2 & 4 \\ -2 & -1 & -1 \end{pmatrix}.$$
Find an orthonormal basis $\{u_1, u_2, u_3, u_4\}$ of $\mathbf{R}^4$ such that $\{u_1, u_2, u_3\}$ is a basis for V. Your answer should be exact, in terms of square roots.

17. For each of the following matrices A and vectors b, find the QR decomposition of A, and find the least-squares solution of $Ax = b$ by back-substitution in $R\hat{x} = Q^T b$. Your answers should be exact, in terms of square roots.

a) $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 2 \end{pmatrix}$, $b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$　　**b)** $A = \begin{pmatrix} 1 & 1 & 2 \\ -1 & -3 & 2 \\ 2 & 2 & 4 \\ -2 & -1 & -1 \end{pmatrix}$, $b = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$

18. In this problem, we use a *QR* decomposition to *quickly* compute the best-fit parabola with specified y-values at $x = -2, -1, 1, 2$, as in Problem 3.

  a) Compute the matrix A such that the least squares solution of $A(C, D, E) = (b_1, b_2, b_3, b_4)$ gives the coefficients of the parabola $y = Cx^2 + Dx + E$ that best fits the data points $(-2, b_1), (-1, b_2), (1, b_3), (2, b_4)$. (Presumably you computed this in Problem 3.)

  b) Find the *QR* decomposition of A.

  c) Find the best-fit parabola through the points $(-2, 3), (-1, -1), (1, 1), (2, 3)$ by back-substitution in $R\hat{x} = Q^T b$. You should get the same answer as in Problem 3.

  Note that we can now repeat part **c)** with new y-values in $O(n^2)$ time.

19. Recall that an *orthogonal* matrix is a square matrix with ortho*normal* columns.

  a) If Q is an orthogonal matrix, show that Q^{-1} is orthogonal.

  b) If Q_1 and Q_2 are orthogonal matrices of the same size, show that $Q_1 Q_2$ is orthogonal.

  c) If Q is orthogonal, show that $\det(Q) = \pm 1$.

20. Decide if each statement is true or false, and explain why.

  a) A matrix with orthogonal columns has full row rank.

  b) If $\{v_1, \ldots, v_n\}$ is a linearly independent set of vectors, then it is orthogonal.

  c) If $\{v_1, v_2\}$ is a basis for a plane V, then for any vector b,
  $$b_V = \frac{b \cdot v_1}{v_1 \cdot v_1} v_1 + \frac{b \cdot v_2}{v_2 \cdot v_2} v_2.$$

  d) If Q has orthonormal columns, then the distance from x to y equals the distance from Qx to Qy.

  e) If $A = QR$ is a *QR*-factorization of a matrix A, then the rows of Q form an orthonormal basis for Row(A).