Orthogonal bases, Gram–Schmidt and the QR decomposition

Ten problems on orthogonality worked by hand: dot products, lengths and angles, coordinates in an orthonormal basis by dot products, a Householder reflection and the properties of an orthogonal matrix, Gram–Schmidt in the plane and in three dimensions, the QR decomposition of a 3×2 and a 3×3 matrix with det A from R, least squares by solving Rx = Q^Tb, the projection QQ^T and the orthogonal complement, a dependent column that Gram–Schmidt turns into zero, and why nearly collinear features break the normal equations but not QR, with worked solutions and the mistakes of skipping the division by a·a, projecting onto a column instead of q, transposing R and taking QQ^T to be I.

BEFORE YOU START

An orthonormal basis turns linear algebra into dot products. Coordinates come from a dot product instead of a linear system, a projection is a sum of dot products, and the inverse of a matrix with orthonormal columns is its transpose. Gram–Schmidt builds such a basis from any independent set of vectors by subtracting, from each new vector, its projections onto the directions already found, and the bookkeeping of that process is the factorization $A = QR$. The mistakes are in the projections: a coefficient not divided by $a^\top a$, a projection onto an original column instead of a finished q, and a tall Q treated as if it were square. These ten problems build one orthonormal basis of $\mathbb{R}^3$ and use it for a reflection, a QR decomposition, a determinant and a least-squares fit, and end with the reason libraries solve least squares with QR rather than the normal equations.

- **Dot products.** $a^\top b = a_1 b_1 + \cdots + a_n b_n$, $\|a\| = \sqrt{a^\top a}$, and the angle θ between nonzero vectors has $\cos\theta = \dfrac{a^\top b}{\|a\|\,\|b\|}$. a and b are orthogonal when $a^\top b = 0$.
- **Orthonormal.** Vectors $q_1, \ldots, q_k$ are orthonormal when $q_i^\top q_j = 0$ for $i \neq j$ and $\|q_i\| = 1$. Stacked as the columns of Q, that says $Q^\top Q = I_k$.
- **Orthogonal matrix.** A square Q with $Q^\top Q = I$. Then $Q^{-1} = Q^\top$, $QQ^\top = I$ as well, $\|Qx\| = \|x\|$ for every x, and $\det Q = \pm 1$. A tall Q ($n \times k$ with $k < n$) still has $Q^\top Q = I_k$, but $QQ^\top$ is not I_n.
- **Projection onto a line.** The projection of b onto the line through $a \neq 0$ is $\dfrac{a^\top b}{a^\top a}\, a$. For a unit vector q it is $(q^\top b)\, q$.
- **Gram–Schmidt.** $u_1 = a_1$; then $u_j = a_j - \sum_{i<j} (q_i^\top a_j)\, q_i$ and $q_j = u_j / \|u_j\|$. Each u_j is a_j minus its projections onto the q's already built. If some $u_j = 0$, then a_j is a combination of the earlier columns.

- **QR.** For A with independent columns, $A = QR$, with Q the Gram–Schmidt vectors as columns and $R = Q^\top A$ upper triangular: $r_{ij} = q_i^\top a_j$ for $i < j$ and $r_{jj} = \|u_j\| > 0$.
- **Least squares.** The $\hat{x}$ that minimizes $\|Ax - b\|$ solves $A^\top A\hat{x} = A^\top b$; with $A = QR$ this becomes $R\hat{x} = Q^\top b$, and the projection of b onto $\mathrm{col}(A)$ is $QQ^\top b$.
- **Condition number.** $\kappa(A) = \sigma_{\max}/\sigma_{\min}$, the ratio of the largest and smallest singular values. Roughly, solving a system with condition number κ in double precision loses $\log_{10}\kappa$ of its 16 significant digits.

Problems

1. For $a = (1, 2, 2)^\top$, $b = (3, 0, 4)^\top$ and $c = (2, 1, -2)^\top$, find $a^\top b$, $\|a\|$, $\|b\|$ and the angle between a and b, and show that a and c are orthogonal. difficulty ·

2. $q_1 = \frac{1}{5}(3, 4)^\top$ and $q_2 = \frac{1}{5}(-4, 3)^\top$ form an orthonormal basis of $\mathbb{R}^2$. Write $x = (2, 11)^\top$ as $c_1 q_1 + c_2 q_2$ without solving a system. difficulty ·

3. Let $x = (1, 2, 2)^\top$ and $v = x - 3e_1 = (-2, 2, 2)^\top$. Compute the Householder reflection $H = I - \dfrac{2vv^\top}{v^\top v}$, and show that $Hx = 3e_1$, $H^\top H = I$, $H^{-1} = H$ and $\det H = -1$. difficulty ··

4. Apply Gram–Schmidt to $a_1 = (3, 4)^\top$ and $a_2 = (2, 11)^\top$. difficulty ··

5. Find the QR decomposition of $A = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 2 & 2 \end{bmatrix}$. difficulty ··

6. Extend Problem 5 with a third column: find the QR decomposition of $A = \begin{bmatrix} 1 & 4 & 5 \\ 2 & 5 & 1 \\ 2 & 2 & 1 \end{bmatrix}$, and use it to find $\det A$. difficulty ···

7. Use the QR decomposition of Problem 5 to find the least-squares solution of $Ax = b$ for $A = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 2 & 2 \end{bmatrix}$ and $b = (3, 3, 6)^\top$, and the residual. difficulty ··

8. For the tall Q of Problem 5, find $P = QQ^{\top}$ and show it is the projection onto the plane spanned by the columns of A. Find the orthogonal complement of that plane, and show $QQ^{\top} \neq I$ although $Q^{\top}Q = I$. difficulty ··

9. Run Gram–Schmidt on $a_1 = (1,2,2)^{\top}$, $a_2 = (4,5,2)^{\top}$, $a_3 = (6,9,6)^{\top}$. What happens at the third vector, and what does it say about $A = [a_1\ a_2\ a_3]$? difficulty ··

10. Two features are almost collinear: difficulty ···

$$A = \begin{bmatrix} 1 & 1 \\ \varepsilon & 0 \\ 0 & \varepsilon \end{bmatrix}, \qquad \varepsilon = 10^{-8}.$$

Show that $\kappa(A^{\top}A) = \kappa(A)^2$ in general, find both for this A, and explain why forming $A^{\top}A$ in double precision fails while QR on A does not.

Answers

1. $a^\top b = 11$, $\|a\| = 3$, $\|b\| = 5$, $\theta = \arccos \frac{11}{15} \approx 42.8^\circ$; $a^\top c = 0$
2. $x = 10q_1 + 5q_2$
3. $H = \frac{1}{3}\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{bmatrix}$, $Hx = 3e_1$, $H^\top H = I$, $H^{-1} = H$, $\det H = -1$
4. $q_1 = \frac{1}{5}(3,4)^\top$, $q_2 = \frac{1}{5}(-4,3)^\top$
5. $Q = \frac{1}{3}\begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 2 & -2 \end{bmatrix}$, $R = \begin{bmatrix} 3 & 6 \\ 0 & 3 \end{bmatrix}$
6. $R = \begin{bmatrix} 3 & 6 & 3 \\ 0 & 3 & 3 \\ 0 & 0 & 3 \end{bmatrix}$ and $\det A = \det Q \, \det R = (-1)(27) = -27$
7. $\hat{x} = \left(3, -\frac{1}{3}\right)^\top$, $A\hat{x} = \frac{1}{3}(5, 13, 16)^\top$, residual $r = \frac{2}{3}(2, -2, 1)^\top$ with $\|r\| = 2$
8. $P = \frac{1}{9}\begin{bmatrix} 5 & 4 & -2 \\ 4 & 5 & 2 \\ -2 & 2 & 8 \end{bmatrix}$; complement spanned by $(2, -2, 1)^\top$; $QQ^\top$ has trace 2, so it is not I_3
9. $u_3 = 0$: $a_3 = 2a_1 + a_2$, the columns are dependent, A has rank 2 and $\det A = 0$
10. $\kappa(A) \approx 1.41 \times 10^{8}$, $\kappa(A^\top A) \approx 2 \times 10^{16}$: the normal equations lose every digit, while QR works on A itself, with condition number $\kappa(R) = \kappa(A)$

Worked solutions

Problem 1

> For $a = (1, 2, 2)^\top$, $b = (3, 0, 4)^\top$ and $c = (2, 1, -2)^\top$, find $a^\top b$, $\|a\|$, $\|b\|$ and the angle between a and b, and show that a and c are orthogonal.

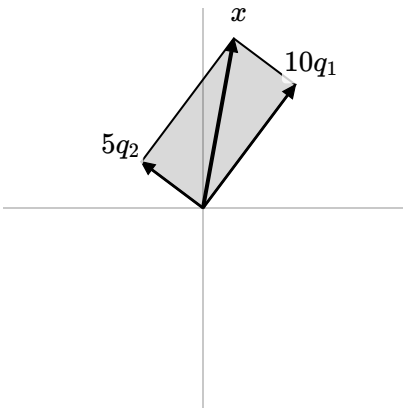
$a^\top b = 1 \cdot 3 + 2 \cdot 0 + 2 \cdot 4 = 11$	Multiply matching entries and add.
$\|a\| = \sqrt{1 + 4 + 4} = 3$ and $\|b\| = \sqrt{9 + 0 + 16} = 5$	The length is the square root of a vector's dot product with itself.
$\cos\theta = \dfrac{11}{3 \cdot 5} = \dfrac{11}{15}$	Divide the dot product by both lengths.
$a^\top c = 2 + 2 - 4 = 0$	A zero dot product is orthogonality.
$a^\top b = 11$, $\|a\| = 3$, $\|b\| = 5$, $\theta = \arccos\frac{11}{15} \approx 42.8°$; $a^\top c = 0$	The cosine is between -1 and 1, as Cauchy–Schwarz, $\|a^\top b\| \le \|a\|\,\|b\|$, guarantees.

Problem 2

> $q_1 = \frac{1}{5}(3, 4)^\top$ and $q_2 = \frac{1}{5}(-4, 3)^\top$ form an orthonormal basis of $\mathbb{R}^2$. Write $x = (2, 11)^\top$ as $c_1 q_1 + c_2 q_2$ without solving a system.

$q_1^\top q_1 = \frac{9+16}{25} = 1$, $q_2^\top q_2 = 1$, $q_1^\top q_2 = \frac{-12+12}{25} = 0$	Check the basis is orthonormal before relying on it.
$q_1^\top x = c_1\, q_1^\top q_1 + c_2\, q_1^\top q_2 = c_1$	Dot both sides of $x = c_1 q_1 + c_2 q_2$ with q_1: the cross term is 0 and $q_1^\top q_1 = 1$.
$c_1 = q_1^\top x = \frac{6+44}{5} = 10$	The coordinate is the dot product.
$c_2 = q_2^\top x = \frac{-8+33}{5} = 5$	The same with q_2.

$x = 10q_1 + 5q_2$



Check: $10q_1 + 5q_2 = (6, 8)^\top + (-4, 3)^\top = (2, 11)^\top$. In a basis that is not orthogonal, the coordinates need a 2×2 solve; here they are two dot products.

Problem 3

Let $x = (1, 2, 2)^\top$ and $v = x - 3e_1 = (-2, 2, 2)^\top$. Compute the Householder reflection $H = I - \dfrac{2vv^\top}{v^\top v}$, and show that $Hx = 3e_1$, $H^\top H = I$, $H^{-1} = H$ and $\det H = -1$.

$v^\top v = 4 + 4 + 4 = 12$ and $vv^\top = \begin{bmatrix} 4 & -4 & -4 \\ -4 & 4 & 4 \\ -4 & 4 & 4 \end{bmatrix}$

Entry (i, j) of $vv^\top$ is $v_i v_j$.

$H = I - \frac{1}{6}vv^\top = \frac{1}{3}\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{bmatrix}$

$1 - \frac{4}{6} = \frac{1}{3}$ on the diagonal, $\frac{4}{6} = \frac{2}{3}$ off it, with the signs of $-vv^\top$.

$Hx = \frac{1}{3}(1 + 4 + 4,\ 2 + 2 - 4,\ 2 - 4 + 2)^\top = (3, 0, 0)^\top$

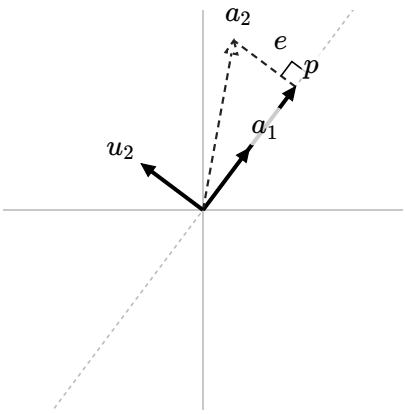
$\|x\| = 3$, so x is sent to the same length along e_1: this is how QR algorithms zero out a column.

$$\underbrace{\begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix}}_{H} \cdot \underbrace{\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}}_{Hx}$$

The columns $\frac{1}{3}(1,2,2)$, $\frac{1}{3}(2,1,-2)$, $\frac{1}{3}(2,-2,1)$ have length 1 and pairwise dot products $\frac{1}{9}(2+2-4) = \frac{1}{9}(2-4+2) = \frac{1}{9}(4-2-2) = 0$	So $H^\top H = I$. H is also symmetric, so $H^{-1} = H^\top = H$: reflecting twice puts every vector back.
$H = \frac{1}{3}\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{bmatrix}$, $Hx = 3e_1$, $H^\top H = I$, $H^{-1} = H$, $\det H = -1$	$Hv = v - 2v = -v$ and $Hw = w$ for every $w \perp v$, so the eigenvalues are $-1, 1, 1$ and $\det H = -1$: a reflection, not a rotation. Lengths are kept: $H(1,1,1)^\top = \frac{1}{3}(5,1,1)^\top$, and both have length $\sqrt{3}$.

Problem 4

Apply Gram–Schmidt to $a_1 = (3,4)^\top$ and $a_2 = (2,11)^\top$.

$\|a_1\| = 5$, so $q_1 = \frac{1}{5}(3,4)^\top$	The first vector only needs normalizing.
$q_1^\top a_2 = \frac{6+44}{5} = 10$	The length of a_2 along q_1.
$u_2 = a_2 - 10q_1 = (2,11)^\top - (6,8)^\top = (-4,3)^\top$ 	Remove that component. Check: $a_1^\top u_2 = -12 + 12 = 0$.
$\|u_2\| = 5$, so $q_2 = \frac{1}{5}(-4,3)^\top$	Normalize what is left.

$q_1 = \frac{1}{5}(3,4)^\top$, $q_2 = \frac{1}{5}(-4,3)^\top$	The basis of Problem 2, and its coordinates of a_2 there, 10 and 5, are exactly the $q_1^\top a_2$ and $\|u_2\|$ found here: $R = \begin{bmatrix} 5 & 10 \\ 0 & 5 \end{bmatrix}$.

Problem 5

Find the QR decomposition of $A = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 2 & 2 \end{bmatrix}$.

$r_{11} = \|a_1\| = \sqrt{1+4+4} = 3$ and $q_1 = \frac{1}{3}(1,2,2)^\top$	The first column of Q is the first column of A, normalized.
$r_{12} = q_1^\top a_2 = \frac{4+10+4}{3} = 6$	The entry above the diagonal in column 2.
$u_2 = a_2 - 6q_1 = (4,5,2)^\top - (2,4,4)^\top = (2,1,-2)^\top$	Check: $q_1^\top u_2 = \frac{1}{3}(2+2-4) = 0$.
$r_{22} = \|u_2\| = 3$ and $q_2 = \frac{1}{3}(2,1,-2)^\top$	The diagonal entries of R are the lengths of the u's, so they are positive.
$Q = \frac{1}{3}\begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 2 & -2 \end{bmatrix}$, $R = \begin{bmatrix} 3 & 6 \\ 0 & 3 \end{bmatrix}$ $\begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & -\frac{2}{3} \end{bmatrix} \cdot \begin{bmatrix} 3 & 6 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 2 & 2 \end{bmatrix}$ Q $\quad$ R $\quad$ A	Check QR: column 1 is $3q_1 = (1,2,2)^\top$ and column 2 is $6q_1 + 3q_2 = (2,4,4)^\top + (2,1,-2)^\top = (4,5,2)^\top$. And $R = Q^\top A$, as it must.

Problem 6

Extend Problem 5 with a third column: find the QR decomposition of $A = \begin{bmatrix} 1 & 4 & 5 \\ 2 & 5 & 1 \\ 2 & 2 & 1 \end{bmatrix}$, and use it to find $\det A$.

$q_1 = \frac{1}{3}(1,2,2)^\top$, $q_2 = \frac{1}{3}(2,1,-2)^\top$, $r_{11} = 3$, $r_{12} = 6$, $r_{22} = 3$	The first two columns are Problem 5 unchanged: Gram–Schmidt never revisits earlier columns.
$r_{13} = q_1^\top a_3 = \frac{5+2+2}{3} = 3$ and $r_{23} = q_2^\top a_3 = \frac{10+1-2}{3} = 3$	Project $a_3 = (5,1,1)^\top$ onto each finished q.
$u_3 = a_3 - 3q_1 - 3q_2 = (5,1,1)^\top - (1,2,2)^\top - (2,1,-2)^\top = (2,-2,1)^\top$	Check: $q_1^\top u_3 = \frac{1}{3}(2-4+2) = 0$ and $q_2^\top u_3 = \frac{1}{3}(4-2-2) = 0$.
$r_{33} = 3$, $q_3 = \frac{1}{3}(2,-2,1)^\top$, so $Q = \frac{1}{3}\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{bmatrix}$	The same matrix as the reflection H of Problem 3, so $\det Q = -1$.
$R = \begin{bmatrix} 3 & 6 & 3 \\ 0 & 3 & 3 \\ 0 & 0 & 3 \end{bmatrix}$ **and** $\det A = \det Q \, \det R = (-1)(27) = -27$	R is triangular, so its determinant is $3 \cdot 3 \cdot 3$. The sign comes from Q: $\lvert\det A\rvert$ is always the product of the r_{jj}, the volume of the box the columns span.

Problem 7

Use the QR decomposition of Problem 5 to find the least-squares solution of $Ax = b$ for $A = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 2 & 2 \end{bmatrix}$ and $b = (3,3,6)^\top$, and the residual.

$A^\top A\hat{x} = A^\top b$ becomes $R^\top Q^\top Q R\hat{x} = R^\top Q^\top b$, so $R\hat{x} = Q^\top b$	$Q^\top Q = I$, and $R^\top$ is invertible because R's diagonal is positive.
$Q^\top b = \left(\frac{3+6+12}{3},\ \frac{6+3-12}{3}\right)^\top = (7,-1)^\top$	One dot product with each q.

$3\hat{x}_2 = -1$ and $3\hat{x}_1 + 6\hat{x}_2 = 7$	R is upper triangular: back substitution, no elimination needed.
$\hat{x}_2 = -\frac{1}{3}$ and $\hat{x}_1 = \frac{7+2}{3} = 3$ $\underset{R}{\begin{bmatrix} 3 & 6 \\ 0 & 3 \end{bmatrix}} \cdot \underset{\hat{x}}{\begin{bmatrix} 3 \\ -\frac{1}{3} \end{bmatrix}} = \underset{Q^\top b}{\begin{bmatrix} 7 \\ -1 \end{bmatrix}}$	From the bottom row up.
$\hat{x} = \left(3, -\frac{1}{3}\right)^\top$, $A\hat{x} = \frac{1}{3}(5, 13, 16)^\top$, **residual** $r = \frac{2}{3}(2, -2, 1)^\top$ **with** $\lVert r \rVert = 2$	$A\hat{x} = 7q_1 - q_2$. The residual is a multiple of q_3 from Problem 6, orthogonal to both columns: $a_1^\top r = \frac{2}{3}(2 - 4 + 2) = 0$ and $a_2^\top r = \frac{2}{3}(8 - 10 + 2) = 0$.

Problem 8

For the tall Q of Problem 5, find $P = QQ^\top$ and show it is the projection onto the plane spanned by the columns of A. Find the orthogonal complement of that plane, and show $QQ^\top \neq I$ although $Q^\top Q = I$.

$Q^\top Q = \frac{1}{9}\begin{bmatrix} 9 & 0 \\ 0 & 9 \end{bmatrix} = I_2$	The columns are orthonormal.
$P = q_1 q_1^\top + q_2 q_2^\top = \frac{1}{9}\begin{bmatrix} 5 & 4 & -2 \\ 4 & 5 & 2 \\ -2 & 2 & 8 \end{bmatrix}$	$QQ^\top$ is the sum of the projections onto each q; for example, its top-left entry is $\frac{1+4}{9}$.
$P^2 = Q(Q^\top Q)Q^\top = QQ^\top = P$ and $Pa_1 = a_1$, $Pa_2 = a_2$	Projecting twice is projecting once, and vectors in the plane are left alone.
The complement is spanned by $(2, -2, 1)^\top$: it is orthogonal to $(1, 2, 2)$ and $(4, 5, 2)$, and $P(2, -2, 1)^\top = 0$	Gram–Schmidt on a third vector gave it in Problem 6. So $I - P = q_3 q_3^\top$, the projection onto the complement.

$P = \frac{1}{9}\begin{bmatrix} 5 & 4 & -2 \\ 4 & 5 & 2 \\ -2 & 2 & 8 \end{bmatrix}$; **complement spanned by** $(2,-2,1)^\top$; $QQ^\top$ **has trace** 2, **so it is not** I_3	The trace of a projection is the dimension it projects onto: $\frac{5+5+8}{9} = 2$, not 3. Only a square Q has $QQ^\top = I$.

Problem 9

Run Gram–Schmidt on $a_1 = (1,2,2)^\top$, $a_2 = (4,5,2)^\top$, $a_3 = (6,9,6)^\top$. What happens at the third vector, and what does it say about $A = [a_1\ a_2\ a_3]$?

$q_1 = \frac{1}{3}(1,2,2)^\top$ and $q_2 = \frac{1}{3}(2,1,-2)^\top$	As in Problem 5.
$q_1^\top a_3 = \frac{6+18+12}{3} = 12$ and $q_2^\top a_3 = \frac{12+9-12}{3} = 3$	The two projection coefficients.
$u_3 = (6,9,6)^\top - (4,8,8)^\top - (2,1,-2)^\top = (0,0,0)^\top$	Nothing is left after removing the components along q_1 and q_2: a_3 lies in their plane.
$a_3 = 12q_1 + 3q_2 = 4a_1 + (a_2 - 2a_1) = 2a_1 + a_2$	Rewrite the q's in terms of the a's with Problem 5: $3q_1 = a_1$ and $3q_2 = a_2 - 2a_1$. Check: $2(1,2,2) + (4,5,2) = (6,9,6)$.
$u_3 = 0$: $a_3 = 2a_1 + a_2$, **the columns are dependent,** A **has rank** 2 **and** $\det A = 0$	There is no q_3 to make, and r_{33} would be 0. In a data matrix this is a feature that is a combination of two others.

Problem 10

Two features are almost collinear:

$$A = \begin{bmatrix} 1 & 1 \\ \varepsilon & 0 \\ 0 & \varepsilon \end{bmatrix}, \qquad \varepsilon = 10^{-8}.$$

> Show that $\kappa(A^\top A) = \kappa(A)^2$ in general, find both for this A, and explain why forming $A^\top A$ in double precision fails while QR on A does not.

$A = U\Sigma V^\top$ gives $A^\top A = V\Sigma^2 V^\top$, so the singular values of $A^\top A$ are σ_i^2	$U^\top U = I$ cancels in the middle.
$\kappa(A^\top A) = \sigma_{\max}^2/\sigma_{\min}^2 = \kappa(A)^2$	Forming $A^\top A$ squares the condition number before any solving starts.
$A^\top A = \begin{bmatrix} 1+\varepsilon^2 & 1 \\ 1 & 1+\varepsilon^2 \end{bmatrix}$ with eigenvalues $2+\varepsilon^2$ and ε^2	Eigenvectors $(1,1)^\top$ and $(1,-1)^\top$. So $\sigma_1(A) = \sqrt{2+\varepsilon^2}$ and $\sigma_2(A) = \varepsilon$.
$\kappa(A) = \sqrt{2+\varepsilon^2}/\varepsilon \approx 1.41 \times 10^8$ and $\kappa(A^\top A) \approx 2 \times 10^{16}$	$1+\varepsilon^2 = 1 + 10^{-16}$ rounds to 1 in double precision (the spacing of doubles near 1 is about 2.2×10^{-16}), so the computed $A^\top A$ is exactly $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, which is singular.
$\kappa(A) \approx 1.41 \times 10^8$, $\kappa(A^\top A) \approx 2 \times 10^{16}$: **the normal equations lose every digit, while QR works on A itself, with condition number $\kappa(R) = \kappa(A)$**	For $b = A(2,1)^\top$, QR returns $(2,1)$ to at least eight digits, and $r_{22} \approx \sqrt{2}\,\varepsilon$ is small but not lost. This is why least-squares solvers use QR or the SVD, and why nearly collinear features make the normal equations unreliable long before they are exactly singular.

Where this goes wrong

1. Projection coefficient not divided by a·a

The projection onto a_1 is remembered as "dot product times a_1", and a_1 is not a unit vector.

$a_1 = (3,4)^\top$, $a_2 = (2,11)^\top$, $a_1^\top a_2 = 50$	Right so far: Problem 4.
"Subtract the dot product times a_1."	The step that causes the mistake: $(a_1^\top a_2)\, a_1$ is the projection only when $\|a_1\| = 1$.

$u_2 = (2, 11)^\top - 50(3, 4)^\top = (-148, -189)^\top$ ✗	Check: $a_1^\top u_2 = -444 - 756 = -1200$, not 0. The coefficient is $\frac{a_1^\top a_2}{a_1^\top a_1} = \frac{50}{25} = 2$, giving $u_2 = (2, 11)^\top - (6, 8)^\top = (-4, 3)^\top$. Dividing by $\|a_1\| = 5$ instead of $\|a_1\|^2$ fails too: $(2, 11)^\top - 10(3, 4)^\top = (-28, -29)^\top$, with $a_1^\top u_2 = -200$.

2. Projecting onto a column of A instead of q

The third step subtracts projections onto the earlier vectors, and the earlier columns of A are the nearest to hand.

$a_1 = (1, 2, 2)^\top$, $a_2 = (4, 5, 2)^\top$, $a_3 = (5, 1, 1)^\top$	Right so far: Problem 6.
"Subtract the projections of a_3 onto a_1 and onto a_2."	The step that causes the mistake: a_1 and a_2 are not orthogonal ($a_1^\top a_2 = 18$), so their projections overlap.
$u_3 = a_3 - \frac{9}{9}a_1 - \frac{27}{45}a_2 = \left(\frac{8}{5}, -4, -\frac{11}{5}\right)^\top$ ✗	Check: $a_1^\top u_3 = \frac{8}{5} - 8 - \frac{22}{5} = -\frac{54}{5}$, not 0. The part of a_3 along the shared direction is removed twice. Project onto q_1 and q_2, which are orthogonal: $u_3 = a_3 - 3q_1 - 3q_2 = (2, -2, 1)^\top$.

3. R computed as A^T Q

R is built from dot products of q's with a's, and the order of the transpose is easy to swap.

$Q = \frac{1}{3}\begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 2 & -2 \end{bmatrix}$ for $A = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 2 & 2 \end{bmatrix}$	Right so far: Problem 5.
"R holds the dot products of the columns, so $R = A^\top Q$."	The step that causes the mistake: $A^\top Q = (Q^\top A)^\top$ is R transposed.

$R = \begin{bmatrix} 3 & 0 \\ 6 & 3 \end{bmatrix}$ ✗	It is lower triangular, which is already the warning. Multiplying back, the first column of QR is $3q_1 + 6q_2 = (5, 4, -2)^\top$, not $a_1 = (1, 2, 2)^\top$. The right order is $R = Q^\top A = \begin{bmatrix} 3 & 6 \\ 0 & 3 \end{bmatrix}$: row i is q_i dotted with each column of A.

4. QQᵀ taken to be I for a tall Q

$Q^\top Q = I$, and the identity is expected to work in the other order too.

$Q^\top Q = I_2$ for the 3×2 matrix Q of Problem 5	Right so far.
"So $Q^{-1} = Q^\top$, and $QQ^\top = I$."	The step that causes the mistake: Q is not square, so it has no inverse, and $Q^\top$ is only a left inverse.
$A\hat{x} = QQ^\top b = b$, so the least-squares fit of Problem 7 is exact and the residual is 0 ✗	$QQ^\top b = 7q_1 - q_2 = \frac{1}{3}(5, 13, 16)^\top$, not $b = (3, 3, 6)^\top$: $QQ^\top$ projects onto a plane, with trace 2. The residual is $\frac{2}{3}(2, -2, 1)^\top$, of length 2.

5. Normalizing a zero vector

Every Gram–Schmidt step ends with a division by a length, and it is done without looking.

$u_3 = a_3 - 12q_1 - 3q_2$ for $a_3 = (6, 9, 6)^\top$	Right so far: Problem 9.
With measured data or inexact arithmetic, a dependent column leaves a u_3 of size around 10^{-15} rather than exactly 0	Rounding leaves crumbs; these small integers happen to cancel exactly.

$q_3 = u_3/\|u_3\|$ is a unit vector, so the columns are independent and A has rank 3 ✗	The direction of u_3 is rounding noise, and dividing by its tiny length blows the noise up to length 1. Compare $\|u_3\|$ with $\|a_3\| = \sqrt{153} \approx 12.4$ before normalizing: a ratio near machine precision means a_3 is a combination of the earlier columns, here $a_3 = 2a_1 + a_2$, and the rank is 2.

More at eulerpad.app. This is one of 82 free practice sets with worked solutions at eulerpad.app/practice, alongside guides to the topics behind them.

EulerPad, the app, reads your handwritten math on iPad, iPhone and Android and checks it line by line. It's free to start.

This set is also Chapter 11 of *Linear Algebra for Machine Learning by Hand*, a 250-problem workbook with worked solutions: eulerpad.app/workbooks

eulerpad.app/practice/gram-schmidt-and-qr