Example 71. A matrix A satisfies $A^T A = I$ if and only if ...

Solution. ... the columns of A are orthonormal.

Why? Let $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots$ be the columns of A. By the way matrix multiplication works, the entries of $A^T A$ are dot products of these columns:

$$\begin{bmatrix} — & \boldsymbol{a}_1^T & — \\ — & \boldsymbol{a}_2^T & — \\ & \vdots & \end{bmatrix} \begin{bmatrix} | & | & \\ \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots \\ | & | & \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \ddots \end{bmatrix}$$

Hence, $A^T A = I$ if and only if the dot products $\boldsymbol{a}_i^T \boldsymbol{a}_j = 0$, for $i \neq j$, and $\boldsymbol{a}_i^T \boldsymbol{a}_i = 1$.

Definition 72. An **orthogonal matrix** is a square matrix with orthonormal columns.

[This is not a typo (but a confusing convention): the columns need to orthonormal, not just orthogonal.]

An $n \times n$ matrix Q is orthogonal $\iff Q^T Q = I$

In other words, $Q^{-1} = Q^T$.

Example 73. What can we say about $\det(Q)$ if Q is orthogonal?

Solution. Write $d = \det(Q)$. Since $Q^{-1} = Q^T$, we have $\frac{1}{d} = d$ (recall that $\det(Q^{-1}) = 1 / \det(Q)$ and $\det(Q^T) = \det(Q)$) or, equivalently, $d^2 = 1$. Hence, $d = \pm 1$.

Both of these are possible as the examples $Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $Q = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ illustrate.

Example 74. Is $Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ orthogonal?

Solution. Yes! The columns are clearly orthogonal. They also have length 1: $\left\| \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \right\| = \sqrt{\cos^2\theta + \sin^2\theta} = 1$.

Comment. Actually, every orthogonal 2×2 matrix Q with $\det(Q) = 1$ is of this form. Geometrically, this is a rotation by the angle θ.

The following is just the Gram–Schmidt orthogonalization (make sure that's obvious!) from last class, except that we immediately normalize each vector $\boldsymbol{q}_i$.

(Gram–Schmidt orthonormalization)

Given a basis $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots$ for W, produce an orthonormal basis $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots$ for W.

- $\boldsymbol{q}_1 = \frac{\boldsymbol{b}_1}{\|\boldsymbol{b}_1\|}$ with $\boldsymbol{b}_1 = \boldsymbol{w}_1$
- $\boldsymbol{q}_2 = \frac{\boldsymbol{b}_2}{\|\boldsymbol{b}_2\|}$ with $\boldsymbol{b}_2 = \boldsymbol{w}_2 - (\boldsymbol{w}_2 \cdot \boldsymbol{q}_1)\boldsymbol{q}_1$
- $\boldsymbol{q}_3 = \frac{\boldsymbol{b}_3}{\|\boldsymbol{b}_3\|}$ with $\boldsymbol{b}_3 = \boldsymbol{w}_3 - (\boldsymbol{w}_3 \cdot \boldsymbol{q}_1)\boldsymbol{q}_1 - (\boldsymbol{w}_3 \cdot \boldsymbol{q}_2)\boldsymbol{q}_2$
- $\boldsymbol{q}_4 = \ldots$

Armin Straub
straub@southalabama.edu

Example 75. Find an orthonormal basis for $W = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}$.

Solution. Let $\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3$ be the vectors spaning W. We then construct an orthonormal basis $\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3$ using Gram–Schmidt orthonormalization as follows:

- $\boldsymbol{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, so that $\boldsymbol{q}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$.
- $\boldsymbol{b}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} - \left(\begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} \cdot \boldsymbol{q}_1 \right) \boldsymbol{q}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$, so that $\boldsymbol{q}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$.
- $\boldsymbol{b}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \left(\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \cdot \boldsymbol{q}_1 \right) \boldsymbol{q}_1 - \left(\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \cdot \boldsymbol{q}_2 \right) \boldsymbol{q}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$, so that $\boldsymbol{q}_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$.

5.6 The QR decomposition

Just like the LU decomposition encodes the steps of Gaussian elimination, the QR decomposition encodes the steps of Gram–Schmidt.

(QR decomposition) Every $m \times n$ matrix A of rank n can be decomposed as $A = QR$, where

- Q has orthonormal columns, $(m \times n)$
- R is upper triangular and invertible. $(n \times n)$

How to find $\boldsymbol{Q}$ and $\boldsymbol{R}$?

- Gram–Schmidt orthonormalization on (columns of) A, to get (columns of) Q
- $R = Q^T A$
  Why? If $A = QR$, then $Q^T A = Q^T Q R$ which simplifies to $R = Q^T A$ (since $Q^T Q = I$).

The decomposition $A = QR$ is unique if we require the diagonal entries of R to be positive (and this is exactly what happens when applying Gram–Schmidt).

Practical comment. Actually, no extra work is needed for computing R. All of its entries have been computed during Gram–Schmidt.

Variations. We can also arrange things so that Q is an $m \times m$ orthogonal matrix and R a $m \times n$ upper triangular matrix. This is a tiny bit more work (and not required for many applications): we need to complement "our" Q with additional orthonormal columns and add corresponding zero rows to R. For square matrices this makes no difference.

Armin Straub
straub@southalabama.edu

Example 76. Determine the QR decomposition of $A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$.

Solution. The first step is Gram–Schmidt orthonormalization on the columns of A. We then use the resulting orthonormal vectors as the columns of Q.

We already did Gram–Schmidt in Example 75: from that work, we have $Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/\sqrt{2} \\ 0 & 0 & 1/\sqrt{2} \end{bmatrix}$.

Hence, $R = Q^T A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & \sqrt{2} \end{bmatrix}$.

Comment. As commented earlier, the entries of R have actually all been computed during Gram–Schmidt, so that, if we pay attention, we could immediately write down R (no extra work required). Looking back at Example 75, can you see this?

Check. Indeed, $QR = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/\sqrt{2} \\ 0 & 0 & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & \sqrt{2} \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ equals A.

Example 77. (homework) Determine the QR decomposition of $A = \begin{bmatrix} 1 & 2 & 4 \\ 0 & 0 & 5 \\ 0 & 3 & 6 \end{bmatrix}$.

Solution. We first apply Gram–Schmidt orthonormalization to the columns of A.

- $\boldsymbol{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, so that $\boldsymbol{q}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.
- $\boldsymbol{b}_2 = \begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix} - \left(\begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix} \cdot \boldsymbol{q}_1 \right) \boldsymbol{q}_1 = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$, so that $\boldsymbol{q}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.
- $\boldsymbol{b}_3 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} - \left(\begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \cdot \boldsymbol{q}_1 \right) \boldsymbol{q}_1 - \left(\begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \cdot \boldsymbol{q}_2 \right) \boldsymbol{q}_2 = \begin{bmatrix} 0 \\ 5 \\ 0 \end{bmatrix}$, so that $\boldsymbol{q}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$.

Therefore, $Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$.

Finally, $R = Q^T A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 4 \\ 0 & 0 & 5 \\ 0 & 3 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 4 \\ 0 & 3 & 6 \\ 0 & 0 & 5 \end{bmatrix}$.

In conclusion, we have found the QR decomposition:

$$\underbrace{\begin{bmatrix} 1 & 2 & 4 \\ 0 & 0 & 5 \\ 0 & 3 & 6 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} 1 & 2 & 4 \\ 0 & 3 & 6 \\ 0 & 0 & 5 \end{bmatrix}}_{R}$$