

Structure of matrix R

On the previous page, we demonstrated how to

- Take a full-column-rank matrix $A = [\vec{a}_1 \dots \vec{a}_n]$
- Obtain matrix $Q = [\vec{e}_1 \dots \vec{e}_r]$ where
 - $\vec{e}_i \cdot \vec{e}_j = 0$ for $i \neq j$
 - $\vec{e}_i \cdot \vec{e}_i = 1$
 - $\vec{e}_j \cdot \vec{a}_i = 0$ for all $j > i$ (eq 1)
- $r =$ number of linearly independent columns in A
- from above, Q satisfies the condition $Q^T Q = I$

Next, will derive matrix $R = Q^T A$,

which comes from $A = Q R$

↓

$$Q^T A = Q^T Q R$$

$$\begin{bmatrix} \vec{e}_1^T \\ \vec{e}_2^T \\ \vec{e}_3^T \end{bmatrix} \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \vec{a}_3 \end{bmatrix} = \begin{bmatrix} \vec{e}_1^T \cdot \vec{a}_1 & \vec{e}_1^T \cdot \vec{a}_2 & \vec{e}_1^T \cdot \vec{a}_3 \\ \vec{e}_2^T \cdot \vec{a}_1 & \vec{e}_2^T \cdot \vec{a}_2 & \vec{e}_2^T \cdot \vec{a}_3 \\ \vec{e}_3^T \cdot \vec{a}_1 & \vec{e}_3^T \cdot \vec{a}_2 & \vec{e}_3^T \cdot \vec{a}_3 \end{bmatrix}$$

From equation 1, zeros appear everywhere below the diagonal:

$$\begin{bmatrix} \vec{e}_1^T \cdot \vec{a}_1 & \vec{e}_1^T \cdot \vec{a}_2 & \vec{e}_1^T \cdot \vec{a}_3 \\ \vec{e}_2^T \cdot \vec{a}_1 & \vec{e}_2^T \cdot \vec{a}_2 & \vec{e}_2^T \cdot \vec{a}_3 \\ \vec{e}_3^T \cdot \vec{a}_1 & \vec{e}_3^T \cdot \vec{a}_2 & \vec{e}_3^T \cdot \vec{a}_3 \end{bmatrix} = \begin{bmatrix} \vec{e}_1^T \cdot \vec{a}_1 & \vec{e}_1^T \cdot \vec{a}_2 & \vec{e}_1^T \cdot \vec{a}_3 \\ 0 & \vec{e}_2^T \cdot \vec{a}_2 & \vec{e}_2^T \cdot \vec{a}_3 \\ 0 & 0 & \vec{e}_3^T \cdot \vec{a}_3 \end{bmatrix}$$

$A = QR$ with orthonormal Q and upper-triangular R with the following dimensions:

- $A: m \times n$
- $Q: m \times r$

- $R: r \times n$

Diagonal entries of R

From the orthogonal decomposition

- $\vec{a}_i = (\text{old directions}) + \vec{u}_i$
- $\vec{u}_i \perp (\text{old directions})$

only \vec{u}_i contributes to the dot product with itself,
so $\vec{a}_i \cdot \vec{u}_i = \vec{u}_i \cdot \vec{u}_i > 0$

Since \vec{e}_i is the normalized version of \vec{u}_i ,

$$r_{ii} = \vec{e}_i \cdot \vec{a}_i = \|\vec{u}_i\| \geq 0$$

- $r_{ii} > 0$ if \vec{a}_i adds a new independent direction
- $r_{ii} = 0$ if \vec{a}_i lies in the span of previous columns

Looking ahead

In the QR iteration for eigenvalues,
each step factors $A = QR$ and forms the next matrix as RQ
If diagonal entries of R were allowed to be negative,
columns of Q could flip sign from step to step,
causing artificial oscillations in the iteration

Requiring $r_{ii} \geq 0$ fixes the orientation of each orthogonal direction
and makes the QR iteration behave consistently



Geometric view of QR factorization

(full column rank)

$$\text{Consider matrix } A = \left[\begin{array}{c|c} 1 & 0 \\ \hline 1 & 2 \end{array} \right]$$

$$\vec{a}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad |\vec{a}_1| = \sqrt{2}$$

↓

$$\vec{e}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$\vec{a}_2 = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$

$$\text{Projection of } \vec{a}_2 \text{ on } \vec{e}_1: \text{proj} = (\vec{a}_2 \cdot \vec{e}_1) \vec{e}_1 = (\sqrt{2}) \vec{e}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\vec{u}_2 = \vec{a}_2 - \text{proj} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$|\vec{u}_2| = \sqrt{2}$$

(\vec{u}_2 is the component of \vec{a}_2 orthogonal to \vec{e}_1)

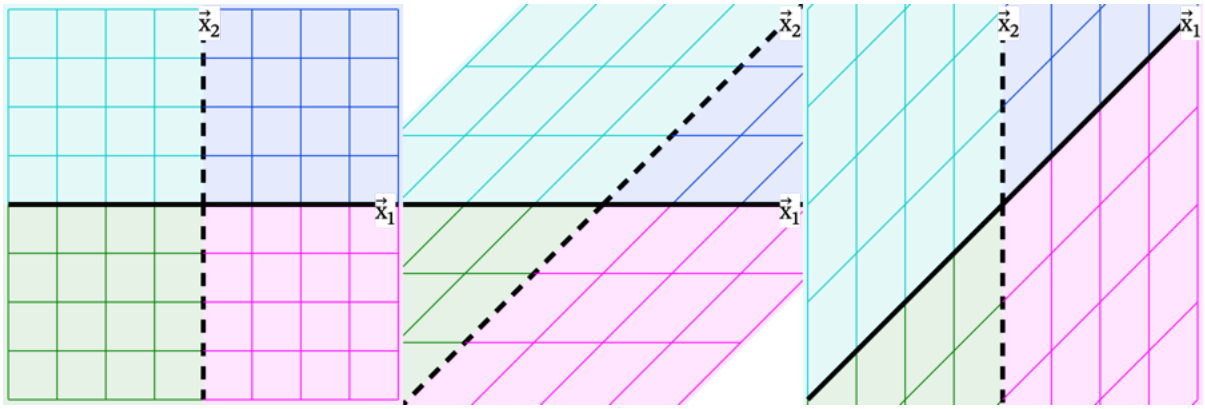
$$\vec{e}_2 = \frac{\vec{u}_2}{|\vec{u}_2|} = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$Q = [\vec{e}_1 | \vec{e}_2] = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$R = Q^T A = \begin{bmatrix} \sqrt{2} & \sqrt{2} \\ 0 & \sqrt{2} \end{bmatrix}$$

Now, as Q and R are derived, we can look at images
and see how this factorization
decouples deformation from rotation:

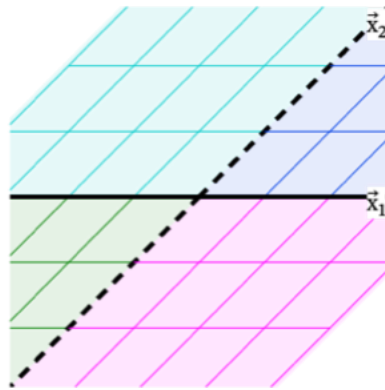
1. original grid
2. hierarchical distortion of the grid by R
(meaning of hierarchical will be discussed on the next page)
3. rotation or reflection of the deformed grid by Q





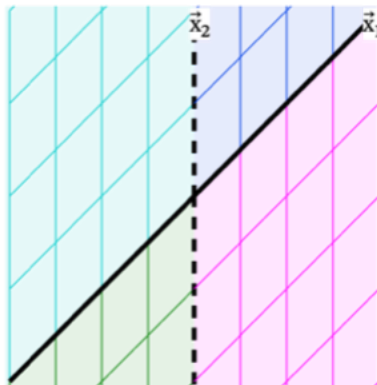
$$R = \left[\begin{array}{c|c} 1.414 & 1.414 \\ \hline 0 & 1.414 \end{array} \right]$$

upper triangular transformation



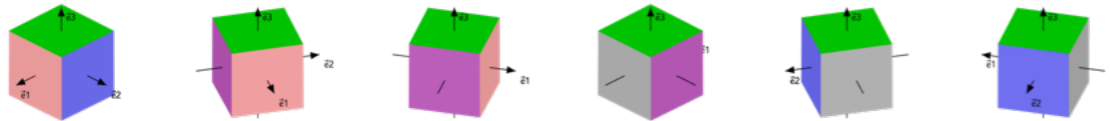
$$Q = \left[\begin{array}{c|c} 0.707 & -0.707 \\ \hline 0.707 & 0.707 \end{array} \right]$$

orthogonal transformation



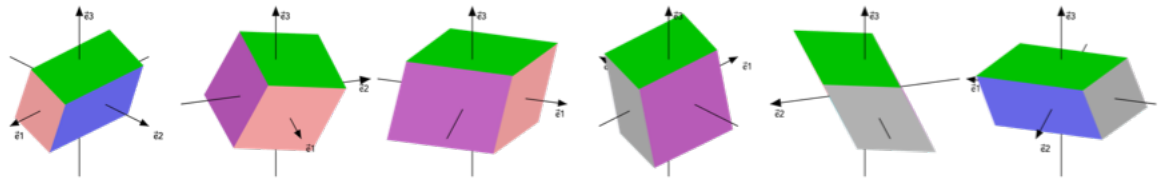
$$A = QR =$$

$$\left[\begin{array}{c|c} 1 & 0 \\ \hline 1 & 2 \end{array} \right]$$



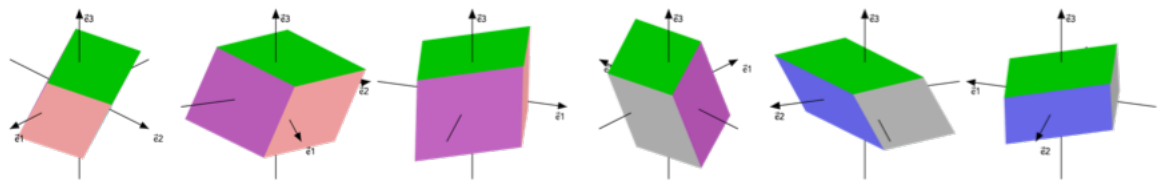
$$R = \begin{bmatrix} 1.5 & 0.35 & 0.15 \\ 0 & 1 & 0.45 \\ 0 & 0 & 1 \end{bmatrix}$$

upper triangular transformation



$$Q = \begin{bmatrix} 0.866 & -0.5 & 0 \\ 0.5 & 0.866 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

orthogonal transformation



$$A = QR =$$

$$\begin{bmatrix} 1.299 & -0.197 & -0.095 \\ 0.75 & 1.041 & 0.465 \\ 0 & 0 & 1 \end{bmatrix}$$

Alternative geometric view
(full column rank case)

$A = QR$ admits two interpretations:

① Transformation sequence:

- deformation by R
- followed by rigid motion (rotation or reflection) by Q

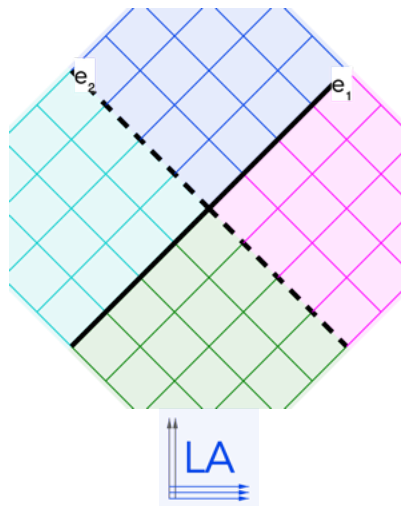
② Basis interpretation:

- Q defines a new orthonormal basis
- R gives coordinates in this basis

For the same matrix shown on the previous page:

$$A = \left[\begin{array}{c|c} 1 & 0 \\ \hline 1 & 2 \end{array} \right] \quad Q = [\vec{e}_1 \mid \vec{e}_2] = \left[\begin{array}{c|c} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \hline \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{array} \right]$$

Q defines the new orthonormal basis used for vectors $R \vec{x}$

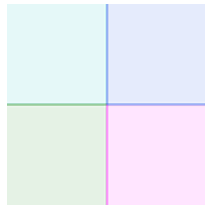


QRF of a 3×2 matrix

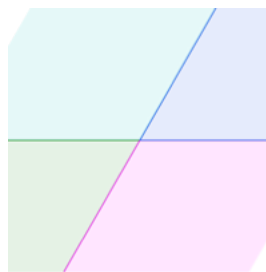
QR factorization of a rank-2 3×2 matrix

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0.5 \times \sqrt{\frac{2}{3}} \\ \frac{1}{2} & -0.5 \times \sqrt{\frac{2}{3}} \\ 0 & \sqrt{\frac{2}{3}} \end{bmatrix} \begin{bmatrix} 2 & \frac{1}{2} \\ 0 & \sqrt{\frac{3}{2}} \end{bmatrix}$$

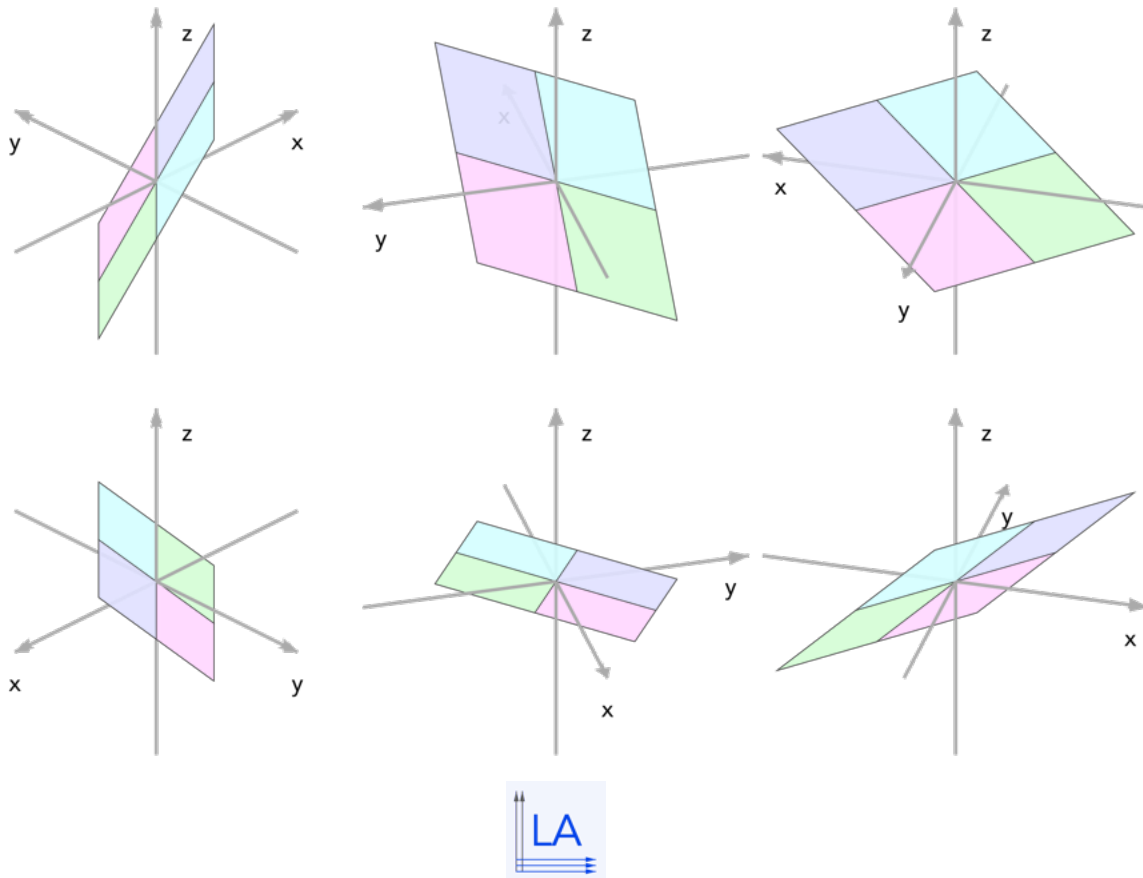
We visualize its domain \mathbb{R}^2 as a 4-color unit square



- ① The 2×2 upper-triangular matrix R deforms the square (shear followed by scaling)



② The 3×2 orthonormal matrix Q isometrically embeds the deformed grid into \mathbb{R}^3 , so the image shows composite transformation by $A = QR$



'Hierarchy' of R

Recall our derivation of R as

$$\begin{bmatrix} \vec{e}_1^T \\ \vec{e}_2^T \\ \vec{e}_3^T \end{bmatrix} \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \vec{a}_3 \end{bmatrix} = \begin{bmatrix} \vec{e}_1^T \cdot \vec{a}_1 & \vec{e}_1^T \cdot \vec{a}_2 & \vec{e}_1^T \cdot \vec{a}_3 \\ \vec{e}_2^T \cdot \vec{a}_1 & \vec{e}_2^T \cdot \vec{a}_2 & \vec{e}_2^T \cdot \vec{a}_3 \\ \vec{e}_3^T \cdot \vec{a}_1 & \vec{e}_3^T \cdot \vec{a}_2 & \vec{e}_3^T \cdot \vec{a}_3 \end{bmatrix}$$

From equation 1, zeros appear everywhere below the diagonal:

$$\left[\begin{array}{c|c|c} \vec{e}_1^T \cdot \vec{a}_1 & \vec{e}_1^T \cdot \vec{a}_2 & \vec{e}_1^T \cdot \vec{a}_3 \\ \hline \vec{e}_2^T \cdot \vec{a}_1 & \vec{e}_2^T \cdot \vec{a}_2 & \vec{e}_2^T \cdot \vec{a}_3 \\ \hline \vec{e}_3^T \cdot \vec{a}_1 & \vec{e}_3^T \cdot \vec{a}_2 & \vec{e}_3^T \cdot \vec{a}_3 \end{array} \right] = \left[\begin{array}{c|c|c} \vec{e}_1^T \cdot \vec{a}_1 & \vec{e}_1^T \cdot \vec{a}_2 & \vec{e}_1^T \cdot \vec{a}_3 \\ \hline 0 & \vec{e}_2^T \cdot \vec{a}_2 & \vec{e}_2^T \cdot \vec{a}_3 \\ \hline 0 & 0 & \vec{e}_3^T \cdot \vec{a}_3 \end{array} \right]$$

Upper triangular R encodes a hierarchy of dependencies
Each new output direction can only use the directions that came before it

For standard basis vectors $\vec{e}_1, \vec{e}_2, \vec{e}_3$ we have

$$R \vec{e}_1 \in \text{span}(\vec{e}_1)$$

$$R \vec{e}_2 \in \text{span}(\vec{e}_1, \vec{e}_2)$$

$$R \vec{e}_3 \in \text{span}(\vec{e}_1, \vec{e}_2, \vec{e}_3)$$

Meaning of hierarchy as forward transformation:

- vector \vec{e}_1 scales only
- vector \vec{e}_2 shears towards \vec{e}_1 and scales
- vector \vec{e}_3 shears towards \vec{e}_1 & \vec{e}_2 and scales

Same hierarchy when solving $R \vec{x} = \vec{b}$ as backward transformation:

$$\left[\begin{array}{c|c|c} r_{11} & r_{12} & r_{13} \\ \hline 0 & r_{22} & r_{23} \\ \hline 0 & 0 & r_{33} \end{array} \right] \left[\begin{array}{c} \vec{x}_1 \\ \vec{x}_2 \\ \vec{x}_3 \end{array} \right] = \left[\begin{array}{c} \vec{b}_1 \\ \vec{b}_2 \\ \vec{b}_3 \end{array} \right]$$

Note the order of dependence:

- forward for space transformation
 - backward for solving



Uniqueness of QR factorization

Each orthogonal direction lies on a line with two possible orientations, so each column of Q could be flipped while preserving orthogonality

However, the Gram–Schmidt construction produces non-negative diagonal entries of R

This sign convention removes the \pm ambiguity

Thin QR constructs an orthonormal basis for the column space of A and enforces an upper triangular R

Therefore, under these requirements, the thin QR factorization is unique even when A is rank-deficient

