Projection of vector $\vec{v}$ on direction of vector $\vec{u}$

If $\vec{u}$ and $\vec{v}$ are vectors in $\mathbb{R}^n$ and $\vec{u} \neq \vec{0}$,
there exist unique vectors $\vec{v}\perp$ and $\vec{v}\|$ satisfying following:

- $\vec{v} = \vec{v}\perp + \vec{v}\|$
- $\vec{u} \cdot \vec{v}\perp = 0$
- $\vec{v}\| = s\,\vec{u}$, where s is a scalar

Then

$$\vec{u} \cdot \vec{v} =$$
$$\vec{u} \cdot (\vec{v}\perp + \vec{v}\|) =$$
$$\vec{u} \cdot \vec{v}\perp + \vec{u} \cdot (s\,\vec{u}) =$$
$$0 + s\,(\vec{u} \cdot \vec{u})$$

$$\text{so } s = \frac{\vec{u} \cdot \vec{v}}{\vec{u} \cdot \vec{u}}$$

- $\vec{v}\| = \left(\dfrac{\vec{u} \cdot \vec{v}}{\vec{u} \cdot \vec{u}}\right) \vec{u}$
- $\vec{v}\perp = \vec{v} - \left(\dfrac{\vec{u} \cdot \vec{v}}{\vec{u} \cdot \vec{u}}\right) \vec{u}$

Projection of vector $\vec{v}$ on columns of $U = [\, \vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r \,]$

This is the same construction as projection of $\vec{v}$ on one vector $\vec{u}$

The only difference is that we replace one direction $\vec{u}$ by r directions $\vec{u}_1, \ldots, \vec{u}_r$

Suppose

- $U = [\ \vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r\ ]$
- $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r \in \mathbb{R}^n$ and are linearly independent
- $\vec{v} \in \mathbb{R}^n$

There exist unique vectors $\vec{v}\perp$ and $\vec{v}||$ such that:

- $\vec{v} = \vec{v}\perp + \vec{v}||$
- $U^T \vec{v}\perp = \vec{0}$

(in other words, $\vec{v}\perp$ is orthogonal to every column of U or every row of U^T)

- $\vec{v}|| = U\,\vec{c}$, where $\vec{c} \in \mathbb{R}^r$

($\vec{c}$ is similar to scalar s used on previous page)

Then

$$U^T \vec{v} = U^T (\vec{v}\perp + \vec{v}||) = U^T \vec{v}\perp + U^T (U\,\vec{c}) = \vec{0} + (U^T U)\,\vec{c}$$

So $\vec{c}$ satisfies

$$(U^T U)\,\vec{c} = U^T \vec{v}$$

Therefore

$$\vec{c} = (U^T U)^{-1} U^T \vec{v}$$

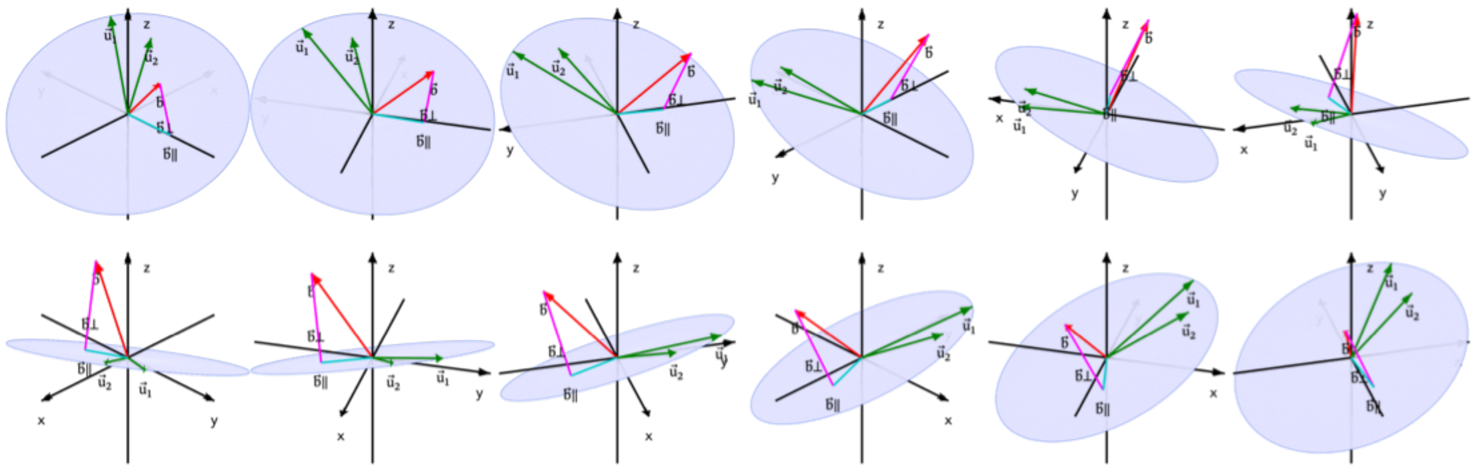
And projection of $\vec{v}$ on columns of U is

$$\vec{v}|| = U\,\vec{c} = U\,(U^T U)^{-1} U^T \vec{v}$$

So the projection matrix onto span($\vec{u}_1, \ldots, \vec{u}_r$) is

$$P = U\,(U^T U)^{-1} U^T$$

The same orthogonality principle appeared earlier in the 'Four Subspaces' chapter where we were solving an inconsistent system $A\,\vec{x} = \vec{b}$ where $\vec{b}$ was outside the column space(A) and used the same fundamental conditions and derivation sequence:



- $\vec{v} = \vec{v}\perp + \vec{v}\|$

$\Leftrightarrow$

$\vec{b} = \vec{b}\perp + \vec{b}\|$

- $U^T \vec{v}\perp = \vec{0}$

$\Leftrightarrow$

$A^T \vec{b}\perp = \vec{0}$

- $A^TA\,\vec{x} = A^T(\vec{b} - \vec{b}\perp) = A^T\vec{b} - A^T\,\vec{b}\perp = A^T\vec{b}$

$\Leftrightarrow$

$U^TU\,\vec{c} = U^T\vec{v}$

- $\vec{x} = (A^TA)^{-1}\,A^T\,\vec{b}$

$\Leftrightarrow$

$\vec{c} = (U^TU)^{-1}\,U^T\,\vec{v}$

While the previous page stopped at solving the normal equation
$\vec{c} = (U^TU)^{-1}\,U^T\,\vec{v}$
this one takes one more step to derive corresponding projection matrix
$P = U\,(U^TU)^{-1}\,U^T$

Idempotency of projection matrix $P = U (U^T U)^{-1} U^T$

P is idempotent which means
it acts once even when applied repeatedly:

$$P^2 = U (U^T U)^{-1} U^T U (U^T U)^{-1} U^T =$$

$$U (U^T U)^{-1} \left((U^T U) (U^T U)^{-1}\right) U^T =$$

$$U (U^T U)^{-1} U^T$$

Gram-Schmidt orthogonalization

On the previous page, we saw an example of
projection of vector $\vec{v}$ onto the direction of vector $\vec{u}$ and derived the following:

$$\vec{v}\perp = \vec{v} - \left(\frac{\vec{u} \cdot \vec{v}}{\vec{u} \cdot \vec{u}}\right) \vec{u}$$

Applying this construction repeatedly allows us to make vectors orthogonal one by one

This time, instead of $\vec{v}$ and $\vec{u}$, we have a set of vectors $\vec{a}_1, \ldots, \vec{a}_n$

We will build new vectors $\vec{q}_1, \ldots, \vec{q}_r$ such that:

- each $\vec{q}_i$ is orthogonal to all previous $\vec{q}_j$
- each $\vec{q}_i$ is a linear combination of $\vec{a}_1, \ldots, \vec{a}_i$

We build the next orthogonal vector by subtracting projections onto all previously built orthogonal vectors

For i = 1…n

Start with $\vec{q}_i = \vec{a}_i$

Then subtract the projection of $\vec{a}_i$ onto $\vec{q}_1$, then onto $\vec{q}_2$, and so on, up to $\vec{q}_{i-1}$

(note that we are subtracting carefully chosen multiples of previous columns)

In one line:

$$\vec{q}_i = \vec{a}_i - \sum_{j=1}^{i-1} \left(\left(\frac{\vec{q}_j \cdot \vec{a}_i}{\vec{q}_j \cdot \vec{q}_j} \right) \vec{q}_j \right)$$

If $\vec{q}_i$ becomes $\vec{0}$, then $\vec{a}_i$ did not add a new direction

it was already a combination of previous vectors, we skip it

Finally, we normalize each non-zero $\vec{q}_i$ to get unit vectors:

$$\vec{e}_i = \frac{\vec{q}_i}{\|\vec{q}_i\|}$$

- $\vec{e}_1, \ldots, \vec{e}_r$ are orthonormal and span the same space as $\vec{a}_1, \ldots, \vec{a}_n$
- $\vec{e}_1$ points in the same direction as $\vec{a}_1$
- r = number of linearly independent columns in the original set

same as $\text{rank}\left(A = [\, \vec{a}_1 \ldots \vec{a}_n \,]\right)$

Structure of matrix Q

On the previous page, we demonstrated an algorithm that

- takes vectors $\vec{a}_1 \ldots \vec{a}_n$
- generates a set of vectors $\vec{e}$ such that

- $\vec{e}_i$ is orthogonal to $\vec{e}_j$

$$\Leftrightarrow$$

$$\vec{e}_i \cdot \vec{e}_j = 0 \text{ for any } i \neq j$$

- each $\vec{e}_i$ is a unit vector

$$\Leftrightarrow$$

$$\vec{e}_i \cdot \vec{e}_i = 1$$

- number of generated vectors r = rank $\left(A = [\, \vec{a}_1 \ldots \vec{a}_n \,]\right)$

We assemble an m×r matrix $Q = [\, \vec{e}_1 \ldots \vec{e}_r \,]$

Because the columns $\vec{e}_i$ are orthonormal,
their dot products form the identity matrix

The (i,j) entry of Q^TQ equals

$$\vec{e}_i \cdot \vec{e}_j$$

But

$$\vec{e}_i \cdot \vec{e}_j = 0 \text{ for } i \neq j$$

$\vec{e}_i \cdot \vec{e}_i = 1$

Therefore

- $Q^T Q = I_r$
- $Q^{-1} = Q^T$ if $r = m$

Properties of Q defined as $Q^T Q = I$

① Multiplication by Q preserves length:
$|\vec{v}| = |Q\,\vec{v}|$ for any $\vec{v}$ of appropriate size

- $|\vec{v}|^2 = \vec{v}^T\,\vec{v}$
- $|Q\,\vec{v}|^2 = (Q\,\vec{v})^T(Q\,\vec{v}) =$
$\vec{v}^T\,Q^T\,Q\,\vec{v} = \vec{v}^T\,\vec{v}$

② Multiplication by Q preserves angles
Suppose $\vec{u}$ and $\vec{v}$ are unit vectors of appropriate size
then angle between them is $\arccos(\vec{u} \cdot \vec{v})$
$\vec{u} \cdot \vec{v} = \vec{u}^T\,\vec{v}$
Will compare it to $(Q\,\vec{u})^T\,(Q\,\vec{v}) = \vec{u}^T\,Q^T\,Q\,\vec{v} = \vec{u}^T\,\vec{v}$

Since multiplication by Q preserves lengths,
angle preservation extends to any non-unit vectors $\vec{u}$ and $\vec{v}$

Geometric view:

- If Q is square,

it preserves distances and angles and therefore acts as reflection or rotation

• If Q is tall with orthonormal columns ($Q^TQ = I$),
then multiplication by Q is an isometric embedding:
it preserves lengths and angles of vectors in its domain

⑤ Special role of square orthogonal matrices

Recall that orthogonality is the strongest geometric form of linear independence
as orthogonal directions share zero overlap

If Q is square and satisfies $Q^TQ = I$, $Q^{-1} = Q^T$,
• Q is invertible in the most stable geometric way as
it preserves lengths and angles
• Q is invertible in the most stable numerical way as
inversion requires only a transpose and does not amplify numerical error

---

Determinant of square Q

Claim: if $Q^TQ = I$, then $|Q| = \{ -1, 1 \}$

Proof:
• $|AB| = |A|\,|B|$
• $Q\,Q^T = I$
• $|\,I\,| = 1$

Taking determinant of $Q\,Q^T = I$:
$|Q\,Q^T| = |\,I\,| = 1$

Determinant of a product equals product of determinants:
$|Q|\,|Q^T| = 1$

$|Q^T| = |Q|$
$\downarrow$
$|Q|^2 = 1$
$\downarrow$
$|Q| = \{-1, 1\}$

- This aligns with Q preserving length and angles
- The converse is not true:

preserving volume alone does not imply preserving lengths or angles

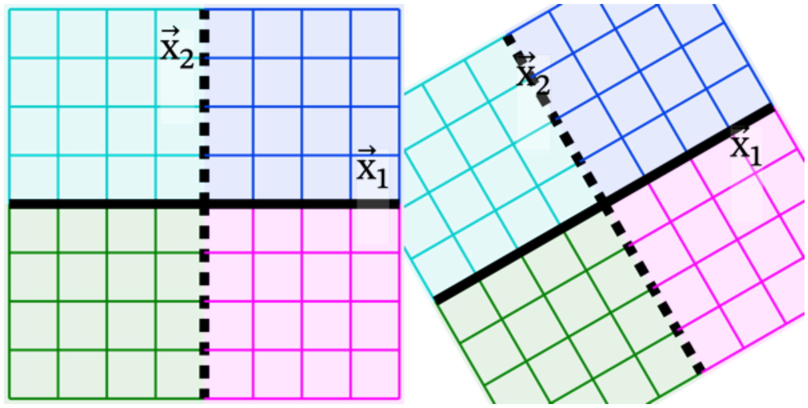
---

Combined rotation and reflection

In the chapter on linear transformations,
we showed examples of rotation & reflection matrices

Here, will show rotation matrix Q(rot) = $\begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}$

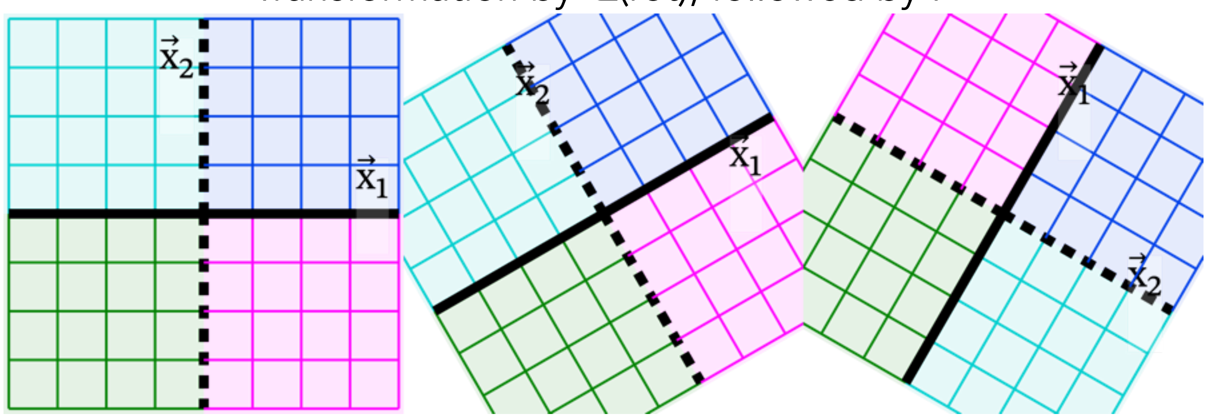
Transformation by Q(rot):



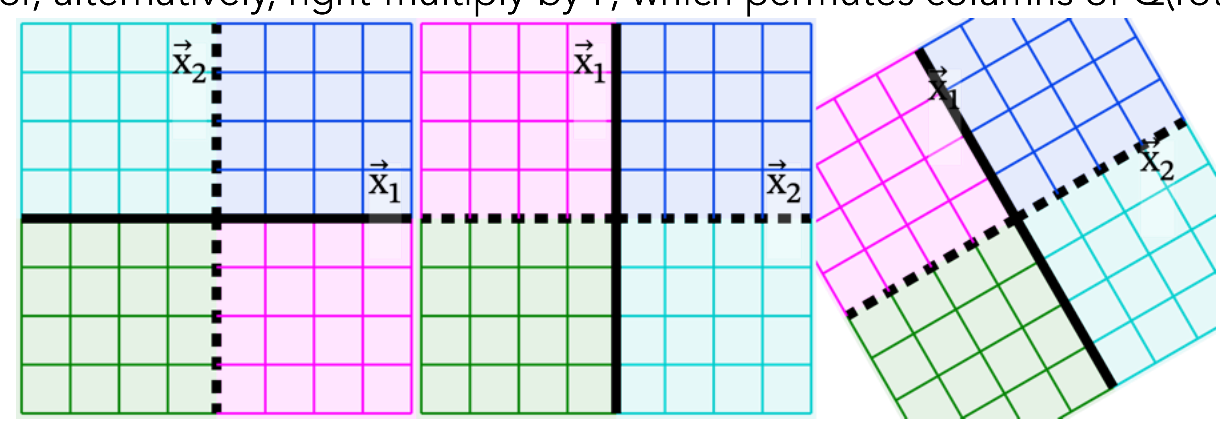
Will left-multiply Q(rot) by reflection matrix $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ $|P| = -1$

- P, by itself, reflects across line (1, 1)
- left-multiplication by P permutes rows of Q(rot)

Transformation by Q(rot), followed by P



or, alternatively, right-multiply by P, which permutes columns of Q(rot)



The two multiplication sequences produce different rotation-reflection matrices with Det = -1

Transformation by a 'tall' Q (m>n)

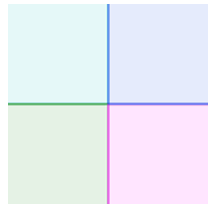
Previously, we explored how a random 3×2 matrix A transformed $\mathbb{R}^2$ into $\mathbb{R}^3$, distorting lengths and angles

Now will consider an example of a 3×2 matrix $Q = [\ \vec{q}_1 \mid \vec{q}_2\ ]$ satisfying

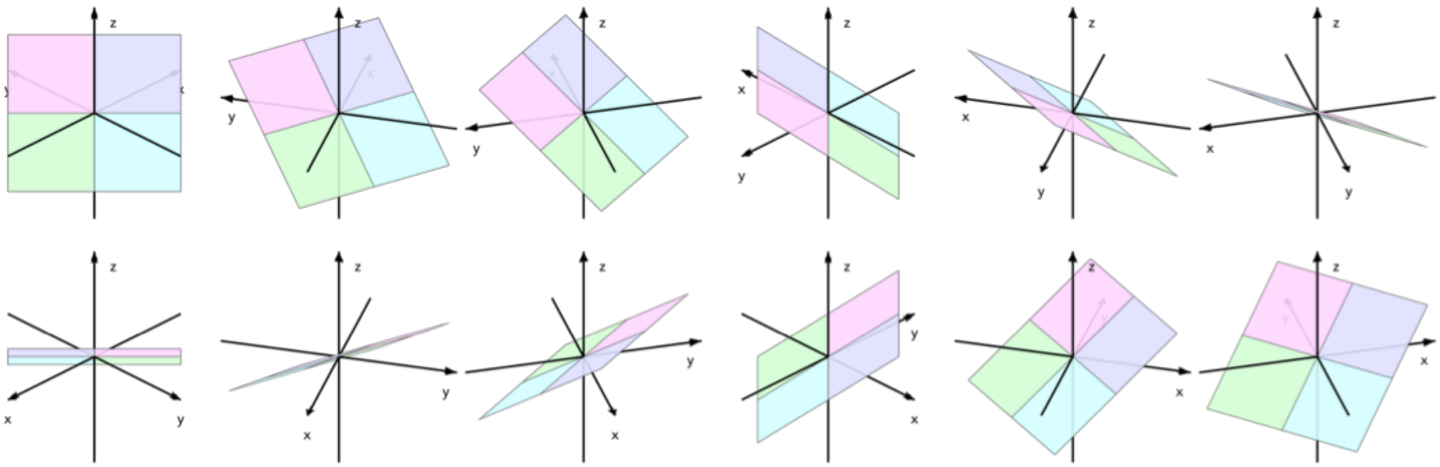
- $|\vec{q}_1| = |\vec{q}_2| = 1$
- $\vec{q}_1 \cdot \vec{q}_2 = 0$
- $Q^T Q = I_2$
- $|Q\,\vec{x}| = |\vec{x}|$
- $(Q\,\vec{u})^T (Q\,\vec{v}) = \vec{u}^T\,\vec{v}$

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

Q acts on a unit square



and embeds it in $\mathbb{R}^3$ without distortion (isometric embedding)



- Q embeds $\mathbb{R}^2$ isometrically into a plane spanned by $\vec{q}_1$ & $\vec{q}_2$ inside $\mathbb{R}^3$
- left-multiplication by Q^T reverses this embedding:

$$Q^T(Q\,\vec{x}) = \vec{x}$$

- $(QQ^T \neq I_3)$
- QQ^T is the orthogonal projection onto the embedded plane (details to follow)

Matrix $Q^T Q$

Assume $Q = [\ \vec{e}_1 \ldots \vec{e}_r\ ]$ has orthonormal columns, so

$$Q^T Q = I_r$$

Earlier we derived the formula for projection of a vector on a matrix

$P = U\,(U^T U)^{-1}\,U^T$

substitute $U = Q$:

$P(\text{orthogonal}) = Q\,(Q^T Q)^{-1}\,Q^T = Q\,I_r\,Q^T = Q\,Q^T$

• $P(\text{orthogonal}) = Q\,Q^T$ is the orthogonal projection onto span($\vec{e}_1$, ..., $\vec{e}_r$)

Matrix P is n×n with rank r

• P(orthogonal) is symmetric:

$P(\text{orthogonal})^T = (Q\,Q^T)^T = Q\,Q^T = P$

• P(orthogonal) is idempotent, as is any projection matrix P:

acts once even when applied repeatedly:

$P^2 = (Q\,Q^T)(Q\,Q^T) = Q\,(Q^T Q)\,Q^T = Q\,Q^T = P$

Structure of matrix R

On the previous page, we demonstrated how to

• Take 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 $Q^{-1} = Q^T$ condition

Next, will derive matrix $R = Q^{-1} A = Q^T A$

Note that this is equivalent to $A = QR$ (the decomposition we are constructing)

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

From equation 1, zeros appear everywhere below 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 following dimensions:

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

Diagonal entries of R

From the orthogonal decomposition

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

Consider matrix $A = \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}$

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

↓

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

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

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

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

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

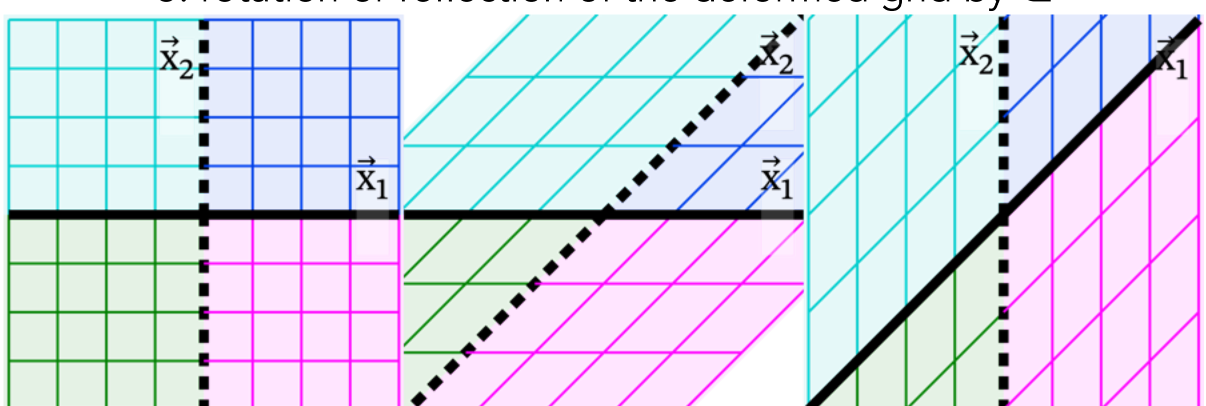
($\vec{u}_2$ is the component of $\vec{a}_2$ orthogonal to $\vec{e}_1$)

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

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

$$R = Q^T A = \left[\begin{array}{c|c} \sqrt{2} & \sqrt{2} \\ \hline 0 & \sqrt{2} \end{array}\right]$$

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



Alternative geometric view
(full column rank)

Recall that earlier we viewed a linear transformation in two ways:
- transforming the geometry while keeping coordinates fixed
- keeping the geometry fixed while changing the coordinate system

QR factorization admits the same dual interpretation

① Transformation interpretation
was shown on the previous page as a sequence:
- deformation by R
- rigid motion (rotation or reflection) by Q

② Basis replacement
QR factorization extracts orthonormal vectors from the columns of A
These vectors form the columns of Q

Suppose the columns of A form a basis of $\mathbb{R}^2$ (or $\mathbb{R}^n$):
- Q will contain the new orthonormal basis for the same space
- R will record the original columns of A expressed in this new basis

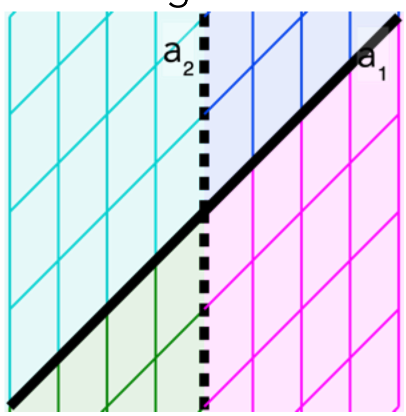
This is equivalent to
$Q^T A = R$

Consider the same matrix shown on the previous page:

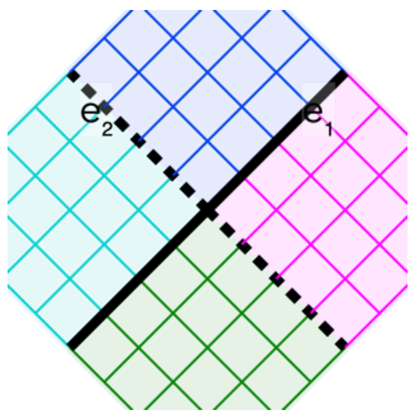
$$A = \left[\begin{array}{c|c} 1 & 0 \\ \hline 1 & 2 \end{array}\right]$$

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

$\mathbb{R}^2$ with columns of A serving as coordinate system or basis



$\mathbb{R}^2$ with orthonormal columns of Q serving as basis



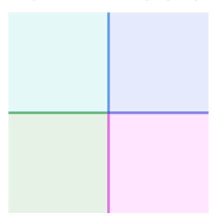
(Gram-Schmidt algorithm works so that
first orthonormal direction comes from the first column of A,
therefore e1 has the same direction as a1)

QRF of a 3×2 matrix

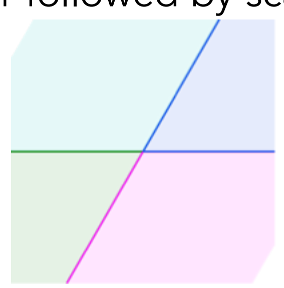
QR factorization of a rank-2 3×2 matrix

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

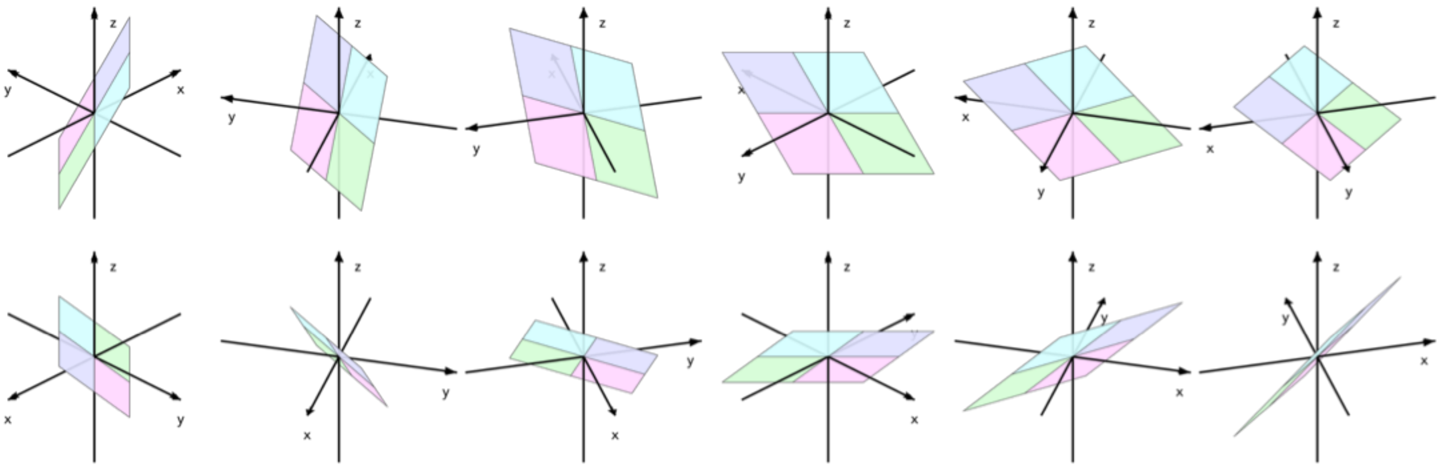
We visualize its domain $\mathbb{R}^2$ as a 4-color unit square



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

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

From equation 1, zeros appear everywhere below diagonal:

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

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:

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{bmatrix} \begin{bmatrix} \vec{x}_1 \\ \vec{x}_2 \\ \vec{x}_3 \end{bmatrix} = \begin{bmatrix} \vec{b}_1 \\ \vec{b}_2 \\ \vec{b}_3 \end{bmatrix}$$

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 ± 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