Algorithm endpoint

For algorithm illustration, we use an n×n matrix (n = 5)

$$A = \left[\begin{array}{c|c|c|c|c} \vec{a}_1 & \vec{a}_2 & \vec{a}_3 & \vec{a}_4 & \vec{a}_5 \end{array}\right] = \left[\begin{array}{c|c|c|c|c} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ \hline a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ \hline a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ \hline a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{array}\right]$$

The algorithm uses orthogonal reflection matrices $H_1 \ldots H_{n-1}$

Each H_i works as follows:

- $H_i \, \vec{a}_j = \vec{a}_j$ for all $j < i$

$\leftrightarrow$

H_i acts as identity on the first i-1 coordinates

$\downarrow$

H_i has the form $\left[\begin{array}{c|c} I_{(i-1)\times(i-1)} & 0_{(i-1)\times(n-i+1)} \\ \hline 0_{(n-i+1)\times(i-1)} & K_{i\,(n-i+1)\times(n-i+1)} \end{array}\right]$

- $H_i \, \vec{a}_i = \vec{a}_i'$
- $\left(\vec{a}_i'\right)_k = 0$ for all $k > i$

$\leftrightarrow$

$\vec{a}_i'$ lies in the span $\left\{ \vec{e}_1 \ldots \vec{e}_i \right\}$

$\downarrow$

$$H_{n-1} \dots H_1 A = \left[\begin{array}{c|c|c|c|c} \vec{a}_1' & \vec{a}_2' & \vec{a}_3' & \vec{a}_4' & \vec{a}_5' \end{array} \right] = R$$

R is upper-triangular

Target vector $\vec{a}_i'$ is chosen so that

- entries 1 ... i-1 stay unchanged
- $\left(\vec{a}_i'\right)_k = 0$ for all $k > i$
- for H_i to function as a reflector,

$\left(\vec{a}_i'\right)_i$ is adjusted so that $|\,\vec{a}_i'\,| = |\,\vec{a}_i\,|$

$\downarrow$

$$\vec{a}_i{}' = \begin{bmatrix} (a_i)_1 \\ \vdots \\ (a_i)_{i-1} \\ \pm\sqrt{\left((\vec{a}_i)_i\right)^2 + \ldots + \left((\vec{a}_i)_n\right)^2} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

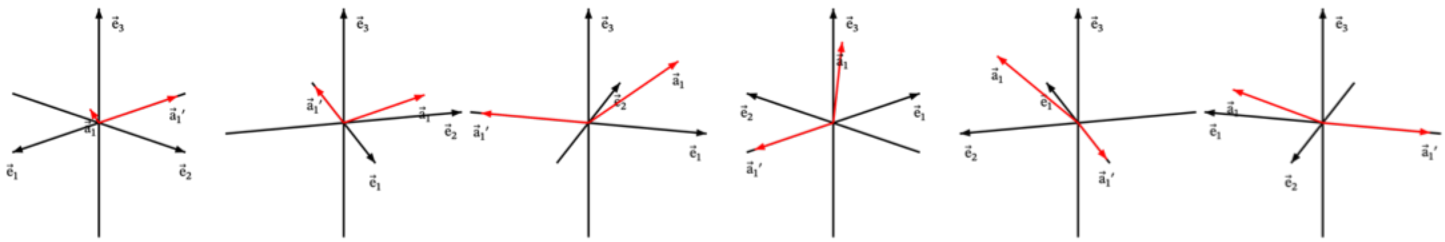
The sign of $(\vec{a}_i{}')_i$ is usually chosen opposite to $(\vec{a}_i)_i$

to use the longer reflector vector $\vec{v} = \vec{a}_i{}' - \vec{a}_i$ and minimize numerical error

$$\text{Numerical example: } A = \begin{bmatrix} 1.2 & 0.4 & 1.1 \\ 1 & 1.6 & -0.3 \\ 0.8 & -0.7 & 1.4 \end{bmatrix}$$

$$\vec{a}_i{}' = \begin{bmatrix} \approx -1.755 \\ 0 \\ 0 \end{bmatrix}$$

- is in the span $\{\vec{e}_1\}$
- sign of $(\vec{a}_1{}')_1$ is opposite to $(\vec{a}_1)_1$
- $|\vec{a}_1{}'| = |\vec{a}_1|$



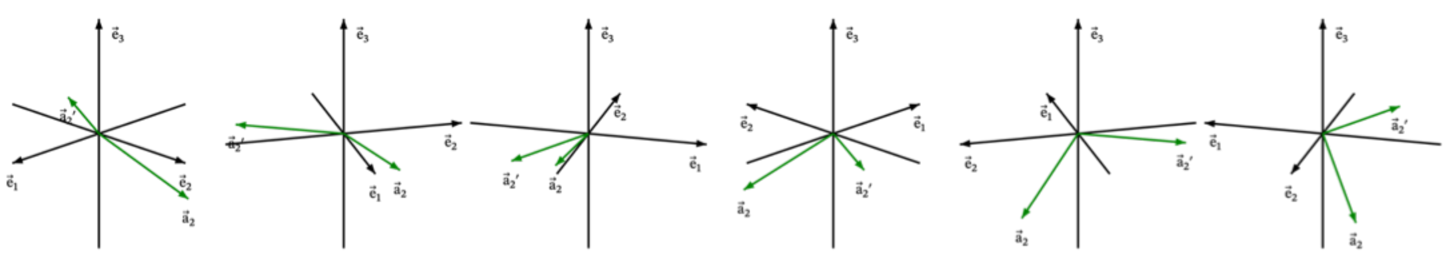
While further steps to obtain H_1 will be shown later,

next set of images shows how $\vec{a}_2{}'$ is obtained from $H_1 A$

$$H_1 A = \begin{bmatrix} \approx -1.755 & \approx -0.866 & \approx -1.219 \\ 0 & \approx 1.172 & \approx -1.085 \\ 0 & \approx -1.043 & \approx 0.772 \end{bmatrix}$$

$$\vec{a}_2{}' = \begin{bmatrix} \approx -0.866 \\ \approx -1.568 \\ 0 \end{bmatrix}$$

- is in the span $\{\vec{e}_1, \vec{e}_2\}$
- $\left(\vec{a}_2{}'\right)_1 = \left(\vec{a}_2\right)_1$
- sign of $\left(\vec{a}_2{}'\right)_2$ is opposite to $\left(\vec{a}_2\right)_2$
- $|\vec{a}_2{}'| = |\vec{a}_2|$



Reflector matrix H_i:

derivation from the target $\vec{a}_i'$

Reflection across fixed subspace S in $\mathbb{R}^n$

- $\mathbb{R}^n = S \oplus S\perp$
- for any $\vec{x} \in \mathbb{R}^n$:

$\vec{x}\| \in S$ is fixed

$\vec{x}\perp \in S\perp$ is negated

↓

Reflection matrix $H = I - 2P$

where P is projection onto the negated subspace $S\perp$

- If $\vec{y} = H\,\vec{x}$, then $H\,\vec{y} = \vec{x}$

$|\,\vec{a}\,| = |\,\vec{a}'\,|$

↓

there should be a matrix H that acts as a reflection:

- $H\,\vec{a} = \vec{a}'$
- $H\,\vec{a}' = \vec{a}$

↓

- $(I - 2P)\,\vec{a} = \vec{a}'$
- $(I - 2P)\,\vec{a}' = \vec{a}$

↓

- $2P\,\vec{a} = \vec{a} - \vec{a}'$
- $2P\,\vec{a}' = \vec{a}' - \vec{a}$

① Adding the equations gives

$$2P\,\vec{a} + 2P\,\vec{a}' = \vec{0}$$

↓

$$P\,(\vec{a} + \vec{a}') = \vec{0}$$

↓

$\vec{a} + \vec{a}'$ is in null(P) or in the fixed subspace

② Subtracting the first equation from the second gives

$$2P\,\vec{a}' - 2P\,\vec{a} = 2\vec{a}' - 2\vec{a}$$

↓

$$2P\,(\vec{a}' - \vec{a}) = 2\,(\vec{a}' - \vec{a})$$

↓

$$P\,(\vec{a}' - \vec{a}) = \vec{a}' - \vec{a}$$

↓

P acts as identity on $\vec{a}' - \vec{a}$

For projection P onto $S\perp$,
$P\,\vec{x} = \vec{x}$ means $\vec{x} \in S\perp$

$$\vec{a}' - \vec{a} \in S\perp$$

↓

P projects onto span($\vec{a}' - \vec{a}$)

if $\vec{v} = \vec{a}' - \vec{a}$,

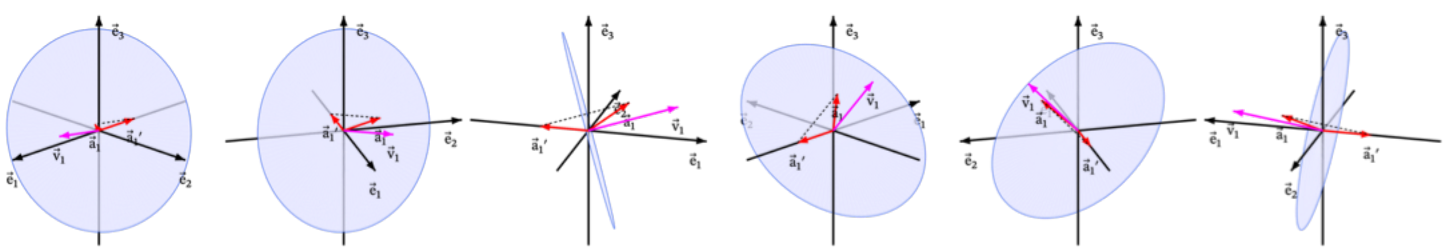
$$P = \frac{\vec{v}\,\vec{v}^T}{\vec{v}^T\,\vec{v}}$$

↓

$$\text{reflector } H = I - \frac{2\,\vec{v}\,\vec{v}^T}{\vec{v}^T\,\vec{v}}$$

$$\text{For col 1 of the same } A = \begin{bmatrix} 1.2 & 0.4 & 1.1 \\ 1 & 1.6 & -0.3 \\ 0.8 & -0.7 & 1.4 \end{bmatrix}$$

$$\vec{a}_i' = \begin{bmatrix} \approx -1.755 \\ 0 \\ 0 \end{bmatrix} \quad \vec{v} = \vec{a}' - \vec{a} = \begin{bmatrix} \approx 2.955 \\ 1 \\ 0.8 \end{bmatrix}$$



- the negated subspace $S\perp$ is always one-dimensional: span($\vec{v}$)
- the fixed subspace S is the n-1 dimensional orthogonal complement (plane in $\mathbb{R}^3$)
- for computation, only $\vec{v}$ is used

$$H = I - \frac{2\,\vec{v}\,\vec{v}^T}{\vec{v}^T\,\vec{v}} = \begin{bmatrix} \approx -0.684 & \approx -0.57 & \approx -0.456 \\ \approx -0.57 & \approx 0.807 & \approx -0.154 \\ \approx -0.456 & \approx -0.154 & \approx 0.877 \end{bmatrix}$$

Summary

Householder QR step summary

At step i:

- choose $\vec{a}_i'$ so that entries below i are zero

- keep $|\vec{a}_i'| = |\vec{a}_i|$ so reflection is possible

- form $\vec{v}_i = \vec{a}_i' - \vec{a}_i$

- build $H_i = I - \dfrac{2\vec{v}_i \vec{v}_i^T}{\vec{v}_i^T \vec{v}_i}$

Then

$$H_i \vec{a}_i = \vec{a}_i'$$

and H_i leaves the previous columns unchanged

After all steps,

$$H_{n-1} \ldots H_1 A = R$$

where R is upper-triangular

↓

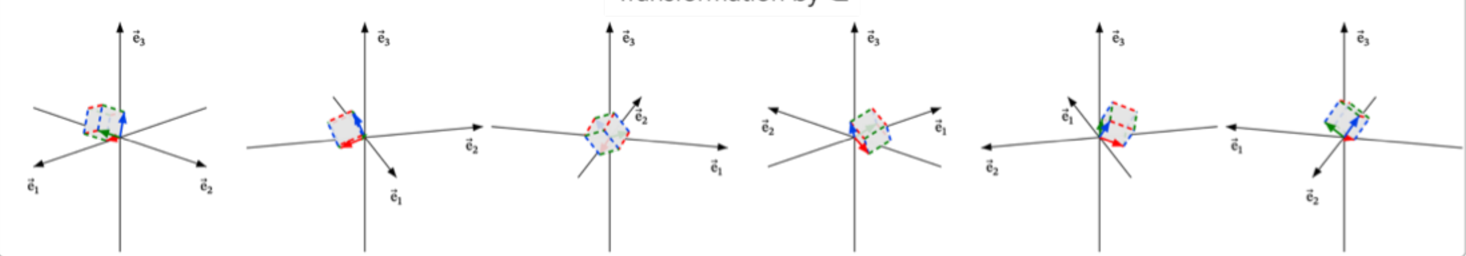
$$A = Q R$$

with $Q = \left(H_{n-1} \ldots H_1\right)^T$

Three geometric routes to QR decomposition:
• Gram-Schmidt orthogonalization
• Givens rotation
• Householder reflection

$$A = \begin{bmatrix} -0.34 & 1.21 & -2.41 \\ -2 & 0.38 & -1.62 \\ -0.89 & 0.88 & 0.03 \end{bmatrix} = Q\,R = \begin{bmatrix} \approx -0.153 & \approx 0.848 & \approx -0.507 \\ \approx -0.903 & \approx -0.329 & \approx -0.277 \\ \approx -0.402 & \approx 0.415 & \approx 0.816 \end{bmatrix} \begin{bmatrix} \approx 2.215 & \approx -0.882 & \approx 1.82 \\ 0 & \approx 1.267 & \approx -1.499 \\ 0 & 0 & \approx 1.695 \end{bmatrix}$$

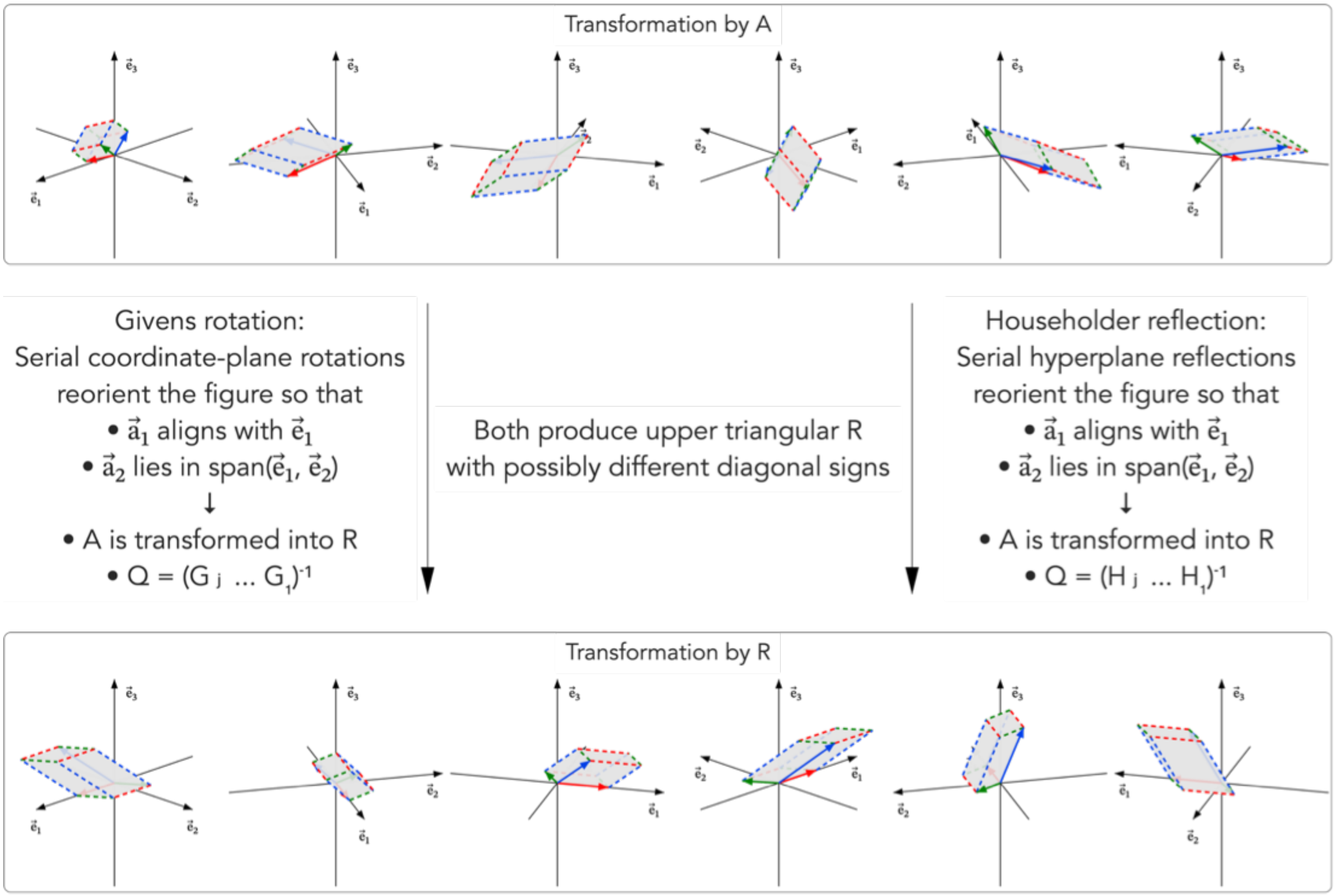
Transformation by Q



GS orthogonalization:
Serial column shear and scaling
reshape the figure so that
• columns become orthogonal
• each column has unit length

↓

• A is transformed into Q
• R is computed as $Q^T A$



Comparison of QR factorization algorithms

	Gram-Schmidt orthogonalization	Givens rotation	Householder reflection
Transforms A into	Q	R	R
Second factor computed as	$R = Q^T A$	$Q = (G_k \dots G_1)^{-1}$	$Q = (H_k \dots H_1)^{-1}$
Algebraic step operation	right-multiplication by scaling and shear matrices	left-multiplication by rotation matrices	left-multiplication by reflection matrices
Geometric step operation	• subtracts projections onto previous directions • normalizes the remaining component	rotates in a coordinate plane	reflects across a hyperplane
Singular A: geometric behavior	dependent column ↓ 0 residual	flat shape rotated into place	flat shape reflected into place
Sign of diagonal entries R_{ii}	non-negative	any	any
Default dimensions of Q & R for A(m×n, rank-r)	• Q is m×r • R is r×n	• Q is m×m • R is m×n	• Q is m×m • R is m×n
Numerical stability	small residual directions amplify error	stable	stable
Efficiency	less efficient	efficient for sparse A	preferred for dense A