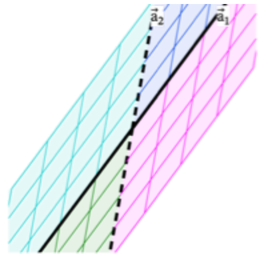
Columns during orthogonalization (2×2)

$$A = \begin{bmatrix} 1.2 & 0.2 \\ 1.6 & 1.1 \end{bmatrix}$$

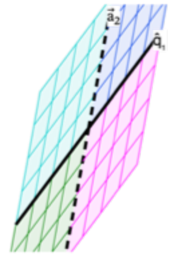
$$\begin{bmatrix} 1.2 & 0.2 \\ 1.6 & 1.1 \end{bmatrix}$$



column 1 normalized:
A2 = A1 × D1

$\vec{a}_1$ scaled so that $\hat{q}_1 = \vec{a}_1 / | \vec{a}_1 |$

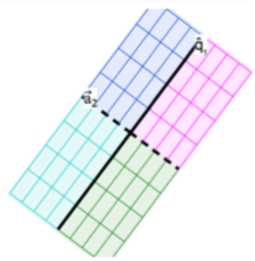
$$\begin{bmatrix} 0.6 & 0.2 \\ 0.8 & 1.1 \end{bmatrix}$$



column 2 orthogonalized:
A3 = A2 × S2

$\vec{a}_2$ orthogonalized so that $\vec{a}_2 \cdot \hat{q}_1 = 0$

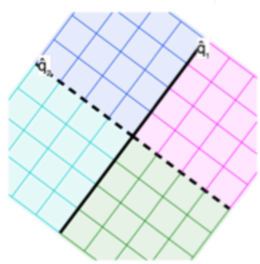
$$\begin{bmatrix} 0.6 & -0.4 \\ 0.8 & 0.3 \end{bmatrix}$$



column 2 normalized:
A4 = A3 × D2

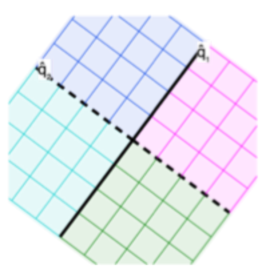
$\vec{a}_2$ scaled so that $\hat{q}_2 = \vec{a}_2 / | \vec{a}_2 |$

$$\begin{bmatrix} 0.6 & -0.8 \\ 0.8 & 0.6 \end{bmatrix}$$



Orthonormal: $Q = [\, \hat{q}_1 \mid \hat{q}_2 \,]$

$$\begin{bmatrix} 0.6 & -0.8 \\ 0.8 & 0.6 \end{bmatrix}$$



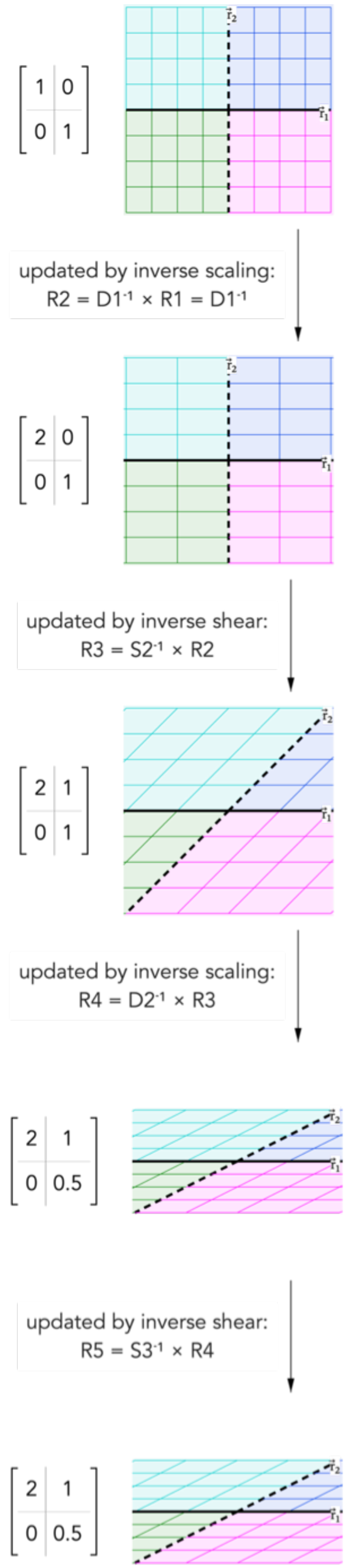
$$S1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad D1 = \begin{bmatrix} \frac{1}{|\vec{a}_1\perp|} & 0 \\ 0 & 1 \end{bmatrix} \quad \vec{a}_1\perp = \vec{a}_1$$

$$S2 = \begin{bmatrix} 1 & -\vec{q}_1^{\top}\vec{a}_2 \\ 0 & 1 \end{bmatrix} \quad D2 = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{|\vec{a}_2\perp|} \end{bmatrix} \quad \vec{a}_2\perp = \vec{a}_2 - (\vec{q}_1^{\top}\vec{a}_2)\,\vec{q}_1$$

$$Q = A4 = A \times D1 \times S2 \times D2$$

Stepwise construction of R during orthogonalization (2×2)

$$A = [\,\vec{a}_1 \mid \vec{a}_2\,] = \begin{bmatrix} 1.2 & 0.2 \\ 1.6 & 1.1 \end{bmatrix}$$



1 0
0 1
updated by inverse scaling:
R2 = $D1^{-1}$ × R1 = $D1^{-1}$
2 0
0 1
updated by inverse shear:
R3 = $S2^{-1}$ × R2
2 1
0 1
updated by inverse scaling:
R4 = $D2^{-1}$ × R3
2 1
0 0.5
updated by inverse shear:
R5 = $S3^{-1}$ × R4
2 1
0 0.5

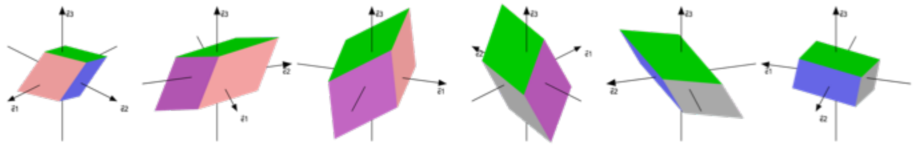
$$S1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad D1 = \begin{bmatrix} \frac{1}{|\vec{a}_1 \perp|} & 0 \\ 0 & 1 \end{bmatrix} \quad \vec{a}_1 \perp = \vec{a}_1$$

$$S2 = \begin{bmatrix} 1 & -\vec{q}_1^{\intercal} \vec{a}_2 \\ 0 & 1 \end{bmatrix} \quad D2 = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{|\vec{a}_2 \perp|} \end{bmatrix} \quad \vec{a}_2 \perp = \vec{a}_2 - (\vec{q}_1^{\intercal} \vec{a}_2)\, \vec{q}_1$$

Columns during orthogonalization (3×3)

$$A = \begin{bmatrix} 1.39 & 0.05 & -0.32 \\ 0.7 & 1.47 & 0.61 \\ 0.7 & 0.33 & 1.21 \end{bmatrix}$$

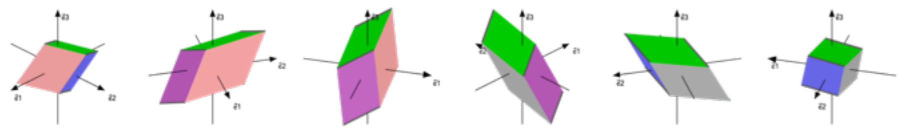
1.39	0.05	-0.32
0.7	1.47	0.61
0.7	0.33	1.21



column 1 normalized:
A2 = A1 × D1

$\vec{a}_1$ scaled so that $\hat{q}_1 = \vec{a}_1 / |\vec{a}_1|$

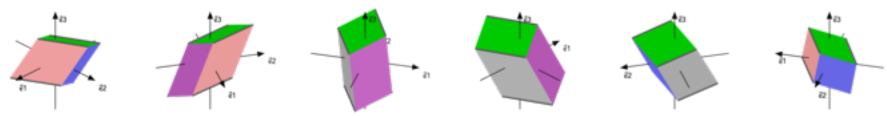
≈0.815	0.05	-0.32
≈0.41	1.47	0.61
≈0.41	0.33	1.21



column 2 orthogonalized:
A3 = A2 × S2

$\vec{a}_2$ orthogonalized so that $\vec{a}_2 \cdot \hat{q}_1 = 0$

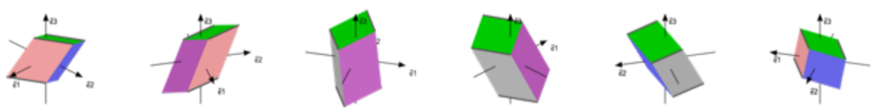
≈0.815	≈-0.585	-0.32
≈0.41	≈1.15	0.61
≈0.41	≈0.01	1.21



column 2 normalized:
A4 = A3 × D2

$\vec{a}_2$ scaled so that $\hat{q}_2 = \vec{a}_2 / |\vec{a}_2|$

≈0.815	≈-0.453	-0.32
≈0.41	≈0.891	0.61
≈0.41	≈0.008	1.21



column 3 orthogonalized:
A5 = A4 × S3

$\vec{a}_3$ orthogonalized so that $\vec{a}_3 \cdot \hat{q}_1 = 0$, $\vec{a}_3 \cdot \hat{q}_2 = 0$

≈0.815	≈-0.453	≈-0.399
≈0.41	≈0.891	≈-0.212
≈0.41	≈0.008	≈1.005

column 3 normalized:
A6 = A5 × D3

$\vec{a}_3$ scaled so that $\hat{q}_3 = \vec{a}_3 / |\vec{a}_3|$

≈0.815	≈-0.453	≈-0.362
≈0.41	≈0.891	≈-0.192
≈0.41	≈0.008	≈0.912

Orthonormal: $Q = [\, \hat{q}_1 \mid \hat{q}_2 \mid \hat{q}_3 \,]$

≈0.815	≈-0.453	≈-0.362
≈0.41	≈0.891	≈-0.192
≈0.41	≈0.008	≈0.912

$$S1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad D1 = \begin{bmatrix} \frac{1}{|\vec{a}_1\perp|} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \vec{a}_1\perp = \vec{a}_1$$

$$S2 = \begin{bmatrix} 1 & -\vec{q}_1^{\top} \vec{a}_2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad D2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{|\vec{a}_2\perp|} & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \vec{a}_2\perp = \vec{a}_2 - (\vec{q}_1^{\top} \vec{a}_2)\, \vec{q}_1$$

$$S3 = \begin{bmatrix} 1 & 0 & -\vec{q}_1^{\top} \vec{a}_3 \\ 0 & 1 & -\vec{q}_2^{\top} \vec{a}_3 \\ 0 & 0 & 1 \end{bmatrix} \quad D3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{|\vec{a}_3\perp|} \end{bmatrix} \quad \vec{a}_3\perp = \vec{a}_3 - (\vec{q}_1^{\top} \vec{a}_3)\, \vec{q}_1 - (\vec{q}_2^{\top} \vec{a}_3)\, \vec{q}_2$$

$$Q = A6 = A \times D1 \times S2 \times D2 \times S3 \times D3$$

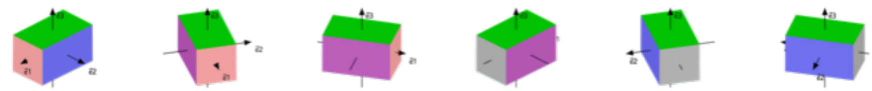
Stepwise construction of R during orthogonalization (3×3)

$$A = [\,\vec{a}_1 \mid \vec{a}_2 \mid \vec{a}_3\,] = \begin{bmatrix} 1.39 & 0.05 & -0.32 \\ 0.7 & 1.47 & 0.61 \\ 0.7 & 0.33 & 1.21 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

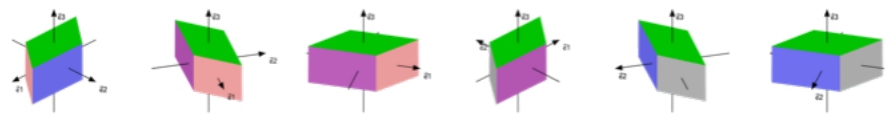
updated by inverse scaling:
R2 = D1⁻¹ × R1 = D1⁻¹

$$\begin{bmatrix} \approx 1.706 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$



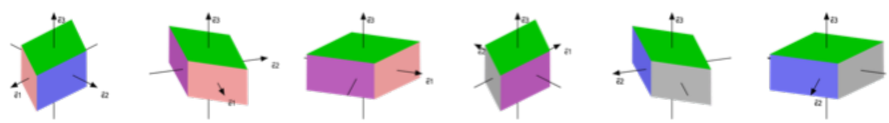
updated by inverse shear:
R3 = S2⁻¹ × R2

$$\begin{bmatrix} \approx 1.706 & \approx 0.779 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$



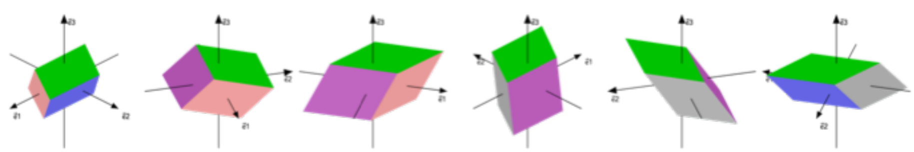
updated by inverse scaling:
R4 = D2⁻¹ × R3

$$\begin{bmatrix} \approx 1.706 & \approx 0.779 & 0 \\ 0 & \approx 1.29 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$



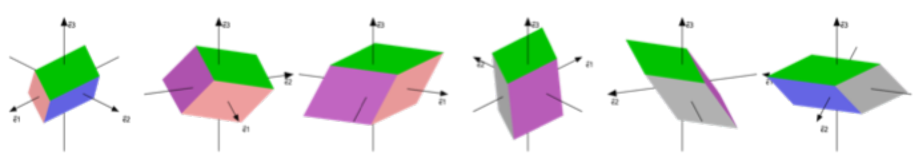
updated by inverse shear:
R5 = S3⁻¹ × R4

$$\begin{bmatrix} \approx 1.706 & \approx 0.779 & \approx 0.486 \\ 0 & \approx 1.29 & \approx 0.699 \\ 0 & 0 & 1 \end{bmatrix}$$

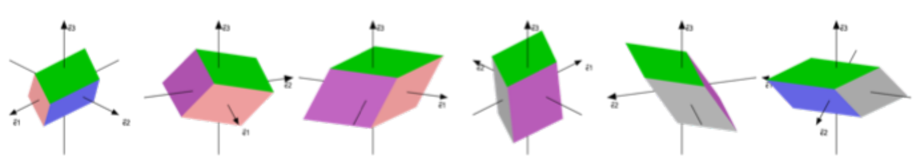


updated by inverse scaling:
R6 = D3⁻¹ × R5

$$\begin{bmatrix} \approx 1.706 & \approx 0.779 & \approx 0.486 \\ 0 & \approx 1.29 & \approx 0.699 \\ 0 & 0 & \approx 1.102 \end{bmatrix}$$



$$\begin{bmatrix} \approx 1.706 & \approx 0.779 & \approx 0.486 \\ 0 & \approx 1.29 & \approx 0.699 \\ 0 & 0 & \approx 1.102 \end{bmatrix}$$



$$S1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad D1 = \begin{bmatrix} \frac{1}{|\vec{a}_1\perp|} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \vec{a}_1\perp = \vec{a}_1$$

$$S2 = \begin{bmatrix} 1 & -\vec{q}_1^{\top}\vec{a}_2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad D2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{|\vec{a}_2\perp|} & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \vec{a}_2\perp = \vec{a}_2 - (\vec{q}_1^{\top}\vec{a}_2)\,\vec{q}_1$$

$$S3 = \begin{bmatrix} 1 & 0 & -\vec{q}_1^{\top}\vec{a}_3 \\ 0 & 1 & -\vec{q}_2^{\top}\vec{a}_3 \\ 0 & 0 & 1 \end{bmatrix} \quad D3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{|\vec{a}_3\perp|} \end{bmatrix} \quad \vec{a}_3\perp = \vec{a}_3 - (\vec{q}_1^{\top}\vec{a}_3)\,\vec{q}_1 - (\vec{q}_2^{\top}\vec{a}_3)\,\vec{q}_2$$

What do we learn from gradual orthogonalization

① Gram-Schmidt orthogonalization is shown as a transformation of space corresponding to the symbolic manipulation usually presented in teaching materials

② Each subsequent matrix A_{k+1} is obtained by right-multiplication

$A_{k+1} = A_k S_k D_k$ where

- for k ≥ 2, S_k is a shear matrix
  - $S_1 = [\ \vec{e}_1 \mid \vec{e}_2 \mid \vec{e}_3\ ]$
  - $S_2 = [\ \vec{e}_1 \mid \vec{s}_2 \mid \vec{e}_3\ ]$
  - $S_3 = [\ \vec{e}_1 \mid \vec{e}_2 \mid \vec{s}_3\ ]$
- D_k is a diagonal scaling matrix
  - $D_1 = [\ \vec{d}_1 \mid \vec{e}_2 \mid \vec{e}_3\ ]$
  - $D_2 = [\ \vec{e}_1 \mid \vec{d}_2 \mid \vec{e}_3\ ]$
  - $D_3 = [\ \vec{e}_1 \mid \vec{e}_2 \mid \vec{d}_3\ ]$

③ Right-multiplication changes columns because the columns of AB are A times the columns of B. For example:

$$Q = (A_3\ S_3)\ D_3$$

$$(A_3\ S_3)\ [\ \vec{e}_1 \mid \vec{e}_2 \mid \vec{d}_3\ ] = [\ A_3\ S_3\ \vec{e}_1 \mid A_3\ S_3\ \vec{e}_2 \mid A_3\ S_3\ \vec{d}_3\ ]$$

↓

since S_k and D_k differ from the identity in one active column, each step changes one column of A_k at a time

④ While the computationally efficient way to compute R is at the end of the algorithm as $R = Q^T A$, R is also the accumulated record of the inverse right-side updates used to orthogonalize the columns:

$$R_{k+1} = (S_k\ D_k)^{-1}\ R_k$$

⑤ The upper-triangular shape of R follows from the upper-triangular shape of every S_k, D_k and their inverses

⑥ Q is obtained by removing scale and shear from A as

$$A_{k+1} = A_k\ S_k\ D_k$$

⑦ R accumulates the removed scale and shear as

$$R_{k+1} = (S_k D_k)^{-1} R_k$$

⑧ The same idea can be applied to both columns and rows

• Column orthogonalization gives the standard

$$A = Q R$$

identity by serial right-multiplication

• Row orthogonalization gives the companion

$$A = L Q$$

identity by serial left-multiplication