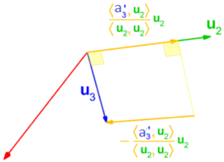
Algorithms and Data Structures

Matrix Decomposition
QR Decomposition



Learning goals

- QR decomposition
- Gram-Schmidt Pprocess

QR DECOMPOSITION

Given $\mathbf{A} \in \mathbb{R}^{n\times n}$. We decompose $\mathbf{A}$ into the product of an orthogonal matrix $\boldsymbol{Q} \in \mathbb{R}^{n\times n}$ and an upper triangular matrix $\boldsymbol{R} \in \mathbb{R}^{n\times n}$

$$\mathbf{A} = \boldsymbol{QR} \quad \text{with} \quad \boldsymbol{Q}^\top \boldsymbol{Q} = \boldsymbol{I},$$

The columns of the matrix $\boldsymbol{Q} = (\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n)$ form an orthonormal basis for the column space of the matrix $\mathbf{A}$ and

$$\boldsymbol{R} = \begin{pmatrix} \langle \boldsymbol{q_1}, \boldsymbol{a_1} \rangle & \langle \boldsymbol{q_1}, \boldsymbol{a_2} \rangle & \langle \boldsymbol{q_1}, \boldsymbol{a_3} \rangle & \cdots \\ 0 & \langle \boldsymbol{q_2}, \boldsymbol{a_2} \rangle & \langle \boldsymbol{q_2}, \boldsymbol{a_3} \rangle & \cdots \\ 0 & 0 & \langle \boldsymbol{q_3}, \boldsymbol{a_3} \rangle & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

The orthonormal basis for $\mathbf{A}$ is calculated by the Gram-Schmidt process.

GRAM-SCHMIDT PROCESS

The process takes a finite, linearly independent set of vectors and generates an orthogonal set of vectors that form an orthonormal basis.$^{(*)}$

Procedure: Projection: $\mathrm{proj}_{\boldsymbol{q}}\boldsymbol{a} = \frac{\langle \boldsymbol{q},\boldsymbol{a}\rangle}{\langle \boldsymbol{q},\boldsymbol{q}\rangle}\boldsymbol{q}$.

$$
\begin{aligned}
\boldsymbol{u}_1 &= \boldsymbol{a}_1 & \boldsymbol{q}_1 &= \frac{\boldsymbol{u}_1}{\|\boldsymbol{u}_1\|} \\
\boldsymbol{u}_2 &= \boldsymbol{a}_2 - \mathrm{proj}_{\boldsymbol{u}_1}\boldsymbol{a}_2 & \boldsymbol{q}_2 &= \frac{\boldsymbol{u}_2}{\|\boldsymbol{u}_2\|} \\
\vdots &= \vdots & \vdots &= \vdots \\
\boldsymbol{u}_k &= \boldsymbol{a}_k - \sum_{j=1}^{k-1}\mathrm{proj}_{\boldsymbol{u}_j}\boldsymbol{a}_k & \boldsymbol{q}_k &= \frac{\boldsymbol{u}_k}{\|\boldsymbol{u}_k\|}
\end{aligned}
$$

The vectors constructed in this way actually form an orthonormal basis of the column space of **A** (can be shown).

$^{(*)}$ If the vector $\boldsymbol{a}_i$ is not independent of $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{i-1}$, then $\boldsymbol{u}_i = \mathbf{0}$.

GRAM-SCHMIDT PROCESS / 2

$\boldsymbol{A}$ can now be represented by the calculated orthonormal basis:

$$
\begin{aligned}
\boldsymbol{a}_1 &= \boldsymbol{q}_1 \langle \boldsymbol{q}_1, \boldsymbol{a}_1 \rangle \\
\boldsymbol{a}_2 &= \boldsymbol{q}_1 \langle \boldsymbol{q}_1, \boldsymbol{a}_2 \rangle + \boldsymbol{q}_2 \langle \boldsymbol{q}_2, \boldsymbol{a}_2 \rangle \\
\vdots &= \vdots \\
\boldsymbol{a}_k &= \sum_{j=1}^{k} \boldsymbol{q}_j \langle \boldsymbol{q}_j, \boldsymbol{a}_k \rangle
\end{aligned}
$$

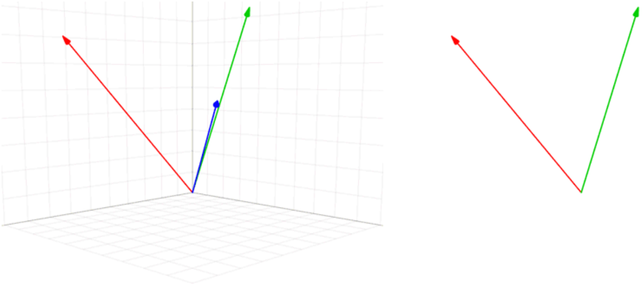
Or in matrix notation:

$$\boldsymbol{QR} = (\boldsymbol{q}_1 \langle \boldsymbol{q}_1, \boldsymbol{a}_1 \rangle, \boldsymbol{q}_1 \langle \boldsymbol{q}_1, \boldsymbol{a}_2 \rangle + \boldsymbol{q}_2 \langle \boldsymbol{q}_2, \boldsymbol{a}_2 \rangle, \cdots) = \boldsymbol{A}$$

GRAM-SCHMIDT VISUALIZED

Given: Three independent vectors a_1, a_2, a_3
Aim: Vectors of an orthonormal basis q_1, q_2, q_3

1. a_1 serves as the first vector of the orthogonal basis (u_1).
2. a_2 is projected onto u_1; projection is substracted from a_2 to obtain u_2.
3. a_3 is projected onto u_1 and u_2, to obtain u_3.
4. u_1, u_2 and u_3 are normalized.



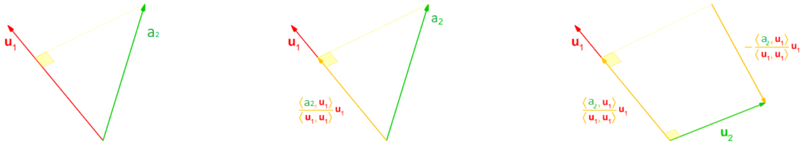
https://commons.wikimedia.org/wiki/File:Gram-Schmidt_orthonormalization_process.gif

GRAM-SCHMIDT VISUALIZED / 2

Given: Three independent vectors a_1, a_2, a_3

Aim: Vectors of an orthonormal basis q_1, q_2, q_3

1. a_1 serves as the first vector of the orthogonal basis (u_1).
2. a_2 is projected onto u_1; projection is substracted from a_2 to obtain u_2.
3. a_3 is projected onto u_1 and u_2, to obtain u_3.
4. u_1, u_2 and u_3 are normalized.

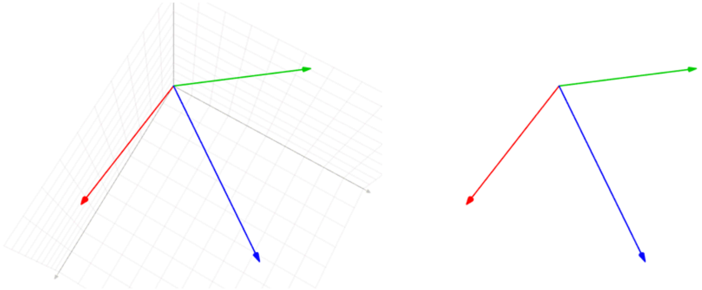


https://commons.wikimedia.org/wiki/File:Gram-Schmidt_orthonormalization_process.gif

GRAM-SCHMIDT VISUALIZED / 3

Given: Three independent vectors a_1, a_2, a_3
Aim: Vectors of an orthonormal basis q_1, q_2, q_3

1. a_1 serves as the first vector of the orthogonal basis (u_1).
2. a_2 is projected onto u_1; projection is substracted from a_2 to obtain u_2.
3. a_3 is projected onto u_1 and u_2, to obtain u_3.
4. u_1, u_2 and u_3 are normalized.



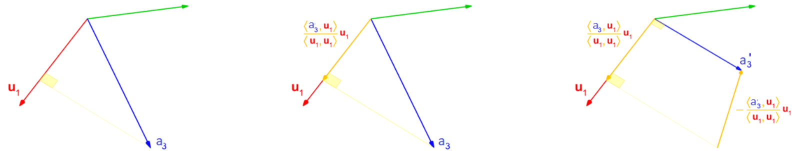
https://commons.wikimedia.org/wiki/File:Gram-Schmidt_orthonormalization_process.gif

GRAM-SCHMIDT VISUALIZED / 4

Given: Three independent vectors a_1, a_2, a_3

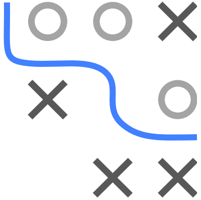
Aim: Vectors of an orthonormal basis q_1, q_2, q_3

1. a_1 serves as the first vector of the orthogonal basis (u_1).
2. a_2 is projected onto u_1; projection is substracted from a_2 to obtain u_2.
3. a_3 is projected onto u_1 and u_2, to obtain u_3.
4. u_1, u_2 and u_3 are normalized.



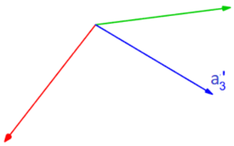
https://commons.wikimedia.org/wiki/File:Gram-Schmidt_orthonormalization_process.gif

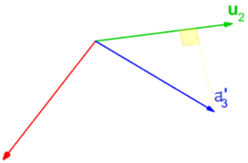
GRAM-SCHMIDT VISUALIZED / 5



Given: Three independent vectors a_1, a_2, a_3
Aim: Vectors of an orthonormal basis q_1, q_2, q_3

1. a_1 serves as the first vector of the orthogonal basis (u_1).
2. a_2 is projected onto u_1; projection is substracted from a_2 to obtain u_2.
3. a_3 is projected onto u_1 and u_2, to obtain u_3.
4. u_1, u_2 and u_3 are normalized.



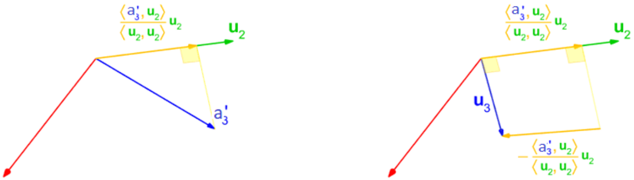


https://commons.wikimedia.org/wiki/File:Gram-Schmidt_orthonormalization_process.gif

GRAM-SCHMIDT VISUALIZED / 6

Given: Three independent vectors a_1, a_2, a_3
Aim: Vectors of an orthonormal basis q_1, q_2, q_3

1. a_1 serves as the first vector of the orthogonal basis (u_1).
2. a_2 is projected onto u_1; projection is substracted from a_2 to obtain u_2.
3. a_3 is projected onto u_1 and u_2, to obtain u_3.
4. u_1, u_2 and u_3 are normalized.

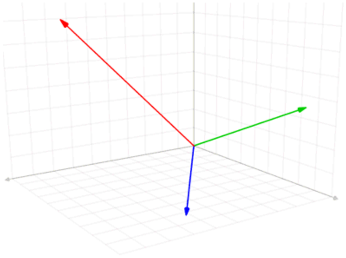


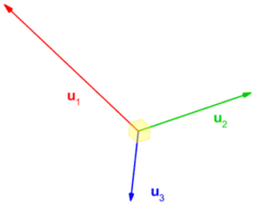
https://commons.wikimedia.org/wiki/File:Gram-Schmidt_orthonormalization_process.gif

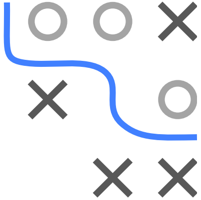
GRAM-SCHMIDT VISUALIZED / 7

Given: Three independent vectors a_1, a_2, a_3
Aim: Vectors of an orthonormal basis q_1, q_2, q_3

1. a_1 serves as the first vector of the orthogonal basis (u_1).
2. a_2 is projected onto u_1; projection is substracted from a_2 to obtain u_2.
3. a_3 is projected onto u_1 and u_2, to obtain u_3.
4. u_1, u_2 and u_3 are normalized.





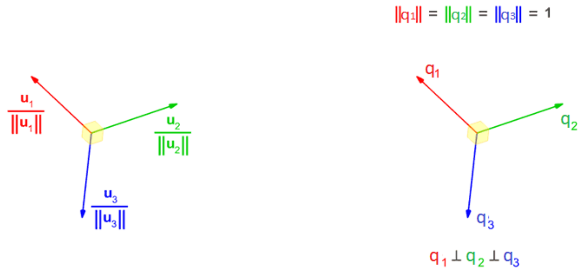


https://commons.wikimedia.org/wiki/File:Gram-Schmidt_orthonormalization_process.gif

GRAM-SCHMIDT VISUALIZED / 8

Given: Three independent vectors a_1, a_2, a_3
Aim: Vectors of an orthonormal basis q_1, q_2, q_3

1. a_1 serves as the first vector of the orthogonal basis (u_1).
2. a_2 is projected onto u_1; projection is substracted from a_2 to obtain u_2.
3. a_3 is projected onto u_1 and u_2, to obtain u_3.
4. u_1, u_2 and u_3 are normalized.



https://commons.wikimedia.org/wiki/File:Gram-Schmidt_orthonormalization_process.gif

QR DECOMPOSITION: EXAMPLE

Calculation of $\boldsymbol{A} = \boldsymbol{QR}$ with $\boldsymbol{A}$ given by

$$\mathbf{A} = \begin{pmatrix} 0 & -20 & -14 \\ 3 & 27 & -4 \\ 4 & 11 & -2 \end{pmatrix}$$

$k = 1$:

$$\boldsymbol{u}_1 = \boldsymbol{a}_1 = \begin{pmatrix} 0 \\ 3 \\ 4 \end{pmatrix}$$

$$\boldsymbol{q}_1 = \frac{\boldsymbol{u}_1}{\|\boldsymbol{u}_1\|} = \frac{\boldsymbol{u}_1}{\sqrt{0+9+16}} = \frac{1}{5}\begin{pmatrix} 0 \\ 3 \\ 4 \end{pmatrix}$$

$$r_{11} = \langle \boldsymbol{q}_1, \boldsymbol{a}_1 \rangle = \frac{1}{5}(0^2 + 3^2 + 4^2) = 5$$

$$r_{12} = \langle \boldsymbol{q}_1, \boldsymbol{a}_2 \rangle = \frac{1}{5}(0 \cdot (-20) + 3 \cdot 27 + 4 \cdot 11) = 25$$

$$r_{13} = \langle \boldsymbol{q}_1, \boldsymbol{a}_3 \rangle = \frac{1}{5}(0 \cdot (-14) + 3 \cdot (-4) + 4 \cdot (-2)) = -4$$

QR DECOMPOSITION: EXAMPLE / 2

$k = 2$:

$$\boldsymbol{u}_2 = \boldsymbol{a}_2 - \frac{\langle \boldsymbol{u}_1, \boldsymbol{a}_2 \rangle}{\langle \boldsymbol{u}_1, \boldsymbol{u}_1 \rangle} \boldsymbol{u}_1$$

$$= \boldsymbol{a}_2 - \frac{125}{25} \begin{pmatrix} 0 \\ 3 \\ 4 \end{pmatrix}$$

$$= \begin{pmatrix} -20 \\ 12 \\ -9 \end{pmatrix}$$

$$\boldsymbol{q}_2 = \frac{\boldsymbol{u}_2}{\|\boldsymbol{u}_2\|} = \frac{\boldsymbol{u}_2}{\sqrt{400 + 144 + 81}} = \frac{1}{25} \begin{pmatrix} -20 \\ 12 \\ -9 \end{pmatrix}$$

$$r_{22} = \langle \boldsymbol{q}_2, \boldsymbol{a}_2 \rangle = \frac{1}{25}((-20) \cdot (-20) + 12 \cdot 27 + (-9) \cdot 11) = 25$$

$$r_{23} = \langle \boldsymbol{q}_2, \boldsymbol{a}_3 \rangle = \frac{1}{25}((-20) \cdot (-14) + 12 \cdot (-4) + (-9) \cdot (-2)) = 10$$

QR DECOMPOSITION: EXAMPLE / 3

$k = 3$:

$$
\begin{aligned}
\boldsymbol{u}_3 &= \boldsymbol{a}_3 - \frac{\langle \boldsymbol{u}_1, \boldsymbol{a}_3 \rangle}{\langle \boldsymbol{u}_1, \boldsymbol{u}_1 \rangle} \boldsymbol{u}_1 - \frac{\langle \boldsymbol{u}_2, \boldsymbol{a}_3 \rangle}{\langle \boldsymbol{u}_2, \boldsymbol{u}_2 \rangle} \boldsymbol{u}_2 \\
&= \boldsymbol{a}_3 - \frac{-20}{25} \begin{pmatrix} 0 \\ 3 \\ 4 \end{pmatrix} - \frac{250}{625} \begin{pmatrix} -20 \\ 12 \\ -9 \end{pmatrix} \\
&= \begin{pmatrix} -6 \\ -6.4 \\ 4.8 \end{pmatrix} \\
\boldsymbol{q}_3 &= \frac{\boldsymbol{u}_3}{\|\boldsymbol{u}_3\|} = \frac{1}{25} \begin{pmatrix} -15 \\ -16 \\ 12 \end{pmatrix} \\
r_{33} &= \langle \boldsymbol{q}_3, \boldsymbol{a}_3 \rangle = \frac{1}{25}((-15) \cdot (-14) + (-16) \cdot (-4) + 12 \cdot (-2)) = 10
\end{aligned}
$$

QR DECOMPOSITION: EXAMPLE / 4

This results in

$$\mathbf{Q} = \frac{1}{25}\begin{pmatrix} 0 & -20 & -15 \\ 15 & 12 & -16 \\ 20 & -9 & 12 \end{pmatrix} \quad \text{and} \quad \mathbf{R} = \begin{pmatrix} 5 & 25 & -4 \\ 0 & 25 & 10 \\ 0 & 0 & 10 \end{pmatrix}.$$

HOUSEHOLDER AND GIVENS MATRIX

Problem in practice:

Q often not really orthogonal when using the above algorithm due to numerical reasons.

Two other methods for QR decomposition

Householder matrix:

For vector $\mathbf{u}$, matrix $\mathbf{U} = \mathbf{I} - d\mathbf{u}\mathbf{u}^\top$ is orthogonal, if $d = 2/\mathbf{u}^\top\mathbf{u}$.
Choose $\mathbf{u} = \mathbf{x} + s\mathbf{e}_1$ with $s = \mathbf{x}^\top\mathbf{x} \quad \Rightarrow \quad \mathbf{U}\mathbf{x} = -s\mathbf{e}_1$.

Successive elimination of column elements yields QR decomposition.

Givens matrix:

Similar to Householder, but orthogonal transformations that eliminate an element of a column vector each, and change a second vector.

For details see Carl D. Meyer *Matrix Analysis and Applied Linear Algebra*.

PROPERTIES OF QR DECOMPOSITION

- Splitting a matrix into an orthogonal matrix **Q** and **R**
- Gram-Schmidt process is numerically unstable, but can be extended and numerically stabilized
- **Existence:** Decomposition exists for each $n \times n$ matrix and can be extended to general $m \times n$, $m \neq n$ matrices
- Runtime behavior: Numerical stable solution of Householder transformation or Givens rotation comes along with higher effort:
  - Decomposition of $n \times n$ matrix using Householder transformation: $\approx \frac{2}{3}n^3$ multiplications
  - Forward and back substitution: n^2

COMPARISON OF METHODS

Procedure	**A**	# Multiplications	Stability
LU	regular	$\approx \frac{1}{3}n^3$	yes, by pivoting
Cholesky	p.d.	$\approx \frac{1}{6}n^3$	yes
QR (Gram Schmidt)	-	$\approx 2n^3$	no
QR (Householder)	-	$\approx \frac{2}{3}n^3$	yes

QR DECOMPOSITION FOR $M \times N$ MATRICES

General $m \times n$, $m \geq n$ matrices can be decomposed as well when using QR decomposition.

$$\mathbf{A} = \boldsymbol{Q}\boldsymbol{R} = \boldsymbol{Q} \begin{bmatrix} \boldsymbol{R}_1 \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \boldsymbol{Q}_1 & \boldsymbol{Q}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{R}_1 \\ \mathbf{0} \end{bmatrix} = \boldsymbol{Q}_1 \boldsymbol{R}_1$$

$\boldsymbol{Q}_1 \in \mathbb{R}^{m \times n}$, $\boldsymbol{Q}_2 \in \mathbb{R}^{m \times (m-n)}$ with orthogonal columns, and $\boldsymbol{R} \in \mathbb{R}^{n \times n}$ upper triangular matrix.

$\boldsymbol{Q}_1 \times \boldsymbol{R}_1$ is known as a **reduced** QR decomposition.