Médard Govoeyi

QR decomposition

August 18, 2025

Faculty of Mathematics
Institute of Analysis and Numerics
Otto-von-Guericke-Universität Magdeburg

DFG-Graduiertenkolleg
MATHEMATISCHE
KOMPLEXITÄTSREDUKTION

Orthogonal matrix and upper triangular matrix

Orthogonal matrix

Definition

An orthogonal matrix is a matrix whose columns are orthonormal. That is, matrices Q such that:

$$QQ^T = I.$$

Orthogonal matrix

Example

- *Rotation matrices,*

$$A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

- *Permutation matrices,*

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

- *Reflection matrices*

Orthogonal matrix

Some properties of orthogonal matrices

- If Q is orthogonal, $\det(Q) = \pm 1$
- Preservation of inner product and norms
  For any vectors $x, y \in \mathbb{R}^m$,
  $$\langle Qx, Qy \rangle = \langle x, y \rangle$$
  In particular,
  $$\|Qx\|_2 = \|x\|_2.$$
- The product of two orthogonal matrices is an orthogonal matrix
- Easy inverse:
  The inverse of an orthogonal matrix is its transpose:
  $$Q^{-1} = Q^\top$$

Orthogonal matrix

- Numerical stability
  Orthogonal matrices are very stable in numerical computations, since they do not change the norm: they neither amplify nor reduce errors.

Definiton

A matrix $R \in \mathbb{R}^{m \times n}$ is called an upper triangular matrix if

$$R = (r_{ij}), r_{ij} = 0 \text{ for all } i > j.$$

Definiton

A matrix $R \in \mathbb{R}^{m \times n}$ is called an upper triangular matrix if

$$R = (r_{ij}), r_{ij} = 0 \text{ for all } i > j.$$

Example (3×3 case)

$$R = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{pmatrix}$$

Some properties of upper triangular matrix

1. Closure under multiplication :
   The product of two upper triangular matrices is also upper triangular.
2. Efficient solving
   A system Rx=b can be solved by back substitution.
3. If R is a square matrix then

$$\det R = \prod_{i=1}^{n} r_{ii}$$

Since orthogonal matrices preserve lengths and angles, and upper triangular matrices facilitate the efficient solution of linear systems, for numerical computations and various matrix operations, it is therefore useful to express a general matrix as **the product of an orthogonal matrix and an upper triangular matrix**.

The QR decomposition

Let $A \in \mathbb{R}^{m \times n}$, $m \geq n$.

Theorem (Full QR decomposition)

There exists an orthogonal matrix $Q \in \mathbb{R}^{m \times m}$ and an upper triangular matrix $R \in \mathbb{R}^{m \times n}$ such that:

$$A = QR.$$

Let $A \in \mathbb{R}^{m \times n}$, $m \geq n$.

Theorem (Full QR decomposition)

There exists an orthogonal matrix $Q \in \mathbb{R}^{m \times m}$ and an upper triangular matrix $R \in \mathbb{R}^{m \times n}$ such that:

$$A = QR.$$

Theorem (Reduced QR decomposition)

If A has full column rank, there exists $\hat{Q} \in \mathbb{R}^{m \times n}$ with orthonormal columns and $\hat{R} \in \mathbb{R}^{n \times n}$ upper triangular such that:

$$A = \hat{Q}\hat{R}$$

This decomposition is unique up to signs of diagonal entries of $\hat{R}$.

Existence and Construction:Gram-Schmidt

Let $A = [a_1, a_2, \cdots, a_n]$, apply Gram-Schmidt to obtain q_i and scalars r_{ij} :
$Q = [q_1, q_2, \cdots, q_n]$,
For $j = 1$:

1. set :
$$r_{11} := \|a_1\|$$
2. Then
$$q_1 := \frac{a_1}{r_{11}}$$
3. General step: For $j = 2, \cdots n$:
   - Orthonormalization
   $$v_j := a_j - \sum_{i=1}^{j-1} r_{ij} q_i, \qquad \text{where } r_{ij} = \langle q_i, a_j \rangle.$$
   - Normalization:
   $$r_{jj} := \|v_j\|_2, \qquad q_j := \frac{v_j}{r_{jj}}$$

Example (Numeric instability of Gram-Schmidt method)

We consider the 3×3 *matrix* $A = [a_1, a_2, a_3]$, *where*

$$a_1 = \begin{pmatrix} 1 \\ \frac{1}{2} \\ \frac{1}{3} \end{pmatrix}, \quad a_2 = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{3} \\ \frac{1}{4} \end{pmatrix}, \quad a_3 = \begin{pmatrix} \frac{1}{3} \\ \frac{1}{4} \\ \frac{1}{5} \end{pmatrix}$$

After applying Gram-schmidt method to the matrix A *in half precision, we obtain*

$$Q = \begin{pmatrix} 0.857 & -0.4995 & 0.274 \\ 0.4285 & 0.569 & -0.7905 \\ 0.2856 & 0.653 & 0.548 \end{pmatrix},$$

and $\|QQ^T - I_3\|_2 = 0.33211513864862696$, *this error proves that the matrix* Q *is not orthogonal.*

Modified Gram-Schmidt (MGS)

The Classical Gram-Schmidt (CGS) process is theoretically correct in exact arithmetic, but suffers from loss of orthogonality in finite-precision computations due to the way round-off errors accumulate.
To address this, the Modified Gram-Schmidt (MGS) algorithm was developed. It performs orthogonalization in a more numerically stable way, by modifying the projection step to prevent error propagation.

Modified Gram-Schmidt (MGS)

Input: $A \in \mathbb{R}^{m \times n}$
Initialize: Set $v_j := a_j$
For $j = 1$ to n:

- **Normalize:**

$$r_{jj} := \|v_j\|_2, \quad q_j := \frac{v_j}{r_{jj}}$$

- **For** $i = j + 1$ to n:

$$r_{ji} := q_j^\top v_i, \quad v_i := v_i - r_{ji} q_j$$

In the Modified Gram-Schmidt (MGS) process, the vector v_i is updated immediately after projecting onto q_j. This avoids reusing a vector that has already accumulated projection errors:

$$v_i := v_i - \langle q_j, v_i \rangle q_j$$

Thus:

- Projections are more accurate
- Orthogonality of Q is preserved longer
- MGS is backward stable in many cases

Example

We consider as in the previous example the same matrix A, using the algorithm (MGS), we obtain

$$Q = \begin{pmatrix} 0.857 & -0.4995 & 0.1514 \\ 0.4285 & 0.569 & -0.661 \\ 0.2856 & 0.653 & 0.733 \end{pmatrix}.$$

And $\|QQ^\top - I\|_2 = 0.06270300839082639$

Compared to previous example , we achieve a (still large) relative error of 6.3%, which is, however, far smaller than the 33% we achieved with the original procedure.

CGS vs MCG

1. CGS and MGS are mathematically equivalent, but not numerically equivalent.
2. MGS is preferred when performing QR decomposition via Gram-Schmidt.
3. For large or ill-conditioned problems, Householder transformations remain the most robust choice.

Decomposition steps

- We proceed in the same way as the Gauss pivot

$$A = \begin{pmatrix} a_{ij} \end{pmatrix}$$

Decomposition steps

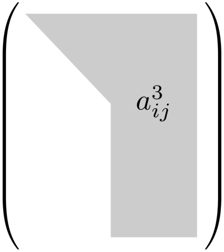
- We proceed in the same way as the Gauss pivot

$$H_1 A = \begin{pmatrix} & a_{ij}^2 \end{pmatrix}$$

Decomposition steps

- We proceed in the same way as the Gauss pivot

$$H_2 H_1 A \quad =$$



- But the matrices H_i are orthogonal matrices, so

$$H_i^{-1} = H_i^\top.$$

Decomposition steps

- We proceed in the same way as the Gauss pivot

$$\dots H_2 H_1 A = \begin{pmatrix} & r_{ij} \\ & \end{pmatrix}$$

- But the matrices H_i are orthogonal matrices, so

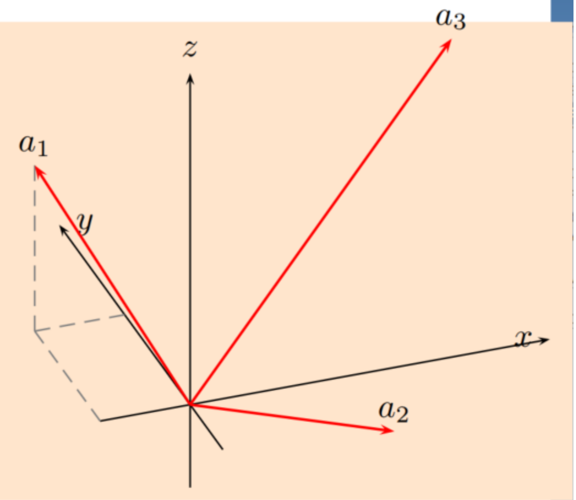
$$H_i^{-1} = H_i^\top .$$

- Then,

$$
\begin{aligned}
A &= (H_{n-1} \cdots H_1)^\top R \\
Q &= (H_{n-1} \cdots H_1)
\end{aligned}
$$

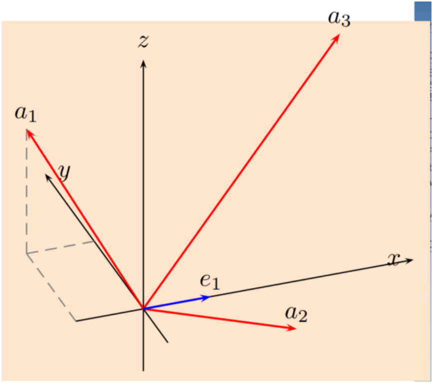
symmetries

- Let $a_1, a_2, \ldots, a_n$ be the column vectors of the matrix A.



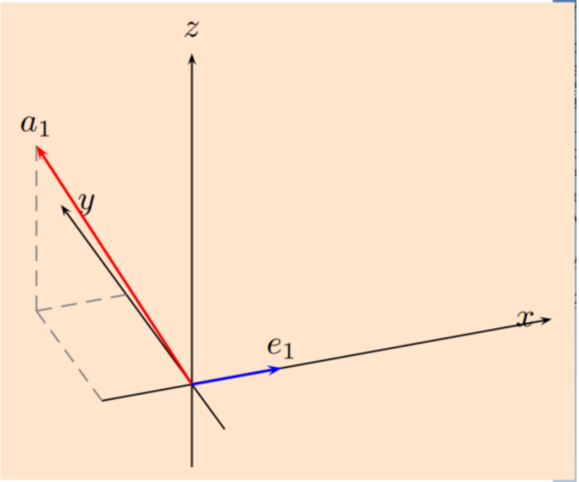
symmetries

- Let $a_1, a_2, \ldots, a_n$ be the column vectors of the matrix A.
- The matrix H_1 must map the vector a_1 to the first canonical basis vector e_1.



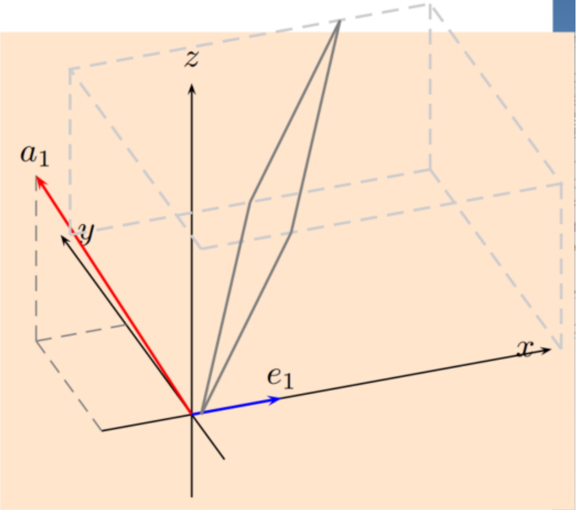
symmetries

- Let $a_1, a_2, \ldots, a_n$ be the column vectors of the matrix A.
- The matrix H_1 must map the vector a_1 to the first canonical basis vector e_1.
- To achieve this, we use the reflection with respect to the hyperplane that bisects the angle between these two vectors.



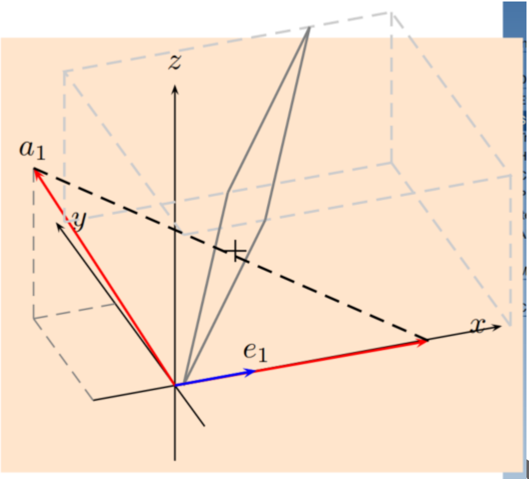
symmetries

- Let $a_1, a_2, \dots, a_n$ be the column vectors of the matrix A.
- The matrix H_1 must map the vector a_1 to the first canonical basis vector e_1.
- To achieve this, we use the reflection with respect to the hyperplane that bisects the angle between these two vectors.



symmetries

- Let $a_1, a_2, \ldots, a_n$ be the column vectors of the matrix A.
- The matrix H_1 must map the vector a_1 to the first canonical basis vector e_1.
- To achieve this, we use the reflection with respect to the hyperplane that bisects the angle between these two vectors.



Householder matrices

Definition (HOUSEHOLDER matrix)

A HOUSEHOLDER matrix is a matrix of the form:

$$H_u = I - 2uu^\top, \text{ where } u \in \mathbb{R}^n, \text{ with } \|u\| = 1.$$

Important properties

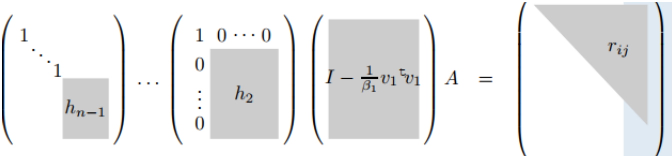
1. H_u is the orthogonal reflection with respect to the hyperplane orthogonal to a given vector u.
2. Symmetry
$$H_u^\top = H_u$$
3. Orthogonality
$$H_u^\top H_u = I$$
4. Involution
$$H_u^2 = I$$
This means H_u is its own inverse:
$$H_u^{-1} = H_u$$
5.
$$H_u u = -u \text{ and for any } v \perp u, H_u v = v.$$

HOUSEHOLDER transformation

1. Symmetry preserves the norm, so $H_1(a_1) = \pm\|a_1\| e_1$
2. Let $\alpha = \|a_1\|$, and $v = a_1 + \text{sign}(a_{11})\alpha e_1$, v is orthogonal to the symmetric plane.
3. $$\begin{aligned}\|v\|^2 &= (a_1 + \text{sign}(a_{11})\alpha e_1)^\top (a_1 + \text{sign}(a_{11})\alpha e_1) \\ &= (a_1^\top + \text{sign}(a_{11})\alpha e_1^\top)(a_1 + \text{sign}(a_{11})\alpha e_1) \\ &= a_1^\top a_1 + \text{sign}(a_{11})\alpha e_1^\top a_1 + \text{sign}(a_{11})\alpha a_1^\top e_1 + \alpha^2 e_1^\top e_1 \\ &= 2\alpha^2 + 2\text{sign}(a_{11})\alpha a_1^\top e_1\end{aligned}$$
4. Set $\beta = \frac{\|v\|^2}{2}$
5. $$\begin{aligned}H_1 &= I - \frac{2}{\|v\|^2} v v^\top \\ &= I - \frac{1}{\beta} v v^\top\end{aligned}$$

Then, we set $A_2 = H_1 A$.
To continue the process,

1. we consider the reduced matrix $\widetilde{A}_2$, which is obtained by removing the first row and the first column of A_2.
2. Then, we apply the same procedure to find a Householder matrix H_2 such that that is, H_2 maps the first column of $\widetilde{A}_2$ to a multiple of e_2.
3. We repeat this process with $\widetilde{A}_2, \widetilde{A}_3, \ldots \widetilde{A}_n$, until we obtain the upper triangular matrix $R = A_n$.



QR Decomposition via Householder Reflections

Require: Matrix $A \in \mathbb{R}^{m \times n}$
Ensure: Matrices $Q \in \mathbb{R}^{m \times m}$, $R \in \mathbb{R}^{m \times n}$ such that $A = QR$

$Q \leftarrow I_m$
$R \leftarrow A$
for $k = 1$ to $\min(m, n)$ **do**
    // Step 1: Extract the subvector $x \in \mathbb{R}^{m-k+1}$
    $x \leftarrow (R_{k,k}, R_{k+1,k}, \dots, R_{m,k})^T$
    // Step 2: Construct the Householder vector v
    $\alpha \leftarrow \|x\|_2$
    $v \leftarrow x + \mathrm{sign}(x_1)\alpha e_1$
    Normalize: $v \leftarrow v / \|v\|_2$
    // Step 3: Form the full $m \times m$ Householder matrix $H_k = I - 2vv^T$
    $H_k \leftarrow I_m$
    for $i = k$ to m **do**
        for $j = k$ to m **do**
            $H_k[i,j] \leftarrow H_k[i,j] - 2v[i-k+1]v[j-k+1]$
        end for
    end for
    // Step 4: Apply the Householder transformation
    $R \leftarrow H_k R$
    $Q \leftarrow Q H_k$
end for
return Q, R

Example

Let consider the same matrix A , by using householder transformations, we obtain

$$Q = \begin{pmatrix} -0.8571 & 0.5016 & -0.117 \\ -0.4286 & -0.5685 & 0.7022 \\ -0.2857 & -0.6421 & -0.7022 \end{pmatrix}$$

and the error

$$\|Q^\top Q - I\| = 4.10^{-8}$$

QR decomposition in overdetermined systems

Overdetermined systems

Let $A \in \mathbb{R}^{m \times n}$ with $m > n$, i.e., A has more rows than columns. Such a system is called *overdetermined*, and the linear system $Ax = b$ typically has no exact solution.

Instead, we seek the *least squares solution*, i.e., a vector $x \in \mathbb{R}^n$ that minimizes the residual norm:

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_2^2.$$

Existence and uniqueness

Theorem

If A has full column rank (i.e., $\operatorname{rank}(A) = n$), then the least squares solution exists and is unique, given by the normal equations:

$$x^* = (A^\top A)^{-1} A^\top b.$$

However, this approach can be numerically unstable due to the squaring of the condition number. To improve numerical stability, we use the QR decomposition.

Overdetermined systems

$$\|Ax - b\|_2^2 = \|QRx - b\|_2^2$$
$$= \|Rx - Q^T b\|_2^2$$

Then let, $R = \begin{pmatrix} \hat{R} \\ 0 \end{pmatrix}, \hat{R} \in \mathcal{M}_n(R)$, $Q^\top b = [c_1, c_2]^\top \quad c_1 \in \mathbb{R}^n, c_2 \in \mathbb{R}^{m-n}$. So ,

$$\|Ax - b\|_2^2 = \|c_1 - \hat{R}x\|_2^2 + \|c_2\|_2^2$$

The minimizer x^* satifies $\hat{R}x^* = c_1$.

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_2^2 = \|c_2\|^2$$

Overdetermined systems

$$\begin{aligned} Ax = b &\implies QRx = b \\ &\implies Rx = Q^\top b \\ &\implies \hat{R}x = c_1 \end{aligned}$$

Since R_1 is upper triangular and invertible (if A is full rank) , the unique least squares error solution:

$$x = \hat{R}^{-1} c_1$$

Overdetermined systems

- Avoids forming $A^\top A$, which can amplify numerical errors.
- More stable than using normal equations.
- Although the **QR decomposition is more expensive than solving the normal equations**, it is usually preferred in practice because of **its superior numerical stability.**

Conclusion

1. **Orthogonal and triangular matrices** have structural properties—orthogonality preserves lengths and angles; upper triangular form enables efficient back-substitution—that make numerical computations stable and well-conditioned.

Conclusion

1. **Orthogonal and triangular matrices** have structural properties—orthogonality preserves lengths and angles; upper triangular form enables efficient back-substitution—that make numerical computations stable and well-conditioned.
2. **Gram–Schmidt** yields a Q, R factorization but is **not numerically stable; Modified Gram–Schmidt (MGS) is significantly more stable**, yet **Householder reflections** provide the **most robust and stable** QR factorization in practice.

Conclusion

1. **Orthogonal and triangular matrices** have structural properties—orthogonality preserves lengths and angles; upper triangular form enables efficient back-substitution—that make numerical computations stable and well-conditioned.
2. **Gram–Schmidt** yields a Q, R factorization but is **not numerically stable; Modified Gram–Schmidt (MGS) is significantly more stable**, yet **Householder reflections** provide the **most robust and stable** QR factorization in practice.
3. This QR factorization via Householder transformations then enables an efficient and reliable solution of overdetermined systems (least-squares problems) by solving $Rx = Q^\top b$

Thank you for your attention !!!