



Dense linear algebra : QR factorization

P. Amestoy, P. Berger, M. Daydé, F.-H. Rouet
(INPT-ENSEEIH)
and
J.-Y. L'Excellent (INRIA/LIP-ENS Lyon)

QR Factorization

2019-2020

1 / 23

2 / 23

QR Factorization

Example

- ▶ Definition of an orthonormal set $\{x_1, \dots, x_k\}$
 - ▶ $x_i^T x_j = 0 \quad \forall i \neq j$
 - ▶ $x_i^T x_i = 1$
- ▶ Orthogonal matrix Q : columns of Q form an orthonormal set
 $Q^T Q = I, Q^{-1} = Q^T$
- ▶ QR Factorization:

$$\begin{array}{|c|} \hline \mathbf{A} \\ \hline \end{array} = \begin{array}{|c|} \hline \mathbf{Q} \\ \hline \end{array} \begin{array}{|c|} \hline \mathbf{R} \\ \hline \end{array}$$

- ▶ Q : orthogonal.
- ▶ R : upper trapezoidal.

$$\begin{bmatrix} 1 & -8 \\ 2 & -1 \\ 2 & 14 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} 3 & 6 \\ 0 & 15 \\ 0 & 0 \end{bmatrix} = Q \times R$$

3 / 23

4 / 23

- Factorization built by applying a succession of orthogonal transformations to the data:

$$Q = Q_1 \dots Q_n$$

with Q_i orthogonal matrices such that $Q^T A = R$

- Transformations:
 - Householder reflections.
 - Givens rotations.
 - Gram-Schmidt process (in this case $Q \in \mathcal{M}_{m,n}(\mathbb{K})$ and $R \in \mathcal{M}_n(\mathbb{K})$)

Exercise 1 (Orthogonal projection)

Let $Q \in \mathcal{M}_{m,n}(\mathbb{R})$, $m \geq n$ be such that $Q^T \cdot Q = I_n \in \mathcal{M}_n$. Let X be the n -dimensional span of the orthogonal set made of the columns of Q .

Show that $P_X = Q \cdot Q^T \in \mathcal{M}_m(\mathbb{R})$ is the orthogonal projection onto X .

5 / 23

6 / 23

QR Factorization - Preamble

Interpretation: $\mathbf{H}_u = \mathbf{I} - \frac{2}{\mathbf{u}^T \mathbf{u}} \mathbf{u} \mathbf{u}^T$.

Definition 1 (Householder matrix)

$\forall \mathbf{u} \in \mathbb{R}^n, \mathbf{u} \neq 0$, the **Householder matrix** $\mathbf{H}_u \in \mathcal{M}_{n,n}(\mathbb{R})$ is defined such that:

$$\mathbf{H}_u = \mathbf{I} - \frac{2}{\mathbf{u}^T \mathbf{u}} \mathbf{u} \mathbf{u}^T.$$

Proposition 1

\mathbf{H}_u is orthogonal ($\mathbf{H}_u^T \mathbf{H}_u = \mathbf{I}$) and symmetric.

Proof.

\mathbf{H}_u is obviously symmetric: $\mathbf{H}_u^T = \mathbf{H}_u$.

One has $\mathbf{H}_u^T \mathbf{H}_u = \left(\mathbf{I} - \frac{2}{\mathbf{u}^T \mathbf{u}} \mathbf{u} \mathbf{u}^T \right)^2 = \mathbf{I} + \frac{4}{(\mathbf{u}^T \mathbf{u})^2} \mathbf{u} \mathbf{u}^T \mathbf{u} \mathbf{u}^T - \frac{4}{\mathbf{u}^T \mathbf{u}} \mathbf{u} \mathbf{u}^T$.

Therefore, $\mathbf{H}_u^T \mathbf{H}_u = \mathbf{I} + \frac{4}{\mathbf{u}^T \mathbf{u}} \mathbf{u} \mathbf{u}^T - \frac{4}{\mathbf{u}^T \mathbf{u}} \mathbf{u} \mathbf{u}^T = \mathbf{I}$. □

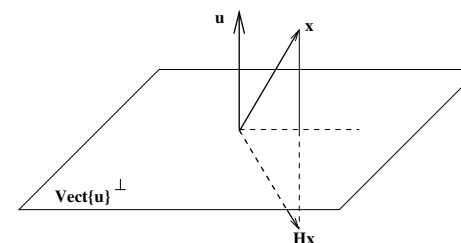
Proposition 2

- $\mathbf{H}_u \mathbf{u} = -\mathbf{u}$
- $\forall \mathbf{x} \in \mathbb{R}^n, \mathbf{H}_u \mathbf{x} = \mathbf{x}$ if \mathbf{x} and \mathbf{u} orthogonal.

Geometric interpretation:

$\forall \mathbf{x} \neq 0$, $\mathbf{H}_u \mathbf{x}$ is the symmetric of \mathbf{x} with respect to $\text{span}(\mathbf{u})^\perp$

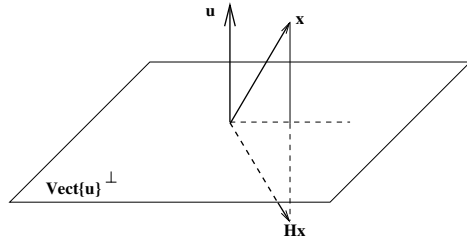
Therefore, \mathbf{H}_u is a reflection.



7 / 23

8 / 23

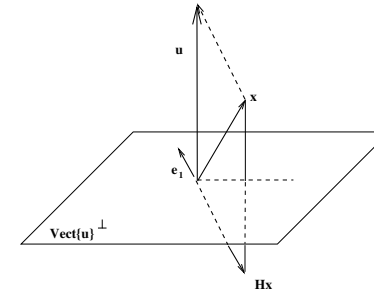
Another view on: $\mathbf{H}_u = \mathbf{I} - \frac{2}{\mathbf{u}^T \mathbf{u}} \mathbf{u} \mathbf{u}^T$.



$\mathbf{H}_u = \mathbf{I} - \frac{2}{\mathbf{u}^T \mathbf{u}} \mathbf{u} \mathbf{u}^T$: towards the QR factorization?

$\forall \mathbf{x} \in \mathbb{R}^n$ ($\mathbf{x} \neq 0$), can we find \mathbf{u} such that $\mathbf{H}_u \mathbf{x}$ and \mathbf{e}_1 collinear?

From a given \mathbf{x} , \mathbf{u} must be easy to compute.



Defining $\mathbf{u} = \mathbf{x} + \sigma \mathbf{e}_1$, we look for σ such that:

$$\mathbf{H}_u \mathbf{x} = \begin{bmatrix} -\sigma \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Let \mathbf{Q}_u be such that $\mathbf{Q}_u = \frac{\mathbf{u}}{\|\mathbf{u}\|_2}$.

From Exercice 1, $\mathbf{Q}_u \mathbf{Q}_u^T$ is a projection onto $\text{span}(\mathbf{u})$: $\frac{\mathbf{u} \mathbf{u}^T}{\mathbf{u}^T \mathbf{u}}$ is the orthogonal projection onto $\text{span}(\mathbf{u})$.

Therefore,

$\mathbf{H}_u = \mathbf{I} - 2\mathbf{Q}_u \mathbf{Q}_u^T$ is a symmetry with respect to $\text{span}(\mathbf{u})^\perp$.

9 / 23

10 / 23

Computation of σ / $\mathbf{H}_u \mathbf{x} = -\sigma \mathbf{e}_1$

Proposition 3

Let $\mathbf{x} \in \mathbb{R}^n$ be such that $\mathbf{x} \neq 0$, $\mathbf{x} = (x_i)_{i=1,n}$ and $\sigma = \text{sign}(x_1) \|\mathbf{x}\|$.

Defining $\mathbf{u} = \mathbf{x} + \sigma \mathbf{e}_1$, then

$\mathbf{H}_u \mathbf{x} = -\sigma \mathbf{e}_1$, with \mathbf{e}_1 the first vector of the standard basis.

Proof.

With $\mathbf{u} = \mathbf{x} + \sigma \mathbf{e}_1$ ($\mathbf{u} \neq 0$) we have

$$\mathbf{H}_u \mathbf{x} = \mathbf{H}_u \mathbf{u} - \sigma \mathbf{H}_u \mathbf{e}_1 = -\mathbf{u} - \sigma \left(\mathbf{e}_1 - \frac{2}{\mathbf{u}^T \mathbf{u}} \mathbf{u} \mathbf{u}^T \mathbf{e}_1 \right) = -\sigma \mathbf{e}_1 + \frac{2\sigma u_1}{\mathbf{u}^T \mathbf{u}} \mathbf{u} - \mathbf{u}.$$

Furthermore,

$$\mathbf{u}^T \mathbf{u} = \mathbf{x}^T \mathbf{u} + \sigma \mathbf{e}_1^T \mathbf{u} = \mathbf{x}^T (\mathbf{x} + \sigma \mathbf{e}_1) + \sigma (\sigma + x_1) = \mathbf{x}^T \mathbf{x} + 2\sigma x_1 + \sigma^2$$

Since $\mathbf{x}^T \mathbf{x} = \sigma^2$ and $u_1 = x_1 + \sigma$,

$$\mathbf{u}^T \mathbf{u} = 2\sigma (x_1 + \sigma) = 2\sigma u_1.$$

Therefore $\mathbf{H}_u \mathbf{x} = -\sigma \mathbf{e}_1$. \square

The choice of σ in Proposition 3 is motivated by $\mathbf{u}^T \mathbf{u} \sim 0$ if $\mathbf{x} \sim \mathbf{e}_1$

Householder reflections: an example

$\mathbf{H} = \mathbf{I} - 2\mathbf{v} \mathbf{v}^T$ with $\mathbf{v} \in \mathbb{R}^n$ such that $\|\mathbf{v}\|_2 = 1$.

\mathbf{H} is orthogonal and symmetric.

It provides a way of vanishing all the entries of a vector except for one component.

► Example :

$$\mathbf{x} = \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix} \quad \mathbf{u} = \mathbf{x} + \begin{bmatrix} \|\mathbf{x}\|_2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix} \quad \text{and } \mathbf{v} = \frac{\mathbf{u}}{\|\mathbf{u}\|_2}$$

Then,

$$\mathbf{H}_u = \mathbf{I} - 2\mathbf{v} \mathbf{v}^T = \frac{1}{15} \times \begin{bmatrix} -10 & 5 & -10 \\ 5 & 14 & 2 \\ -10 & 2 & 11 \end{bmatrix}$$

Therefore,

$$\mathbf{H}_u \mathbf{x} = \begin{bmatrix} -3 \\ 0 \\ 0 \end{bmatrix}$$

11 / 23

12 / 23

Theorem 1

Let $\mathbf{A} \in M_n(\mathbb{R})$ be a full rank matrix.

Then, there are $\mathbf{Q} \in M_n(\mathbb{R})$ orthogonal and $\mathbf{R} \in M_n(\mathbb{R})$ upper triangular such that

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

Proof.

see Exercise 2. \square

What about rectangular matrices?

The previous theorem can easily be extended to any matrix $\mathbf{A} \in M_{m,n}(\mathbb{R})$ $m \geq n$, with a rank equal to n . It can be shown that $\exists \mathbf{R} \in M_n(\mathbb{R})$ upper triangular such that

$$\mathbf{A} = \mathbf{Q} \begin{pmatrix} \mathbf{R} \\ 0 \end{pmatrix}$$

13 / 23

Householder Factorization: example

From a Householder vector: $\mathbf{u} = \mathbf{x} \pm \|\mathbf{x}\|_2 \mathbf{e}_1$, and its normalization $\mathbf{v} = \mathbf{u} / \|\mathbf{u}\|_2$ we can obtain a matrix that reads:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \\ 0 & a_{42} & a_{43} \\ 0 & a_{52} & a_{53} \end{bmatrix}$$

Let \mathbf{H} be such that:

$$\mathbf{H} \times \begin{bmatrix} a_{22} \\ a_{32} \\ a_{42} \\ a_{52} \end{bmatrix} = \begin{bmatrix} a'_{22} \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\text{If } \mathbf{H}' = \begin{bmatrix} 1 & 0 \\ 0 & \mathbf{H} \end{bmatrix} \text{ then } \mathbf{H}' \times \mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a'_{22} & a'_{23} \\ 0 & 0 & a'_{33} \\ 0 & 0 & a'_{43} \\ 0 & 0 & a'_{53} \end{bmatrix}$$

15 / 23

Exercise 2

Let $\mathbf{A} \in M_n(\mathbb{R})$ be a full rank matrix.

1. Show there exists $\mathbf{H}_1 \in M_n(\mathbb{R})$ orthogonal and $\alpha_1 \in \mathbb{R}$ such that $\mathbf{H}_1 \mathbf{A}(\mathbf{e}_1) = \alpha_1 \mathbf{e}_1$
2. Deduce from 1. that there exists a sequence of orthogonal matrices $\mathbf{H}_i, 1 \leq i \leq n-1$ such that $\mathbf{H}_{n-1} \dots \mathbf{H}_1 \mathbf{A} = \mathbf{R}$ with $\mathbf{R} \in M_n(\mathbb{R})$ triangular.

Definition 2

With $\mathbf{Q}^T = \mathbf{H}_{n-1} \dots \mathbf{H}_1$, exercise 2 leads to the Householder factorization $\mathbf{A} = \mathbf{Q}\mathbf{R}$

14 / 23

- Triangularization of a matrix 4×3 : $\mathbf{Q} = \mathbf{H}_3 \times \mathbf{H}_2 \times \mathbf{H}_1$

$$\begin{array}{l|l|l|l} | & x & x & x & | & | & x & x & x & | & | & x & x & x & | \\ | & x & x & x & | & \mathbf{H}_1 & | & 0 & x & x & | & \mathbf{H}_2 & | & 0 & x & x & | \\ | & x & x & x & | & \rightarrow & | & 0 & x & x & | & \rightarrow & | & 0 & 0 & x & | \\ | & x & x & x & | & & | & 0 & x & x & | & & | & 0 & 0 & x & | \end{array}$$

$$\begin{array}{l|l|l|l} & & & & | & x & x & x & | \\ & & & & \mathbf{H}_3 & | & 0 & x & x & | \\ \rightarrow & & & & | & 0 & 0 & x & | & = \mathbf{R} \\ & & & & | & 0 & 0 & 0 & | \end{array}$$

- QR backward stable, with a lower backward error than LU.

16 / 23

Exercise 3 (Memory consumption)

Compute the memory space required to do a QR factorization.

Exercise 4 (Computational complexity)

Estimate the amount of operations that are done during the Householder factorization.

Remark: the most expensive part of the algorithm occurs, at each step, when updating the reduce matrix $\mathbf{A}_k \in M_{n-k, n-k}(\mathbb{R})$.

$$\mathbf{A}_k = \mathbf{A}_k - \beta \mathbf{u}_k (\mathbf{u}_k^T \mathbf{A}_k) \quad \text{avec} \quad \beta = \frac{2}{\mathbf{u}_k^T \mathbf{u}_k} \quad \text{et} \quad \mathbf{u}_k \in \mathbb{R}^{n-k}. \quad (1)$$

Memory consumption

We do not compute \mathbf{H}_i : only the vectors \mathbf{u}_i are required.

The triangular part of \mathbf{A} is used to store these vectors (plus a vector of size n for the diagonal terms).

Therefore, the factorization can be done using the memory space allocated for the initial matrix (and a vector of size n).

Remark 1: Q-matrix product

Matrix product with \mathbf{Q} : sequence of Householder transformations.

Remark 2: Computational complexity

We need to specify the QR algorithm.

17 / 23

18 / 23

Householder QR factorization algorithm

Function: $[\mathbf{v}, \beta] = \text{house}(x)$

Compute \mathbf{v} and β such that $\mathbf{H} = \mathbf{I} - \beta \mathbf{v} \mathbf{v}^T$ is orthogonal, $v(1) = 1$ and $\mathbf{H}\mathbf{x} = \|\mathbf{x}\|_2 \mathbf{e}_1$
 $v(1)$ is not stored (known value). We need a vector of size n for β .

Algorithm

Let $\mathbf{A} \in M_{m,n}(\mathbb{R})$ $m \geq n$, be a rectangular matrix.
 The step computing the Householder matrices reads

```

for  $k = 1, n$  do
   $[\mathbf{v}, \beta] = \text{house}(\mathbf{A}(k : m, k))$ 
   $\mathbf{A}(k : m, k+1 : n) = (\mathbf{I}_{m-k+1} - \beta \mathbf{v} \mathbf{v}^T) \mathbf{A}(k : m, k+1 : n)$ 
   $\mathbf{A}(k, k) = -\text{sign}(\mathbf{A}(k, k)) \|\mathbf{A}(k : m, k)\|$ 
  if  $k \leq m$  then
     $\mathbf{A}(k+1 : m, k) = \mathbf{v}(2 : m-k+1)$ 
  end if
end for
```

19 / 23

Computational complexity of the QR factorization

Step factorizing $\mathbf{A} = \mathbf{QR}$ with $\mathbf{A} \in M_{m,n}(\mathbb{R})$

```

for  $k = 1, n$  do
  (1) Compute  $\beta$  (inner product)
  for  $j = 1, n-k$  do
    /* Let  $\text{Col}_j$  be the  $j$ th column of  $\mathbf{A}_k$  (size  $m-k$ )* /
  (2)  $\text{tmp} = \mathbf{v}_k^T \times \text{Col}_j$  ( $\mathbf{v}_k \in M_{m-k}(\mathbb{R})$ )
  (3)  $\text{tmp} = \beta \times \text{tmp}$ 
  (4)  $\text{Col}_j = \text{Col}_j - \text{tmp} \times \mathbf{v}_k$ 
  end for
end for
```

$2(m-k)$ flops for step (2) and (4), so $4(m-k)(n-k)$ flops per step k .
 It results in

$$4 \sum_{k=1}^n (m-k)(n-k) \simeq 4 \left(\frac{n^3}{3} + (m-n)n(n-1)/2 \right)$$

It leads to a computational complexity of $\frac{4}{3}n^3$ for square matrices of size n

- ▶ 2×2 rotation:

$$G(\theta) = \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \text{ orthogonal}$$

with $c = \cos(\theta)$ and $s = \sin(\theta)$.

- ▶ Application: $x = \{x_1, x_2\}$

$$c = \frac{x_1}{(x_1^2 + x_2^2)^{\frac{1}{2}}} \text{ et } s = \frac{-x_2}{(x_1^2 + x_2^2)^{\frac{1}{2}}}$$

$$y = (y_1, y_2) = G^T x \text{ then } y_2 = 0$$

- ▶ Specify the entries of a matrix that will vanish.

21 / 23

- ▶ Example: QR factorization

$$A = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \\ v_1 & v_2 & v_3 \end{bmatrix}$$

- ▶ We find (c, s) such that :

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix}^T \times \begin{bmatrix} r_{11} \\ v_1 \end{bmatrix} = \begin{bmatrix} r'_{11} \\ 0 \end{bmatrix}$$

22 / 23

- ▶ Rotation in the $(1,4)$ plane:

$$G(1,4) = \begin{bmatrix} c & 0 & 0 & s \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -s & 0 & 0 & c \end{bmatrix}$$

$$G(1,4)^T \times A = \begin{bmatrix} r'_{11} & r'_{12} & r'_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \\ 0 & v'_2 & v'_3 \end{bmatrix}$$

- ▶ Computational complexity: $2n^3$ flops for the triangularization of the matrix.

23 / 23