4.2. QR - decomposition

A QR decomposition describes the decomposition of a matrix into a product of two matrices with special properties. This decomposition exists for every matrix and can be calculated with different algorithms where the best known are:

- Gram–Schmidt
- Householder transformation
- Givens rotations

Definition 4.4. *Let $A \in \mathbb{R}^{m \times n}$ with $m \geq n$ be a matrix. The decomposition of A into a product $A = QR$ with an orthogonal matrix $Q \in \mathbb{R}^{m \times m}$ and an upper triangular matrix $R \in \mathbb{R}^{m \times n}$ is called a QR-decomposition of A.*

Considering the fact that $m \geq n$ and the matrix R is always quadratic, it often makes sense to partition both R and Q in a way such that the special structure of these matrices can be used advantageously. Since R is an upper triangular matrix, the last $m - n$ rows of matrix R consist only of zeros. Therefore it makes sense to split the matrix Q into two parts Q_1 and Q_2 which have n and $m - n$ columns respectively. The QR decomposition is reduced by the use of the special characteristics described above to:

$$\underbrace{A}_{m\times n} = \underbrace{Q}_{m\times m} \underbrace{\begin{bmatrix} R_1 \\ 0 \end{bmatrix}}_{m\times n} = [\underbrace{Q_1}_{m\times n} \quad \underbrace{Q_2}_{m\times(m-n)}] \underbrace{\begin{bmatrix} R_1 \\ 0 \end{bmatrix}}_{m\times n} = Q_1 R_1 \tag{4.11}$$

This notation is often referred to as reduced QR decomposition of A ([**trefethen1997numer**

Remark 4.15. *The QR - decomposition is unique if $rank(A) = n$ and the diagonal elements of R_1 are required to be positive.*

Remark 4.16. *Let $Q \in R^{n \times n}$ be an orthogonal matrix and $x \in \mathbb{R}^n$ a vector, then the following properties can be derived:*

- $\|Qx\|_2^2 = (Qx)^T(Qx) = x^T Q^T Q x = x^T I x = \|x\|_2^2$ *(length-invariant)*
- $Q^\top Q = I$

The QR-decomposition of matrix A can now be used to reduce the numerical instabilities that can occur in algorithm (4).

Remark 4.17. *Let A^P be the restricted matrix from algorithm (4) and QR the corresponding decomposition such that $A^P = QR$. Then steps e) and iv. from algorithm (4) can be written as:*

$$\begin{aligned} s^P &= ((A^P)^\top A^P)^{-1}(A^P)^\top b \\ &= ((QR)^\top QR)^{-1}(QR)^\top b \\ &= (R^\top Q^\top QR)^{-1}R^\top Q^\top b \\ &= R^{-1}(R^\top)^{-1}R^\top Q^\top b \\ &= R^{-1}Q^\top b \end{aligned} \tag{4.12}$$

A multiplication of (4.12) from the left with R results in a formula which is particularly easy for the calculation of the coefficients s_i^P.

$$Rs^P = Q^\top b \tag{4.13}$$

Since R is an upper triangular matrix, this system of equations can be solved very easily. Under the assumption that s^P has dimension $n \times 1$ and $Q^\top b$ is denoted as $\tilde{b}$, the parameters can be calculated by backward substitution following the rule:

$$\begin{aligned} s_n^P &= \frac{\tilde{b}}{r_{nn}} \\ s_i^P &= \frac{1}{r_{ii}}\left(\tilde{b}_i - \sum_{j=i+1}^{n} r_{ij}s_j^P\right) \qquad i = n-1, ..., 1 \end{aligned}$$

After showing how algorithm (4) benefits by using the QR decomposition, a way how such a decomposition can be calculated will be presented. In the following, the Householder-transformation, one of the most widespread methods, is derived.

4.2.1. Householder Transformation

The aim of the Householder transformation is to transform matrix A into an upper triangular matrix R by iterative multiplications of so-called Householder matrices H_i. The procedure is schematically shown in the following example:

Example 4.2. *Let $A \in \mathbb{R}^{m \times n}$, $m \geq n$ be the matrix for which a QR decomposition is to be performed. Let H_1 and H_2 be special Householder matrices. Then the QR decomposition is methodically calculated according to the following pattern.*

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ a_{3,1} & a_{3,2} & \cdots & a_{3,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix}$$

$$H_1 A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ 0 & a_{2,2} & \cdots & a_{2,n} \\ 0 & a_{3,2} & \cdots & a_{3,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ 0 & & & \\ 0 & & & \\ 0 & & A_2 & \\ 0 & & & \end{pmatrix}$$

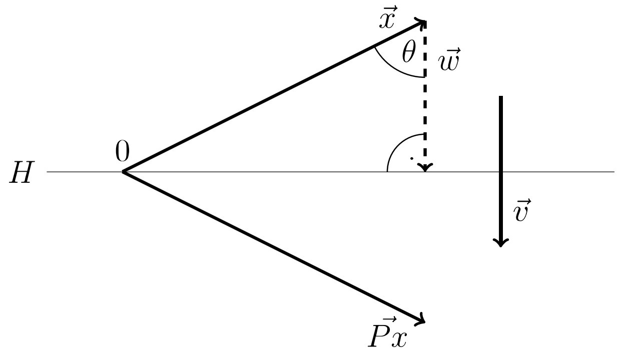
$$H_2 H_1 A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ 0 & a_{2,2} & \cdots & a_{2,n} \\ 0 & 0 & \cdots & a_{3,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{m,n} \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ 0 & a_{2,2} & \cdots & a_{2,n} \\ 0 & 0 & & \\ 0 & \vdots & A_3 & \\ 0 & 0 & & \end{pmatrix}$$

As shown in example (4.2), by applying the Householder matrix H_1 to matrix A, the first column of Matrix A is transformed to a multiple of the first unit vector. This transformation is implemented by a mirroring which is derived in the following section. After the transformation of the first column only the submatrix A_2 of the matrix $H_1 A$ is considered. This matrix consists of one row and one column less, but has a decisive advantage. Considering A_2 on its own, the first column can be mirrored to a multiple of the first unit vector as before and the same logic can be used iteratively. Altogether this means that the use of Householder matrices H_i iteratively generates an upper triangular matrix.

Remark 4.18. *Since the sub matrices (i.e. A_2, A_3, ...) that are to be transformed become in each step one row and one column smaller, this is also the case for the Householder matrices $\tilde{H}_i$. In order to preserve the transformations already carried out in the previous steps, the matrices $\tilde{H}_i$ are therefore enlarged in a way such that:*

$$H_i := \begin{bmatrix} I & 0 \\ 0 & \tilde{H}_i \end{bmatrix}$$

Figure 4.3.: Mirroring of $\vec{x}$ to $\vec{Px}$ through hyperplane H.



The task now is to find a matrix P that represents the desired reflection. To achieve this, a step-by-step approach is chosen. In a first step the reflection of a vector at a hyperplane through the origin in Euclidean space is constructed. Once this general case has been derived, it can be used to construct a reflection such that the the first column of matrix A is transformed to a multiple of the first unit vector.

For the construction of the mirroring matrix P the case shown in figure (4.3) is considered. Let $\vec{x}$ be a vector in an Euclidean space and $\vec{Px}$ the vector into which $\vec{x}$ is to be transferred by a mirroring. Furthermore, H is the hyperplane at which the reflection should take place. H, that mirror-hyperplane which runs through the origin is defined by the normal vector $\vec{v}$, thus a vector which is orthogonal to the hyperplane. The difference between the vector $\vec{x}$ and the hyperplane H is called $\vec{w}$. The angle enclosed by the vectors $\vec{-x}$ and $\vec{w}$ is called θ. The goal is to identify a relation between $\vec{Px}$ and $\vec{x}$. In the sense of better readability and since there can be no misunderstandings in the following, the vector arrows are omitted from now on. The length of $\vec{w}$ is then given by:

$$\|w\| = \|x\| \cos(\theta) = \|x\| \frac{\langle -x, v \rangle}{\|-x\| \|v\|} = \frac{-x^\top v}{\|v\|}$$

Where in the second step the definition of the dot product was used. The vector w is then characterized by the length and the direction which leads to:

$$w = \frac{-x^\top v}{\|v\|} \frac{v}{\|v\|} = -v \frac{x^\top v}{v^\top v}$$

By referencing to figure (4.3), the mirrored vector Px can now be defined as:

$$Px = x + 2w = x - 2v\frac{x^\top v}{v^\top v} = x - 2\frac{vv^\top x}{v^\top v} = \left(I - 2\frac{vv^\top}{v^\top v}\right)x$$

The matrix constructed, representing the linear mapping described above is called Householder matrix. Householder matrices are defined by a normal vector v, i.e. a vector that is orthogonal to the mirror hyperplane and are typically denoted by H.

$$H = I - 2\frac{vv^\top}{v^\top v} \tag{4.15}$$

Where I in equation (4.15) is the identity matrix. In case that v is normalized to length one (4.15) simplifies to:

$$H = I - 2vv^\top \tag{4.16}$$

The concept of Householder matrices can now be used to formalize the process described in example 4.2. Let x be a vector which is to be mirrored to a multiple of the first unit vector e_1. This means that a vector v is required, so that with the corresponding Householder-Matrix H_v the following linear transformation can be achieved.

$$H_v x = ce_1$$

The required reflection vector v now results from normalizing the difference vector and is given by:

$$v = \frac{x - ce_1}{\|x - ce_1\|}$$

A basic property that is important for the construction of a QR decomposition is the fact that Householder matrices are orthogonal.

Remark 4.19. *Let H be a Householder matrix. Then H is symmetric and orthogonal.*

$$H^\top = (I - 2\frac{vv^\top}{v^\top v})^\top = I^\top - \left(2\frac{vv^\top}{v^\top v}\right)^\top = I - 2\frac{\left(vv^\top\right)^\top}{\left(v^\top v\right)^\top} = I - 2\frac{vv^\top}{v^\top v} = H$$

$$
\begin{aligned}
H^\top H = HH &= (I - 2\frac{vv^\top}{v^\top v})(I - 2\frac{vv^\top}{v^\top v}) \\
&= I^2 - 2I\frac{vv^\top}{v^\top v} - 2I\frac{vv^\top}{v^\top v} + 4\frac{vv^\top}{v^\top v}\frac{vv^\top}{v^\top v} \\
&= I - 4\frac{vv^\top}{v^\top v} + 4\frac{vv^\top vv^\top}{v^\top vv^\top v} \\
&= I - 4\frac{vv^\top}{v^\top v} + 4\frac{(v^\top v)vv^\top}{(v^\top v)^2} \\
&= I
\end{aligned}
$$

Remark 4.20. *Let $H_1, H_2, ..., H_n$ be Householder matrices as stated in example 4.2. Then by using the properties proofed in the previous remark the QR decomposition is given by:*

$$
\begin{aligned}
& H_n H_{n-1} \cdot \ldots \cdot H_1 A = R \\
\Leftrightarrow \quad & Q^\top A = R \\
\Leftrightarrow \quad & QQ^\top A = QR \\
\Leftrightarrow \quad & A = QR
\end{aligned}
$$

After demonstrating how a QR decomposition can be performed using Householder transformations, the next section is dedicated to the question of performance.