MA4230 Matrix Computation

Chapter 3: QR factorization

Notation: upper-triangular matrix

<u>Note:</u> In this chapter, we restrict ourselves to "tall" matrices $A \in \mathbb{R}^{m\times n}$ with $m \geq n$.

Let $m, n \in \mathbb{N}$ with $m \geq n$. A matrix $R = (r_{ij}) \in \mathbb{R}^{m\times n}$ is called **upper-triangular** iff $r_{ij} = 0$ whenever $i > j$, i.e., iff

$$R = \begin{pmatrix} \hat{R} \\ \hline 0_{(m-n)\times n} \end{pmatrix} \in \mathbb{R}^{m\times n}, \quad \text{where} \quad \hat{R} = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & r_{nn} \end{pmatrix} \in \mathbb{R}^{n\times n}.$$

Examples of upper-triangular matrices

- an upper-triangular 4×4 matrix looks like $\begin{pmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & 0 & * \end{pmatrix}$,
- an upper-triangular 5×3 matrix looks like $\begin{pmatrix} * & * & * \\ 0 & * & * \\ 0 & 0 & * \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$,
- an upper-triangular 2×1 matrix looks like $\begin{pmatrix} * \\ 0 \end{pmatrix}$.

3.1 Definition of full and reduced QR factorization

Definition of QR factorization

Definition (QR factorization)

Let $m, n \in \mathbb{N}$ with $m \geq n$, and let $A \in \mathbb{R}^{m\times n}$. If there exist

$$Q = (q_1|\cdots|q_m) \ \in \mathbb{R}^{m\times m} \text{ orthogonal},$$
$$R = \begin{pmatrix} \hat{R} \\ \hline 0_{(m-n)\times n} \end{pmatrix} \in \mathbb{R}^{m\times n} \ \text{ upper-triangular},$$

such that there holds

$$A = QR,$$

then we call this a (full) **QR factorization** of A.

Reduced QR factorization

Suppose $A \in \mathbb{R}^{m\times n}$, $m \geq n$, has a QR factorization $A = QR$ with

$$Q = (q_1|\cdots|q_m) \quad \in \mathbb{R}^{m\times m} \text{ orthogonal,}$$
$$R = \begin{pmatrix} \hat{R} \\ \hline 0_{(m-n)\times n} \end{pmatrix} \in \mathbb{R}^{m\times n} \text{ upper-triangular.}$$

Observe:

$$A = QR = (q_1|\cdots|q_m)\begin{pmatrix} \hat{R} \\ \hline 0_{(m-n)\times n} \end{pmatrix} = (q_1|\cdots|q_n)\hat{R} =: \hat{Q}\hat{R}.$$

This is a reduced QR factorization of A in the sense of the following defn:

Definition: Given $A \in \mathbb{R}^{m\times n}$, $m \geq n$, we call a factorization $A = \hat{Q}\hat{R}$ with $\hat{Q} \in \mathbb{R}^{m\times n}$ having orthonormal columns and $\hat{R} \in \mathbb{R}^{n\times n}$ being upper-triangular a **reduced QR factorization** of A.

Example

An example of a QR factorization is

$$\begin{pmatrix} 1 & 1 \\ -1 & 1 \\ 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} & 0 & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{3} & \frac{1}{\sqrt{3}} \\ 0 & \frac{4}{\sqrt{6}} \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

with corresponding reduced QR factorization

$$\begin{pmatrix} 1 & 1 \\ -1 & 1 \\ 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \sqrt{3} & \frac{1}{\sqrt{3}} \\ 0 & \frac{4}{\sqrt{6}} \end{pmatrix}.$$

3.2 Existence and uniqueness

Towards a reduced QR factorization: an observation

Let $A \in \mathbb{R}^{m\times n}$, $m \geq n$. Finding a reduced QR factorization $A = \hat{Q}\hat{R}$ with $\hat{Q} \in \mathbb{R}^{m\times n}$ having orthonormal columns and $\hat{R} \in \mathbb{R}^{n\times n}$ upper-triangular,

$$A = (a_1|\cdots|a_n) = (q_1|\cdots|q_n)\begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & r_{nn} \end{pmatrix} = \hat{Q}\hat{R},$$

is equivalent to finding n orthonormal vectors $q_1, \ldots, q_n \in \mathbb{R}^m$ and $\frac{n(n+1)}{2}$ real numbers $\{r_{ij}\}_{1\leq i\leq j\leq n} \subseteq \mathbb{R}$ such that

$$\begin{cases} a_1 &= r_{11}q_1, \\ a_2 &= r_{12}q_1 + r_{22}q_2, \\ &\vdots \\ a_n &= r_{1n}q_1 + r_{2n}q_2 + \cdots + r_{nn}q_n. \end{cases}$$

$\Rightarrow$ find orthonormal $q_1, \ldots, q_n \in \mathbb{R}^m$ s.t. $a_i \in \operatorname{span}(q_1, \ldots, q_i)\ \forall 1 \leq i \leq n$.

Towards reduced QR: Gram–Schmidt orthogonalization

$\Rightarrow$ find orthonormal $q_1, \ldots, q_n \in \mathbb{R}^m$ s.t. $a_i \in \text{span}(q_1, \ldots, q_i)$ $\forall 1 \leq i \leq n$.

Now focus on the case $A = (a_1 | \cdots | a_n) \in \mathbb{R}^{m \times n}$, $m \geq n$, with $\text{rk}(A) = n$.

Gram–Schmidt orthogonalization is a method to find orthonormal vectors $q_1, \ldots, q_n \in \mathbb{R}^m$ s.t.

$$\text{span}(q_1, \ldots, q_i) = \text{span}(a_1, \ldots, a_i) \qquad \forall 1 \leq i \leq n.$$

First step is easy: Find a unit vector $q_1 \in \mathbb{R}^m$ s.t. $\text{span}(q_1) = \text{span}(a_1)$.

$$q_1 := \frac{a_1}{\|a_1\|_2} \quad \Longrightarrow \quad a_1 = r_{11} q_1 \quad \text{with} \quad r_{11} := \|a_1\|_2.$$

(Note $\text{rk}(A) = n \Longrightarrow a_1 \neq 0 \in \mathbb{R}^m \Longrightarrow \|a_1\|_2 > 0$.)

If $n = 1$, done. If $n \geq 2$:

Let $A = (a_1|\cdots|a_n) \in \mathbb{R}^{m\times n}$, $m \geq n \geq 2$, $\mathrm{rk}(A) = n$.

Suppose we have found orthonormal $q_1, \dots, q_{k-1} \in \mathbb{R}^m$ $(2 \leq k \leq n)$ s.t.

$$\mathrm{span}(q_1, \dots, q_i) = \mathrm{span}(a_1, \dots, a_i) \quad \forall 1 \leq i \leq k-1.$$

Then, define

$$q_k := \pm\frac{\tilde{q}_k}{\|\tilde{q}_k\|_2}, \quad \text{where} \quad \tilde{q}_k := a_k - \sum_{l=1}^{k-1}\langle q_l, a_k\rangle q_l.$$

Note

- $\tilde{q}_k \neq 0 \in \mathbb{R}^m$ (Pf: $\tilde{q}_k = 0 \implies a_k \in \mathrm{span}(q_1, \dots, q_{k-1}) \implies a_k \in \mathrm{span}(a_1, \dots, a_{k-1})$, contradiction to $\mathrm{rk}(A) = n$.)
- $\|q_k\|_2 = 1, \quad \{q_k\} \perp \{q_1, \dots, q_{k-1}\}$.
- $q_k \in \mathrm{span}(a_1, \dots, a_k)$, $a_k \in \mathrm{span}(q_1, \dots, q_k) \implies \mathrm{span}(q_1, \dots, q_k) = \mathrm{span}(a_1, \dots, a_k)$.

$\implies$ Have $q_1, \dots, q_k \in \mathbb{R}^m$ orthonormal and $\mathrm{span}(q_1, \dots, q_i) = \mathrm{span}(a_1, \dots, a_i) \quad \forall 1 \leq i \leq k$. Done (iterate)!

Recall from previous slide:

$$q_k := \pm\frac{\tilde{q}_k}{\|\tilde{q}_k\|_2}, \quad \text{where} \quad \tilde{q}_k := a_k - \sum_{l=1}^{k-1}\langle q_l, a_k\rangle q_l.$$

This allows us to write

$$a_k = \sum_{l=1}^{k} r_{lk} q_l, \qquad r_{lk} := \begin{cases} \langle q_l, a_k\rangle & \text{, if } 1 \le l \le k-1, \\ \pm\|\tilde{q}_k\|_2 & \text{, if } l = k. \end{cases}.$$

$\Longrightarrow$ Found orthonormal vectors $q_1, \ldots, q_n \in \mathbb{R}^m$ and numbers $\{r_{ij}\}_{1\le i\le j\le n} \subseteq \mathbb{R}$ s.t. $A = \hat{Q}\hat{R}$ with $\hat{Q} = (q_1|\cdots|q_n)$ and $\hat{R} = (r_{ij})$:

$$\forall\, 1 \le k \le n: \quad q_k = \frac{1}{r_{kk}}\left(a_k - \sum_{l=1}^{k-1} r_{lk} q_l\right),$$

$$\forall\, 1 \le i \le j \le n: \quad r_{ij} = \begin{cases} \langle q_i, a_j\rangle & \text{, if } i \le j-1, \\ \pm\|a_j - \sum_{l=1}^{j-1} r_{lj} q_l\|_2 & \text{, if } i = j. \end{cases}$$

The sign of the values r_{jj}, $1 \le j \le n$, is not determined and we use the convention to choose $r_{jj} > 0$ for all j.

Algorithm: Gram–Schimdt orthogonalization

Let $m, n \in \mathbb{N}$, $m \geq n$, and $A = (a_1 | \cdots | a_n) \in \mathbb{R}^{m \times n}$ with $\mathrm{rk}(A) = n$.
Then, A has the reduced QR factorization $A = \hat{Q}\hat{R}$ with

$$\hat{Q} := (q_1 | \cdots | q_n) \in \mathbb{R}^{m \times n}, \qquad \hat{R} := \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & r_{nn} \end{pmatrix} \in \mathbb{R}^{n \times n}$$

determined as follows:

1) Compute

$$\tilde{q}_1 := a_1 \in \mathbb{R}^m, \quad r_{11} := \|\tilde{q}_1\|_2 > 0, \quad q_1 := \frac{1}{r_{11}}\tilde{q}_1 \in \mathbb{R}^m.$$

If $n = 1$, we stop. If $n \geq 2$, we continue as follows.

2) Compute $r_{12} := \langle q_1, a_2 \rangle \in \mathbb{R}$. Then, compute

$$\tilde{q}_2 := a_2 - r_{12} q_1 \in \mathbb{R}^m, \quad r_{22} := \|\tilde{q}_2\|_2 > 0, \quad q_2 := \frac{1}{r_{22}}\tilde{q}_2 \in \mathbb{R}^m.$$

$\vdots$

j) Compute $r_{ij} := \langle q_i, a_j \rangle \in \mathbb{R}$ for $i \in \{1, \ldots, j-1\}$. Then, compute

$$\tilde{q}_j := a_j - \sum_{l=1}^{j-1} r_{lj} q_l \in \mathbb{R}^m, \quad r_{jj} := \|\tilde{q}_j\|_2 > 0, \quad q_j := \frac{1}{r_{jj}}\tilde{q}_j \in \mathbb{R}^m.$$

$\vdots$

n) Compute $r_{in} := \langle q_i, a_n \rangle \in \mathbb{R}$ for $i \in \{1, \ldots, n-1\}$. Then, compute

$$\tilde{q}_n := a_n - \sum_{l=1}^{n-1} r_{ln} q_l \in \mathbb{R}^m, \quad r_{nn} := \|\tilde{q}_n\|_2 > 0, \quad q_n := \frac{1}{r_{nn}}\tilde{q}_n \in \mathbb{R}^m.$$

Example

Consider $A := (a_1|a_2|a_3) := \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & 2 & 1 \end{pmatrix} \in \mathbb{R}^{4\times 3}$. (Note $\mathrm{rk}(A) = 3$.)

1) $\tilde{q}_1 := a_1 = (1,-1,1,1)^{\mathrm{T}}$. Then, $r_{11} := \|\tilde{q}_1\|_2 = 2$ and we set $q_1 := r_{11}^{-1}\tilde{q}_1 = \frac{1}{2}(1,-1,1,1)^{\mathrm{T}}$.
2) $r_{12} := \langle q_1, a_2\rangle = 1$, $\quad \tilde{q}_2 := a_2 - r_{12}q_1 = \frac{1}{2}(-1,3,1,3)^{\mathrm{T}}$. Then, $r_{22} := \|\tilde{q}_2\|_2 = \sqrt{5}$ and we set $q_2 := r_{22}^{-1}\tilde{q}_2 = \frac{1}{2\sqrt{5}}(-1,3,1,3)^{\mathrm{T}}$.
3) $r_{13} := \langle q_1, a_3\rangle = 0$, $r_{23} := \langle q_2, a_3\rangle = \frac{2}{\sqrt{5}}$, $\tilde{q}_3 := a_3 - r_{13}q_1 - r_{23}q_2 = \frac{2}{5}(3,1,-3,1)^{\mathrm{T}}$. Then, $r_{33} := \|\tilde{q}_3\|_2 = \frac{4}{\sqrt{5}}$ and we set $q_3 := r_{33}^{-1}\tilde{q}_3 = \frac{1}{2\sqrt{5}}(3,1,-3,1)^{\mathrm{T}}$.

$$\Longrightarrow \quad A = \hat{Q}\hat{R} \quad \text{with} \quad \hat{Q} := \begin{pmatrix} \frac{1}{2} & -\frac{1}{2\sqrt{5}} & \frac{3}{2\sqrt{5}} \\ -\frac{1}{2} & \frac{3}{2\sqrt{5}} & \frac{1}{2\sqrt{5}} \\ \frac{1}{2} & \frac{1}{2\sqrt{5}} & -\frac{3}{2\sqrt{5}} \\ \frac{1}{2} & \frac{3}{2\sqrt{5}} & \frac{1}{2\sqrt{5}} \end{pmatrix}, \quad \hat{R} := \begin{pmatrix} 2 & 1 & 0 \\ 0 & \sqrt{5} & \frac{2}{\sqrt{5}} \\ 0 & 0 & \frac{4}{\sqrt{5}} \end{pmatrix}.$$

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

$$\hat{Q} := \begin{pmatrix} \frac{1}{2} & -\frac{1}{2\sqrt{5}} & \frac{3}{2\sqrt{5}} \\ -\frac{1}{2} & \frac{3}{2\sqrt{5}} & \frac{1}{2\sqrt{5}} \\ \frac{1}{2} & \frac{1}{2\sqrt{5}} & -\frac{3}{2\sqrt{5}} \\ \frac{1}{2} & \frac{3}{2\sqrt{5}} & \frac{1}{2\sqrt{5}} \end{pmatrix}, \qquad \hat{R} := \begin{pmatrix} 2 & 1 & 0 \\ 0 & \sqrt{5} & \frac{2}{\sqrt{5}} \\ 0 & 0 & \frac{4}{\sqrt{5}} \end{pmatrix}$$

is a reduced QR factorization of A. How to obtain a full QR factorization?

"Fill up" $\hat{Q}$ with additional orthonormal column and $\hat{R}$ with additional row of zeros: can take, e.g.,

$$Q := \begin{pmatrix} \frac{1}{2} & -\frac{1}{2\sqrt{5}} & \frac{3}{2\sqrt{5}} & \frac{1}{2} \\ -\frac{1}{2} & \frac{3}{2\sqrt{5}} & \frac{1}{2\sqrt{5}} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2\sqrt{5}} & -\frac{3}{2\sqrt{5}} & \frac{1}{2} \\ \frac{1}{2} & \frac{3}{2\sqrt{5}} & \frac{1}{2\sqrt{5}} & -\frac{1}{2} \end{pmatrix}, \qquad R := \begin{pmatrix} 2 & 1 & 0 \\ 0 & \sqrt{5} & \frac{2}{\sqrt{5}} \\ 0 & 0 & \frac{4}{\sqrt{5}} \\ 0 & 0 & 0 \end{pmatrix}$$

to find that $A = QR$ is a (full) QR factorization of A.

From reduced to full QR

From a reduced QR factorization, we can always obtain a full QR factorization:

Let $A \in \mathbb{R}^{m \times n}$, $m \geq n$, and suppose $A = \hat{Q}\hat{R}$ is a reduced QR factn.

- If $m = n$, this is already a full QR factorization.

- If $m > n$, choose arbitrary orthonormal vectors $q_{n+1}, \ldots, q_m \in \mathbb{R}^m$ satisfying $\{q_{n+1}, \ldots, q_m\} \perp \{q_1, \ldots, q_n\}$, and obtain

$$A = (\hat{Q}|q_{n+1}|\cdots|q_m)\left(\frac{\hat{R}}{0_{(m-n)\times n}}\right) =: QR$$

is a (full) QR factorization of A.

Existence

Theorem (Existence result for QR)

Let $m, n \in \mathbb{N}$ with $m \geq n$. Then, every $A \in \mathbb{R}^{m \times n}$ has a QR factorization.

Proof: We know every full-rank matrix $A \in \mathbb{R}^{m \times n}$, $m \geq n$, has a reduced QR factorization (Gram–Schmidt Algorithm) and hence, also a full QR factorization.

It remains to consider the case of rank-deficient matrices: To this end, let $A \in \mathbb{R}^{m \times n}$, $m \geq n$, with $0 \leq \mathrm{rk}(A) < n$.

Then, running the Gram–Schmidt Algorithm, there will be at least one step j, where $\tilde{q}_j = 0$. Whenever this happens, set $r_{jj} = 0$ and take $q_j \in \mathbb{R}^m$, $\|q_j\|_2 = 1$, satisfying $\{q_j\} \perp \{q_1, \ldots, q_{j-1}\}$, and continue the Algorithm.

This yields a reduced QR factorization for A, from which we can then obtain a full QR factorization. □

$\Longrightarrow$ We now have a way to compute reduced and full QR factorizations to arbitrary real $m \times n$ matrices with $m \geq n$.

Exercises can be found on the problem sheets.

Next: Uniqueness?

Is the QR factorization unique?

No. • In 1D: Let $A = (a) \in \mathbb{R}^{1\times 1}$. Then, A has the QR factorizations

$$(a) = \underbrace{(1)}_{Q}\underbrace{(a)}_{R}, \qquad (a) = \underbrace{(-1)}_{Q}\underbrace{(-a)}_{R}.$$

• Let $A \in \mathbb{R}^{m\times n}$, $m \geq n$, and suppose $A = QR$ is a QR factorization of A. Then, $A = (-Q)(-R)$ is also a QR factorization of A.

• Let $A \in \mathbb{R}^{m\times n}$, $m \geq n$, and suppose $A = QR$ is a QR factorization of A (recall $Q \in \mathbb{R}^{m\times m}$ orthogonal and $R \in \mathbb{R}^{m\times n}$ upper-triangular). Write $Q = (q_1|\cdots|q_m)$, $R^{\mathrm{T}} = (b_1|\cdots|b_m)$, and let $s_1, \dots, s_m \in \{-1, 1\}$. Then,

$$A = QR = (q_1|\cdots|q_m)\begin{pmatrix} b_1^{\mathrm{T}} \\ \hline \vdots \\ \hline b_m^{\mathrm{T}} \end{pmatrix} = (s_1 q_1|\cdots|s_m q_m)\begin{pmatrix} s_1 b_1^{\mathrm{T}} \\ \hline \vdots \\ \hline s_m b_m^{\mathrm{T}} \end{pmatrix} =: \tilde{Q}\tilde{R}.$$

$\Longrightarrow$ Given a QR factorization, we can construct new QR factorizations by multiplying the i-th column of Q and the i-th row of R by $s_i \in \{-1, 1\}$.

$\Longrightarrow$ There is hope: we only used signs to construct new factorizations.

Uniqueness result for QR

Theorem (Uniqueness result for QR)

Let $m, n \in \mathbb{N}$ with $m \geq n$. Then, every $A \in \mathbb{R}^{m \times n}$ with $\mathrm{rk}(A) = n$ has a unique reduced QR factn $A = \hat{Q}\hat{R}$ with $\hat{R}$ having positive diagonal entries.

Proof: Let $A \in \mathbb{R}^{m \times n}$, $m \geq n$, be a matrix of full rank, i.e., $\mathrm{rk}(A) = n$. For any reduced QR factorization $A = \hat{Q}\hat{R}$ with $\hat{Q} = (q_1 | \cdots | q_n) \in \mathbb{R}^{m \times n}$ having orthonormal columns and $\hat{R} = (r_{ij}) \in \mathbb{R}^{n \times n}$ upper-triangular, have

$$\begin{cases} a_1 & = r_{11} q_1, \\ a_2 & = r_{12} q_1 + r_{22} q_2, \\ & \vdots \\ a_n & = r_{1n} q_1 + r_{2n} q_2 + \cdots + r_{nn} q_n. \end{cases} \implies \quad q_j = \frac{a_j - \sum_{l=1}^{j-1} r_{lj} q_l}{r_{jj}} \quad \forall\, 1 \leq j \leq n$$

Note $r_{jj} \neq 0 \ \forall j$ $(\mathrm{rk}(A) = n)$. Left-multiply by q_i^{T}, $i < j$: $0 = \frac{\langle q_i, a_j \rangle - r_{ij}}{r_{jj}}$.

$$\forall\, 1 \leq i \leq j \leq n : \quad r_{ij} = \begin{cases} \langle q_i, a_j \rangle & \text{, if } i < j, \\ \pm \| a_j - \sum_{l=1}^{j-1} r_{lj} q_l \|_2 & \text{, if } i = j. \end{cases}$$

$\implies$ Requiring $r_{jj} > 0 \ \forall j$ makes $\hat{Q}, \hat{R}$ uniquely determined. $\square$

Application of QR: solving linear systems

The QR factorization provides a method to solve linear systems.

Given $A \in \mathbb{R}^{m\times n}$, $m \geq n$, and $b \in \mathbb{R}^m$. Problem: find $x \in \mathbb{R}^n$ s.t.

$$Ax = b.$$

If we have a QR factorization $A = QR$, we have

$$Ax = b \quad \Longleftrightarrow \quad QRx = b \quad \Longleftrightarrow \quad Rx = Q^{\mathrm{T}}b.$$

$\Longrightarrow$ compute $\tilde{b} := Q^{\mathrm{T}}b \in \mathbb{R}^m$ and then solve the upper-triangular system

$$Rx = \tilde{b}.$$

(solve by backward substitution, cheap!)

3.3 Projectors

What is a projector?

Definition (Projector/projection matrix)

A square matrix $P \in \mathbb{R}^{n \times n}$ is called a **projector**, or a **projection matrix**, iff

$$P^2 = P$$

(i.e., iff $P \in \mathbb{R}^{n \times n}$ is idempotent).

Note that for $P \in \mathbb{R}^{n \times n}$:

$$P^2 = P \quad \Longleftrightarrow \quad L_P \circ L_P = L_P$$

(recall defn of associated linear map: $L_P : \mathbb{R}^n \to \mathbb{R}^n$, $x \mapsto Px$).

$\Longrightarrow$ So, why are those matrices named projectors?

Why is a square matrix P with $P^2 = P$ called projector?

Two crucial observations: If $P \in \mathbb{R}^{n\times n}$ is a projector (i.e., $P^2 = P$), then

- there holds

$$Py = y \qquad \forall y \in \mathcal{R}(P).$$

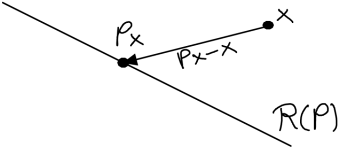
Proof: Let $y \in \mathcal{R}(P)$. Then, $y = Px$ for some $x \in \mathbb{R}^n$. Hence, $Py = P^2x = Px = y$. □

- there holds

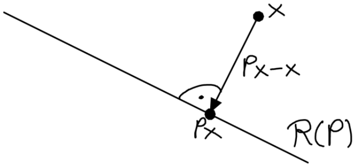
$$Px - x \in \mathcal{N}(P) \qquad \forall x \in \mathbb{R}^n.$$

Proof: $\forall x \in \mathbb{R}^n$: $P(Px - x) = P^2x - Px = Px - Px = 0$. □

We say **the projector P projects onto $\mathcal{R}(P)$ along $\mathcal{N}(P)$**.



"**oblique projector**"



"**orthogonal projector**"

The complementary projector

Let $P \in \mathbb{R}^{n\times n}$ be a projector. Then,

$$(I_n - P)^2 = I_n^2 - 2P + P^2 = I_n - 2P + P = I_n - P,$$

i.e., $I_n - P \in \mathbb{R}^{n\times n}$ is a projector.

Definition (Complementary projector)

Let $P \in \mathbb{R}^{n\times n}$ be a projector. Then, $I_n - P \in \mathbb{R}^{n\times n}$ is called the **complementary projector** to P.

We are going to see that the complementary projector to P is the projector onto $\mathscr{N}(P)$ along $\mathscr{R}(P)$.

Before we prove this, let's introduce the following:

Definition (Complementary subspaces)

Let $S_1, S_2 \subseteq \mathbb{R}^n$ be subspaces of $\mathbb{R}^n$. Then, S_1 and S_2 are called **complementary subspaces** of $\mathbb{R}^n$ iff

$$S_1 + S_2 = \mathbb{R}^n \quad \text{and} \quad S_1 \cap S_2 = \{0\}.$$

Projectors separate $\mathbb{R}^n$ into two complementary subspaces

Theorem (A fundamental result for projectors)

Let $P \in \mathbb{R}^{n\times n}$ be a projector. Then,

(i) $\mathscr{R}(I_n - P) = \mathscr{N}(P)$ *and* $\mathscr{N}(I_n - P) = \mathscr{R}(P)$.

(ii) $\mathscr{R}(P)$ *and* $\mathscr{N}(P)$ *are complementary subspaces of* $\mathbb{R}^n$. *Further, for any* $x \in \mathbb{R}^n$,

$$x = Px + (I_n - P)x \in \mathscr{R}(P) + \mathscr{N}(P)$$

is the unique way of writing $x = x_1 + x_2$ *with* $x_1 \in \mathscr{R}(P), x_2 \in \mathscr{N}(P)$.

Proof of (i): Start by showing $\mathscr{R}(I_n - P) = \mathscr{N}(P)$.
"$\subseteq$" Let $y \in \mathscr{R}(I_n - P)$. Then, $\exists x \in \mathbb{R}^n$: $y = x - Px$. We find $Py = Px - P^2x = Px - Px = 0$, i.e., $y \in \mathscr{N}(P)$.
"$\supseteq$": Let $x \in \mathscr{N}(P)$. Then, $Px = 0$ and hence, $x = x - Px \in \mathscr{R}(I_n - P)$.

Next, show $\mathscr{N}(I_n - P) = \mathscr{R}(P)$. We know that $\tilde{P} := I_n - P \in \mathbb{R}^{n\times n}$ is a projector. Hence, $\mathscr{R}(I_n - \tilde{P}) = \mathscr{N}(\tilde{P})$, i.e., $\mathscr{R}(P) = \mathscr{N}(I_n - P)$. □

Claim (ii): $\mathscr{R}(P)$ and $\mathscr{N}(P)$ are complementary subspaces of $\mathbb{R}^n$. Further, for any $x \in \mathbb{R}^n$,

$$x = Px + (I_n - P)x \in \mathscr{R}(P) + \mathscr{N}(P)$$

is the unique way of writing $x = x_1 + x_2$ with $x_1 \in \mathscr{R}(P), x_2 \in \mathscr{N}(P)$.

Proof of (ii): We only show $\mathscr{R}(P)$ and $\mathscr{N}(P)$ are complementary subspaces of $\mathbb{R}^n$ as the second part is a consequence by a later result.

First, note $\mathscr{R}(P)$ and $\mathscr{N}(P)$ are subspaces of $\mathbb{R}^n$.

Let us show that $\mathscr{R}(P) + \mathscr{N}(P) = \mathbb{R}^n$:

"$\subseteq$" ✓

"$\supseteq$" Let $x \in \mathbb{R}^n$. Then,

$$x = Px + (I_n - P)x \in \mathscr{R}(P) + \mathscr{R}(I_n - P) = \mathscr{R}(P) + \mathscr{N}(P). \quad ✓$$

Next, let us show that $\mathscr{R}(P) \cap \mathscr{N}(P) = \{0\}$:

"$\supseteq$" ✓

"$\supseteq$" Let $x \in \mathscr{R}(P) \cap \mathscr{N}(P)$. Then, $x = P\tilde{x}$ for some $\tilde{x} \in \mathbb{R}^n$, and $Px = 0$. Hence, $0 = Px = P^2\tilde{x} = P\tilde{x} = x$. ✓ □

$\Longrightarrow$ A projector $P \in \mathbb{R}^{n\times n}$ separates $\mathbb{R}^n$ into two complementary subspaces, namely $\mathcal{R}(P)$ and $\mathcal{N}(P)$.

$\Longrightarrow$ What about the converse? Given two complementary subspaces S_1, S_2 of $\mathbb{R}^n$, can we find a projector $P \in \mathbb{R}^{n\times n}$ s.t. $\mathcal{R}(P) = S_1$, $\mathcal{N}(P) = S_2$?

Yes!

Theorem (Projector onto S_1 along S_2)

Let $S_1, S_2 \subseteq \mathbb{R}^n$ be two complementary subspaces of $\mathbb{R}^n$. Then, there exists a unique projector $P \in \mathbb{R}^{n\times n}$ such that $\mathcal{R}(P) = S_1$ and $\mathcal{N}(P) = S_2$. We call this projector **the projector onto S_1 along S_2**.

Claim: Let $S_1, S_2 \subseteq \mathbb{R}^n$ be two complementary subspaces of $\mathbb{R}^n$. Then, there exists a unique projector $P \in \mathbb{R}^{n\times n}$ s.t. $\mathscr{R}(P) = S_1$, $\mathscr{N}(P) = S_2$.

Proof: *Step 1*: We show any $x \in \mathbb{R}^n$ has a unique decomposition

$$x = x_1 + x_2 \quad \text{with} \quad x_1 \in S_1,\ x_2 \in S_2.$$

Existence: ✓, since $S_1 + S_2 = \mathbb{R}^n$.

Uniqueness: Suppose $\exists\, x_1, \tilde{x}_1 \in S_1, x_2, \tilde{x}_2 \in S_2$ s.t.

$$x = x_1 + x_2 = \tilde{x}_1 + \tilde{x}_2.$$

$$\implies \underbrace{x_1 - \tilde{x}_1}_{\in S_1} = \underbrace{\tilde{x}_2 - x_2}_{\in S_2} \in S_1 \cap S_2 = \{0\}.$$

$$\implies x_1 = \tilde{x}_1 \text{ and } x_2 = \tilde{x}_2.$$

Claim: Let $S_1, S_2 \subseteq \mathbb{R}^n$ be two complementary subspaces of $\mathbb{R}^n$. Then, there exists a unique projector $P \in \mathbb{R}^{n\times n}$ s.t. $\mathscr{R}(P) = S_1$, $\mathscr{N}(P) = S_2$.

Proof: *Step 2*: Existence (construction) of P.

Define map

$$L : \mathbb{R}^n \to \mathbb{R}^n, \qquad x = \underbrace{x_1}_{\in S_1} + \underbrace{x_2}_{\in S_2} \mapsto x_1.$$

(well-defined by Step 1.)

- Claim: L is linear.

- Proof: Given $x = x_1 + x_2 \in \mathbb{R}^n$, $y = y_1 + y_2 \in \mathbb{R}^n$ with $x_1, y_1 \in S_1$, $x_2, y_2 \in S_2$, and $\alpha \in \mathbb{R}$, we have

$$L(\alpha x + y) = L(\underbrace{(\alpha x_1 + y_1)}_{\in S_1} + \underbrace{(\alpha x_2 + y_2)}_{\in S_2}) = \alpha x_1 + y_1 = \alpha L(x) + L(y).$$

$\Longrightarrow \exists\ P \in \mathbb{R}^{n\times n}$ s.t. $L = L_P$ (i.e., $L(x) = Px$ for all $x \in \mathbb{R}^n$).

P is a projector: For any $x = x_1 + x_2 \in \mathbb{R}^n$ with $x_1 \in S_1$, $x_2 \in S_2$:

$$P^2 x = L(L(x)) = L(x_1) = L(x_1 + 0) = x_1 = L(x) = Px.$$

$\implies P^2 = P$ ✓

$\mathcal{R}(P) = S_1$: Note $\mathcal{R}(P) = \{L(x) \,|\, x \in \mathbb{R}^n\} \subseteq S_1$ (recall $L(x) = x_1$).
Conversely, for $y \in S_1$: $y = \underbrace{y}_{\in S_1} + \underbrace{0}_{\in S_2}$, thus $y = L(y) = Py \in \mathcal{R}(P)$. ✓

$\mathcal{N}(P) = S_2$: For any $x = x_1 + x_2 \in \mathbb{R}^n$ with $x_1 \in S_1$, $x_2 \in S_2$:
$Px = 0 \iff L(x) = 0 \iff x_1 = 0 \iff x \in S_2$. ✓

$\implies$ We have found a projector $P \in \mathbb{R}^{n \times n}$ with $\mathcal{R}(P) = S_1$, $\mathcal{N}(P) = S_2$.

Remains to show uniqueness.

Claim: Let $S_1, S_2 \subseteq \mathbb{R}^n$ be two complementary subspaces of $\mathbb{R}^n$. Then, there exists a unique projector $P \in \mathbb{R}^{n\times n}$ s.t. $\mathscr{R}(P) = S_1$, $\mathscr{N}(P) = S_2$.

Proof: *Step 3*: Uniqueness of P.

Suppose $\exists$ another projector $\tilde{P} \in \mathbb{R}^{n\times n}$ with $\mathscr{R}(\tilde{P}) = S_1$, $\mathscr{N}(\tilde{P}) = S_2$.

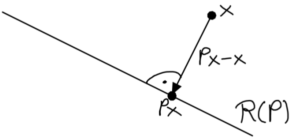
Then, must have $\tilde{P}y = y$ for any $y \in \mathscr{R}(\tilde{P}) = S_1$.

$\Longrightarrow$ For any $x \in \mathbb{R}^n$ with $x = x_1 + x_2$ where $x_1 \in S_1$, $x_2 \in S_2$:

$$\tilde{P}x = \tilde{P}x_1 + \tilde{P}x_2 = x_1 + 0 = x_1 = L(x) = Px.$$

$\Longrightarrow \tilde{P} = P$. □

Orthogonal projectors



Definition (Orthogonal projector)

A projector $P \in \mathbb{R}^{n\times n}$ is called an **orthogonal projector** iff it projects onto S_1 along S_2 for some subspaces S_1, S_2 of $\mathbb{R}^n$ with $S_1 \perp S_2$. A projector which is not an orthogonal projector is called oblique projector.

$\Longrightarrow$ $P \in \mathbb{R}^{n\times n}$ is orthogonal projector iff $P^2 = P$ and $\mathcal{R}(P) \perp \mathcal{N}(P)$.

Theorem (Characterization of orthogonal projectors)

A matrix $P \in \mathbb{R}^{n\times n}$ is an orthogonal projector iff $P^2 = P = P^{\mathrm{T}}$.

WARNING: Orthogonal projectors do not need to be orthogonal matrices. Actually, I_n is the only matrix in $\mathbb{R}^{n\times n}$ that is orthogonal and an orthogonal projector. (Pf: $P^2 = P = P^{\mathrm{T}} = P^{-1} \Longleftrightarrow P = I_n$.)

$P \in \mathbb{R}^{n \times n}$ is orthogonal projector $\iff P^2 = P = P^{\mathrm{T}}$.

Proof: "$\Longleftarrow$": Let $P \in \mathbb{R}^{n \times n}$ with $P^2 = P = P^{\mathrm{T}}$. Then,

- P is a projector ✓
- $\mathcal{R}(P) \perp \mathcal{N}(P)$: Let $y \in \mathcal{R}(P)$, $x \in \mathcal{N}(P)$. Need to show $\langle y, x \rangle = 0$.

$y \in \mathcal{R}(P) \implies \exists v \in \mathbb{R}^n : y = Pv$,

$x \in \mathcal{N}(P) \implies Px = 0$.

Then,

$$\langle y, x \rangle = \langle Pv, x \rangle = \langle v, P^{\mathrm{T}} x \rangle = \langle v, Px \rangle = 0.$$

$\implies P$ is an orthogonal projector.

$P \in \mathbb{R}^{n\times n}$ is orthogonal projector $\iff P^2 = P = P^{\mathrm{T}}$.

<u>Proof:</u> "$\Longrightarrow$": Let $P \in \mathbb{R}^{n\times n}$ orthogonal projector, i.e.,

$$P^2 = P, \qquad \mathcal{R}(P) \perp \mathcal{N}(P).$$

Need to prove $P^{\mathrm{T}} = P$. If $P = 0_{n\times n}$, done. So, suppose $P \neq 0_{n\times n}$.

Write $r := \mathrm{rk}(P) = \dim(\mathcal{R}(P)) \in \{1, \ldots, n\}$. Note $\dim(\mathcal{N}(P)) = n - r$.

- Let $\{q_1, \ldots, q_r\}$ ONB of $\mathcal{R}(P)$. Note $Pq_i = q_i\ \forall 1 \le i \le r$.
- Let $\{q_{r+1}, \ldots, q_n\}$ ONB of $\mathcal{N}(P)$. Note $Pq_i = 0\ \forall r+1 \le i \le n$.

$\Longrightarrow$ As $\mathcal{R}(P) \perp \mathcal{N}(P)$, have $\{q_1, \ldots, q_n\}$ is orthonormal basis (ONB) of $\mathbb{R}^n$.

Set $Q := (q_1|\cdots|q_n) \in \mathbb{R}^{n\times n}$ and note Q is orthogonal. Then,

$$PQ = Q\Sigma, \quad \text{where} \quad \Sigma := \mathrm{diag}_{n\times n}(\underbrace{1, \ldots, 1}_{\text{r times}}, \underbrace{0, \ldots, 0}_{\text{(n-r) times}}).$$

$\Longrightarrow P = Q\Sigma Q^{\mathrm{T}}$. **We found a SVD and an eigval decomposition of P!**

$\Longrightarrow P^{\mathrm{T}} = Q\Sigma^{\mathrm{T}}Q^{\mathrm{T}} = Q\Sigma Q^{\mathrm{T}} = P$. □

Singular values of projectors

Theorem (Singular values of projectors)

Let $P \in \mathbb{R}^{n \times n} \backslash \{0\}$ be a projector with rank $r := \mathrm{rk}(P)$ and singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$. Then,

(i) *$\sigma_i \geq 1$ for all $i \in \{1, \ldots, r\}$.*

(ii) *P is an orthogonal projector $\Longleftrightarrow \sigma_1 = \|P\|_2 = 1$.*

Proof of (ii): (proof of (i) is an exercise)

"$\Longrightarrow$": If P is an orthogonal projector, $\sigma_i = 1 \ \forall \ 1 \leq i \leq r \Longrightarrow \sigma_1 = 1$.

Singular values of projectors

Theorem (Singular values of projectors)

Let $P \in \mathbb{R}^{n\times n}\backslash\{0\}$ be a projector with rank $r := \mathrm{rk}(P)$ and singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$. Then,

(i) *$\sigma_i \geq 1$ for all $i \in \{1, \ldots, r\}$.*

(ii) *P is an orthogonal projector $\iff \sigma_1 = \|P\|_2 = 1$.*

Proof of (ii): "$\Longleftarrow$": Suppose $P \in \mathbb{R}^{n\times n}\backslash\{0\}$ is a projector with $\sigma_1 = 1$.
Let $P = U\Sigma V^{\mathrm{T}} = (u_1|\cdots|u_n)\mathrm{diag}_{n\times n}(\sigma_1, \ldots, \sigma_n)(v_1|\cdots|v_n)^{\mathrm{T}}$ SVD.

By (i), $1 = \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \geq 1 \Longrightarrow \sigma_i = 1 \; \forall\; 1 \leq i \leq r$.

$\Longrightarrow P = \sum_{i=1}^r \sigma_i u_i v_i^{\mathrm{T}} = \sum_{i=1}^r u_i v_i^{\mathrm{T}}, \quad P^{\mathrm{T}} = \sum_{i=1}^r v_i u_i^{\mathrm{T}}$.

Note $\forall 1 \leq j \leq r$: $Pu_j = u_j \Longrightarrow \sum_{i=1}^r \langle v_i, u_j\rangle u_i = u_j \Longrightarrow \langle v_j, u_j\rangle = 1$.

$\Longrightarrow v_j = u_j \; \forall 1 \leq j \leq r$ as $\|v_j - u_j\|_2^2 = \|v_j\|_2^2 + \|u_j\|_2^2 - 2\langle v_j, u_j\rangle = 0$.

$\Longrightarrow P = \sum_{i=1}^r u_i u_i^{\mathrm{T}} = P^{\mathrm{T}}$, i.e., P is an orthogonal projector. $\square$

Projection with orthonormal basis

Let $\{q_1, \dots, q_n\}$ orthonormal basis of $\mathbb{R}^n$, and consider the complementary subspaces $S_1 := \mathrm{span}(q_1, \dots, q_r)$ and $S_2 := \mathrm{span}(q_{r+1}, \dots, q_n)$ of $\mathbb{R}^n$, where $1 \leq r \leq n-1$.

Then, the unique projector $P \in \mathbb{R}^{n\times n}$ onto S_1 along S_2 is given by

$$P = \hat{Q}\hat{Q}^{\mathrm{T}} = \sum_{i=1}^{r} q_i q_i^{\mathrm{T}}, \quad \text{where} \quad \hat{Q} := (q_1|\cdots|q_r) \in \mathbb{R}^{n\times r},$$

and P is actually an orthogonal projector.
Indeed, $\mathscr{R}(P) = \mathscr{R}(\hat{Q}) = S_1$, $\mathscr{N}(P) = \mathscr{N}(\hat{Q}^{\mathrm{T}}) = S_2$, and $P^2 = P = P^{\mathrm{T}}$.

The corresponding linear map

$$L_P : \mathbb{R}^n \to \mathbb{R}^n, \quad x \mapsto \sum_{i=1}^{r} q_i q_i^{\mathrm{T}} x = \sum_{i=1}^{r} \langle x, q_i \rangle q_i$$

projects the vector space $\mathbb{R}^n$ orthogonally onto S_1 along S_2, i.e., it isolates the components of a vector in directions $q_1, \dots, q_r$.

The complementary projector $I_n - P$ is also an orthogonal projector: it is the projector onto $S_2 = \mathrm{span}(q_{r+1}, \ldots, q_n)$ along $S_1 = \mathrm{span}(q_1, \ldots, q_r)$, i.e., it isolates the components of a vector in directions $q_{r+1}, \ldots, q_n$. The corresponding linear map is

$$L_{I_n-P} : \mathbb{R}^n \to \mathbb{R}^n, \quad x \mapsto (I_n - \hat{Q}\hat{Q}^{\mathrm{T}})x = \sum_{i=r+1}^{n} q_i q_i^{\mathrm{T}} x = \sum_{i=r+1}^{n} \langle x, q_i \rangle q_i.$$

Observe that we can decompose any $x \in \mathbb{R}^n$ uniquely into $x = x_1 + x_2$ with $x_1 \in S_1$, $x_2 \in S_2$, where $x_1 = \hat{Q}\hat{Q}^{\mathrm{T}}x$ and $x_2 = (I_n - \hat{Q}\hat{Q}^{\mathrm{T}})x$.

Projection with arbitrary basis

Let S_1 be a subspace of $\mathbb{R}^m$ spanned by $n \le m$ linearly independent vectors $a_1, \ldots, a_n \in \mathbb{R}^m$. We set $A := (a_1|\cdots|a_n) \in \mathbb{R}^{m\times n}$ so that $S_1 = \mathscr{R}(A)$, and construct an orthogonal projector $P \in \mathbb{R}^{m\times m}$ onto S_1.

For $x \in \mathbb{R}^m$ we must have $Px \in S_1$, i.e., $Px = Ay$ for some $y \in \mathbb{R}^n$, and $\{Px - x\} \perp S_1$, i.e.,

$$0_{n\times 1} = \begin{pmatrix} \langle a_1, Px - x\rangle \\ \vdots \\ \langle a_n, Px - x\rangle \end{pmatrix} = A^{\mathrm{T}}(Px - x) = A^{\mathrm{T}}Ay - A^{\mathrm{T}}x.$$

Note that $\mathrm{rk}(A^{\mathrm{T}}A) = \mathrm{rk}(A) = n \implies A^{\mathrm{T}}A \in \mathbb{R}^{n\times n}$ is invertible.

$$\implies y = (A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}x \quad \implies \quad Px = Ay = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}x.$$

The orthogonal projector onto $S_1 = \mathscr{R}(A)$ is given by

$$P = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}} \in \mathbb{R}^{m\times m}.$$

Rk: if $A = \hat{Q}$ has orthonormal columns, this reduces to $P = \hat{Q}\hat{Q}^{\mathrm{T}}$.

Some remarks regarding orthogonal projectors

Let S be a subspace of $\mathbb{R}^n$.

- The orthogonal projector onto S is unique.

  Proof: Suppose $P_1, P_2 \in \mathbb{R}^{n\times n}$ are orthogonal projectors with $\mathcal{R}(P_1) = \mathcal{R}(P_2) = S$. Then, we have

  $$\underbrace{P_1x}_{\in S} - \underbrace{P_2x}_{\in S} = \underbrace{(I_n - P_2)x}_{\in S^\perp} - \underbrace{(I_n - P_1)x}_{\in S^\perp} \in S \cap S^\perp = \{0\} \qquad \forall x \in \mathbb{R}^n,$$

  i.e., $P_1x = P_2x \; \forall x \in \mathbb{R}^n$ and thus, $P_1 = P_2$. Here, we have used that

  $$\mathcal{R}(I_n - P_i) = \mathcal{N}(P_i) \perp \mathcal{R}(P_i) = S \qquad \forall i \in \{1, 2\}.$$

  (Recall definition of orthogonal complement: $S^\perp := \{\, x \in \mathbb{R}^n \,|\, \langle x, s\rangle = 0 \; \forall s \in S\}$.)

- The orthogonal projector onto S is the projector onto S along $S^\perp$.

  Proof: Let $P \in \mathbb{R}^{n\times n}$ be the orthogonal projector onto $S = \mathcal{R}(P)$. Then, using $\mathcal{R}(A)^\perp = \mathcal{N}(A^{\mathrm{T}}) \; \forall A \in \mathbb{R}^{m\times n}$ (exercise) and $P^{\mathrm{T}} = P$, we find $\mathcal{N}(P) = \mathcal{R}(P^{\mathrm{T}})^\perp = \mathcal{R}(P)^\perp = S^\perp$. □

3.4 QR via Gram–Schmidt orthogonalization

Let $A = (a_1 | \cdots | a_n) \in \mathbb{R}^{m \times n}$, $m \geq n$, and assume $\mathrm{rk}(A) = n$.

1) Compute

$$\tilde{q}_1 := a_1 \in \mathbb{R}^m, \quad r_{11} := \|\tilde{q}_1\|_2 > 0, \quad q_1 := \frac{1}{r_{11}} \tilde{q}_1 \in \mathbb{R}^m.$$

If $n = 1$, we stop. If $n \geq 2$, we continue as follows.

2) Compute $r_{12} := \langle q_1, a_2 \rangle \in \mathbb{R}$. Then, compute

$$\tilde{q}_2 := a_2 - r_{12} q_1 \in \mathbb{R}^m, \quad r_{22} := \|\tilde{q}_2\|_2 > 0, \quad q_2 := \frac{1}{r_{22}} \tilde{q}_2 \in \mathbb{R}^m.$$

⋮

j) Compute $r_{ij} := \langle q_i, a_j \rangle \in \mathbb{R}$ for $i \in \{1, \ldots, j-1\}$. Then, compute

$$\tilde{q}_j := a_j - \sum_{l=1}^{j-1} r_{lj} q_l \in \mathbb{R}^m, \quad r_{jj} := \|\tilde{q}_j\|_2 > 0, \quad q_j := \frac{1}{r_{jj}} \tilde{q}_j \in \mathbb{R}^m.$$

⋮

n) Compute $r_{in} := \langle q_i, a_n \rangle \in \mathbb{R}$ for $i \in \{1, \ldots, n-1\}$. Then, compute

$$\tilde{q}_n := a_n - \sum_{l=1}^{n-1} r_{ln} q_l \in \mathbb{R}^m, \quad r_{nn} := \|\tilde{q}_n\|_2 > 0, \quad q_n := \frac{1}{r_{nn}} \tilde{q}_n \in \mathbb{R}^m.$$

Gram–Schmidt and projectors

Let $A = (a_1|\cdots|a_n) \in \mathbb{R}^{m\times n}$, $m \geq n$, and assume $\mathrm{rk}(A) = n$. Let $q_1, \ldots, q_n \in \mathbb{R}^m$ be the orthonormal vectors obtained through Gram–Schmidt and define

$P_1 := I_m$

$P_i := I_m - \hat{Q}_{i-1}\hat{Q}_{i-1}^{\mathrm{T}}$, where $\hat{Q}_{i-1} := (q_1|\cdots|q_{i-1}) \in \mathbb{R}^{m\times(i-1)}$, $2 \leq i \leq n$.

Note that $P_i \in \mathbb{R}^{m\times m}$ projects the vector space $\mathbb{R}^m$ onto the space orthogonal to $\mathrm{span}(q_1, \ldots, q_{i-1})$. Then,

$$q_1 = \frac{P_1 a_1}{\|P_1 a_1\|_2}, \quad q_2 = \frac{P_2 a_2}{\|P_2 a_2\|_2}, \quad \cdots \quad , \quad q_n = \frac{P_n a_n}{\|P_n a_n\|_2},$$

i.e., q_i is precisely the normalized orthogonal projection of a_i onto the space orthogonal to $\mathrm{span}(q_1, \ldots, q_{i-1})$.

Classical Gram–Schmidt iteration

Let $A = (a_1|\cdots|a_n) \in \mathbb{R}^{m\times n}$, $m \geq n$, and $\mathrm{rk}(A) = n$.

for $j = 1, \ldots, n$ **do**
  $\tilde{q}_j = a_j$
  for $i = 1, \ldots, j-1$ **do**
    $r_{ij} = \langle q_i, a_j \rangle$
    $\tilde{q}_j = \tilde{q}_j - r_{ij} q_i$
  end for
  $r_{jj} = \|\tilde{q}_j\|_2$
  $q_j = \frac{1}{r_{jj}} \tilde{q}_j$
end for

Drawback: numerically unstable. However, a simple modification leads to improved stability.

Key observation: projector $P_i = I_m - \hat{Q}_{i-1}\hat{Q}_{i-1}^{\mathrm{T}} \in \mathbb{R}^{m\times m}$ of rank $m-(i-1)$ can be decomposed as product of $i-1$ rank $m-1$ projectors:

$$P_i = (I_m - q_{i-1}q_{i-1}^{\mathrm{T}})(I_m - q_{i-2}q_{i-2}^{\mathrm{T}})\cdots(I_m - q_1 q_1^{\mathrm{T}}), \qquad 2 \leq i \leq n.$$

Modified Gram–Schmidt

Let $A = (a_1|\cdots|a_n) \in \mathbb{R}^{m\times n}$, $m \geq n$, and $\operatorname{rk}(A) = n$. The **modified Gram–Schmidt iteration** does the following:

for $i = 1, \ldots, n$ **do**
  $\tilde{q}_i = a_i$
end for
for $i = 1, \ldots, n$ **do**
  $r_{ii} = \|\tilde{q}_i\|_2$
  $q_i = \frac{1}{r_{ii}}\tilde{q}_i$
  for $j = i+1, \ldots, n$ **do**
    $r_{ij} = \langle q_i, \tilde{q}_j\rangle$
    $\tilde{q}_j = \tilde{q}_j - r_{ij}q_i$
  end for
end for

Theorem

This algorithm requires $\sim 2mn^2$ *flops, i.e.,* $\lim_{m,n\to\infty} \frac{\#\textit{flops}}{2mn^2} = 1$.

Gram–Schmidt = triangular orthogonalization

Schematically, (modified) Gram–Schmidt does the following:

1. $$AR_1 = (a_1|\cdots|a_n)\begin{pmatrix} \frac{1}{r_{11}} & -\frac{r_{12}}{r_{11}} & -\frac{r_{13}}{r_{11}} & \cdots & -\frac{r_{1n}}{r_{11}} \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix} = (q_1|*|\cdots|*),$$
2. $$AR_1R_2 = (q_1|*|\cdots|*)\begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & \frac{1}{r_{22}} & -\frac{r_{23}}{r_{22}} & \cdots & -\frac{r_{2n}}{r_{22}} \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix} = (q_1|q_2|*|\cdots|*),$$

$\vdots$

n. $AR_1R_2\cdots R_n = (q_1|q_2|\cdots|q_n) = \hat{Q}$, i.e., $A = \hat{Q}\hat{R}$ with $\hat{R} = (R_1\cdots R_n)^{-1}$.

Gram–Schmidt is a **triangular orthogonalization** method.

3.5 QR via Householder triangularization

Two different ideologies

As before, consider "tall" matrices $A \in \mathbb{R}^{m \times n}$, $m \geq n$.

Gram–Schmidt: **triangular orthogonalization**, i.e., construct $R_1, \ldots, R_n \in \mathbb{R}^{n \times n}$ upper-triangular s.t.

$$AR_1 R_2 \cdots R_n = \hat{Q} \in \mathbb{R}^{m \times n}$$

is matrix with orthonormal columns.

$\Longrightarrow$ yields reduced QR factorization $A = \hat{Q}\hat{R}$ with $\hat{R} := (R_1 R_2 \cdots R_n)^{-1}$.

Householder: **orthogonal triangularization**, i.e., construct $Q_1, \ldots, Q_n \in \mathbb{R}^{m \times m}$ orthogonal s.t.

$$Q_n \cdots Q_2 Q_1 A = R \in \mathbb{R}^{m \times n}$$

is upper-triangular.

$\Longrightarrow$ yields (full) QR factorization $A = QR$ with $Q := Q_1^{\mathrm{T}} Q_2^{\mathrm{T}} \cdots Q_n^{\mathrm{T}}$.

So, how do we find such orthogonal matrices Q_i?

The idea

We construct orthogonal matrices $Q_1, \dots, Q_n$ in a way so that $A \in \mathbb{R}^{m\times n}$, $m \geq n$ is transformed as follows: (illustration for $m = 4$, $n = 3$)

$$A = \begin{pmatrix} * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \end{pmatrix} \implies Q_1 A = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ 0 & * & * \\ 0 & * & * \\ 0 & * & * \end{pmatrix}$$

$$\implies Q_2 Q_1 A = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & * \\ 0 & 0 & * \end{pmatrix} \implies Q_3 Q_2 Q_1 A = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \\ 0 & 0 & 0 \end{pmatrix}.$$

So, left-multiplication by Q_i should leave the first $(i-1)$ rows and columns unchanged and introduce zeros below the i-th main diagonal entry, thus leading to an upper-triangular matrix $R = Q_n \cdots Q_2 Q_1 A$ after n such steps.

$$A = \begin{pmatrix} * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \end{pmatrix} \implies Q_1 A = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ 0 & * & * \\ 0 & * & * \\ 0 & * & * \end{pmatrix}$$

$$\implies Q_2 Q_1 A = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & * \\ 0 & 0 & * \end{pmatrix} \implies Q_3 Q_2 Q_1 A = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \\ 0 & 0 & 0 \end{pmatrix}.$$

We choose Q_i, $i \in \{1, \dots, n\}$, to be an orthogonal matrix of the form

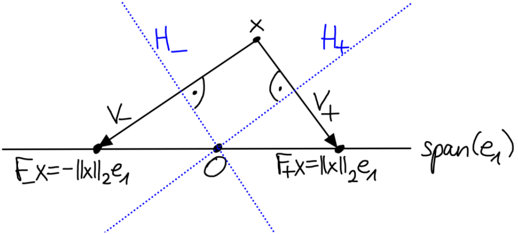
$$Q_i = \left(\begin{array}{c|c} I_{i-1} & 0_{(i-1)\times(m-i+1)} \\ \hline 0_{(m-i+1)\times(i-1)} & F \end{array} \right) \in \mathbb{R}^{m\times m},$$

with $F \in \{F_-, F_+\} \in \mathbb{R}^{(m-i+1)\times(m-i+1)}$ s.t.

$$x = \begin{pmatrix} * \\ * \\ \vdots \\ * \end{pmatrix} \in \mathbb{R}^{m-i+1} \implies F_{\pm} x = \begin{pmatrix} \pm\|x\|_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \pm\|x\|_2 e_1.$$

Geometric illustration of $F_{\pm}$

Householder reflectors



Note $I_{m-i+1} - \frac{vv^{\mathrm{T}}}{\|v\|_2^2}$ is the orthogonal projector onto the hyperplane orthogonal to $v \in \mathbb{R}^{m-i+1}$. Therefore,

$$F = I_{m-i+1} - 2\frac{vv^{\mathrm{T}}}{\|v\|_2^2}$$

is as required. We call F a **Householder reflector**.
For numerical stability, choose reflector which moves x the larger distance:

$$v = \mathrm{sign}(\langle x, e_1\rangle)\|x\|_2 e_1 + x,$$

where $\mathrm{sign}(\alpha) = 1$ for $\alpha \geq 0$ and $\mathrm{sign}(\alpha) = -1$ otherwise.
(Rk: If $\langle x, e_1\rangle \geq 0$, then $v = -v_-$. If $\langle x, e_1\rangle < 0$, then $v = -v_+$.)

Example: QR via Householder triangularization

Task: Compute a QR factorization of $A := \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & 2 & 1 \end{pmatrix}$.

Step 1: Set $x_1 := \begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \end{pmatrix}$ and $v_1 := \operatorname{sign}(\langle x_1, e_1 \rangle)\|x_1\|_2 e_1 + x_1 = \begin{pmatrix} 3 \\ -1 \\ 1 \\ 1 \end{pmatrix}$.

Take $Q_1 := I_4 - 2\frac{v_1 v_1^{\mathrm{T}}}{\|v_1\|_2^2} = \frac{1}{6}\begin{pmatrix} -3 & 3 & -3 & -3 \\ 3 & 5 & 1 & 1 \\ -3 & 1 & 5 & -1 \\ -3 & 1 & -1 & 5 \end{pmatrix}$. Then,

$$Q_1 A = \begin{pmatrix} -2 & -1 & 0 \\ 0 & 4/3 & 4/3 \\ 0 & 2/3 & -4/3 \\ 0 & 5/3 & 2/3 \end{pmatrix}.$$

$$Q_1 A = \begin{pmatrix} -2 & -1 & 0 \\ 0 & 4/3 & 4/3 \\ 0 & 2/3 & -4/3 \\ 0 & 5/3 & 2/3 \end{pmatrix}.$$

Step 2: $x_2 := \begin{pmatrix} 4/3 \\ 2/3 \\ 5/3 \end{pmatrix}$, $v_2 := \operatorname{sign}(\langle x_2, e_1 \rangle) \|x_2\|_2 e_1 + x_2 = \begin{pmatrix} \sqrt{5} + \frac{4}{3} \\ 2/3 \\ 5/3 \end{pmatrix}$.

Take

$$Q_2 := \left(\begin{array}{c|c} 1 & 0_{1\times 3} \\ \hline 0_{3\times 1} & I_3 - 2\frac{v_2 v_2^{\mathrm{T}}}{\|v_2\|_2^2} \end{array}\right) = \frac{\sqrt{5}}{435} \begin{pmatrix} \frac{435}{\sqrt{5}} & 0 & 0 & 0 \\ 0 & -116 & -58 & -145 \\ 0 & -58 & 75\sqrt{5}+16 & -(30\sqrt{5}-40) \\ 0 & -145 & -(30\sqrt{5}-40) & 12\sqrt{5}+100 \end{pmatrix}.$$

Then,

$$Q_2 Q_1 A = \begin{pmatrix} -2 & -1 & 0 \\ 0 & -\sqrt{5} & -\frac{2}{\sqrt{5}} \\ 0 & 0 & -\frac{24\sqrt{5}+200}{145} \\ 0 & 0 & -\frac{12\sqrt{5}-16}{29} \end{pmatrix}.$$

$$Q_2Q_1A = \begin{pmatrix} -2 & -1 & 0 \\ 0 & -\sqrt{5} & -\frac{2}{\sqrt{5}} \\ 0 & 0 & -\frac{24\sqrt{5}+200}{145} \\ 0 & 0 & -\frac{12\sqrt{5}-16}{29} \end{pmatrix}.$$

Step 3: $x_3 := \begin{pmatrix} -\frac{24\sqrt{5}+200}{145} \\ -\frac{12\sqrt{5}-16}{29} \end{pmatrix}$, $v_3 := \mathrm{sign}(\langle x_3, e_1\rangle)\|x_3\|_2 e_1 + x_3 = -\frac{4}{29}\begin{pmatrix} 7\sqrt{5}+10 \\ 3\sqrt{5}-4 \end{pmatrix}$.

Take

$$Q_3 := \left(\begin{array}{c|c} I_2 & 0_{2\times 2} \\ \hline 0_{2\times 2} & I_2 - 2\frac{v_3 v_3^{\mathrm{T}}}{\|v_3\|_2^2} \end{array}\right) = \frac{1}{29}\begin{pmatrix} 29 & 0 & 0 & 0 \\ 0 & 29 & 0 & 0 \\ 0 & 0 & -10\sqrt{5}-6 & 4\sqrt{5}-15 \\ 0 & 0 & 4\sqrt{5}-15 & 10\sqrt{5}+6 \end{pmatrix}.$$

Then,

$$Q_3Q_2Q_1A = \begin{pmatrix} -2 & -1 & 0 \\ 0 & -\sqrt{5} & -\frac{2}{\sqrt{5}} \\ 0 & 0 & \frac{4}{\sqrt{5}} \\ 0 & 0 & 0 \end{pmatrix} =: R.$$

$$Q_3Q_2Q_1A = \begin{pmatrix} -2 & -1 & 0 \\ 0 & -\sqrt{5} & -\frac{2}{\sqrt{5}} \\ 0 & 0 & \frac{4}{\sqrt{5}} \\ 0 & 0 & 0 \end{pmatrix} =: R.$$

Noting that Q_1, Q_2, Q_3 are symmetric orthogonal matrices, we find that $A = QR$ with

$$Q := Q_1Q_2Q_3 = \begin{pmatrix} -\frac{1}{2} & \frac{1}{2\sqrt{5}} & \frac{3}{2\sqrt{5}} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{3}{2\sqrt{5}} & \frac{1}{2\sqrt{5}} & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2\sqrt{5}} & -\frac{3}{2\sqrt{5}} & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{3}{2\sqrt{5}} & \frac{1}{2\sqrt{5}} & \frac{1}{2} \end{pmatrix}, \quad R := \begin{pmatrix} -2 & -1 & 0 \\ 0 & -\sqrt{5} & -\frac{2}{\sqrt{5}} \\ 0 & 0 & \frac{4}{\sqrt{5}} \\ 0 & 0 & 0 \end{pmatrix}$$

is a QR factorization of A.

Algorithm

For a matrix $A \in \mathbb{R}^{m \times n}$, $m \geq n$, the Householder triangularization produces the factor R of a QR factorization $A = QR$ and goes as follows:

for $i = 1, \ldots, n$ **do**
  $x = A_{i:m,i}$
  $v_i = \text{sign}(x_1)\|x\|_2 e_1 + x$ (x_1 denotes the first entry of x)
  $v_i = \frac{1}{\|v_i\|_2} v_i$
  $A_{i:m,i:n} = A_{i:m,i:n} - 2v_i(v_i^{\mathrm{T}} A_{i:m,i:n})$
end for

This algorithm stores the result R in place of A. The reflection vectors $v_1, \ldots, v_n$ are stored for applying and forming Q.

Theorem

The above algorithm requires $\sim 2mn^2 - \frac{2}{3}n^3$ *flops.*

What about Q?

For practical applications, there is often no need to construct Q explicitly. However, e.g. to solve linear systems $Ax = b$ using QR, we need to be able to compute matrix-vector products $Q^{\mathrm{T}}b$.

Noting $Q^{\mathrm{T}} = Q_n \cdots Q_2 Q_1$ (recall $Q = Q_1 Q_2 \cdots Q_n$ and that the Q_i are symmetric and orthogonal), a product $Q^{\mathrm{T}}b$ with a given $b \in \mathbb{R}^m$ can be calculated via:

for $i = 1, \ldots, n$ **do**
$b_{i:m} = b_{i:m} - 2v_i(v_i^{\mathrm{T}} b_{i:m})$
end for,

leaving the result $Q^{\mathrm{T}}b$ in place of b.

If it is required to explicitly form $Q = Q_1 Q_2 \cdots Q_n$, compute $Qe_1, \ldots Qe_m$. A product Qx with a given $x \in \mathbb{R}^m$ can be calculated via:

for $i = n, n-1, \ldots, 1$ **do**
$x_{i:m} = x_{i:m} - 2v_i(v_i^{\mathrm{T}} x_{i:m})$
end for,

leaving the result Qx in place of x.

3.6 QR via Givens rotations

Givens rotations: the idea

Givens rotations is a good alternative method for sparse matrices.

Key observation: Recall any orthogonal matrix $Q \in \mathbb{R}^{2\times 2}$ with $\det(Q) = 1$ is of the form

$$Q(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}, \quad \theta \in [0, 2\pi),$$

and that $L_{Q(\theta)}$ rotates the plane $\mathbb{R}^2$ anticlockwise by the angle θ.

Given $x \in \mathbb{R}^2$, we can find a θ s.t.

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \Longrightarrow \quad Q(\theta)x = \begin{pmatrix} \|x\|_2 \\ 0 \end{pmatrix}.$$

Indeed, take $\theta \in [0, 2\pi)$ s.t.

$$\cos(\theta) = \frac{x_1}{\|x\|_2}, \qquad \sin(\theta) = -\frac{x_2}{\|x\|_2}.$$

Then,

$$Q(\theta)x = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \sqrt{x_1^2 + x_2^2} \\ 0 \end{pmatrix} = \begin{pmatrix} \|x\|_2 \\ 0 \end{pmatrix}.$$

Illustration of the method at an explicit example

Consider $A := \begin{pmatrix} -2 & -1 & 1 \\ 3 & 2 & -1 \\ 4 & 1 & 4 \end{pmatrix}$. Givens rotations in 3D:

$$G_1(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix}, \; L_{G_1(\theta)} : \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ \tilde{x}_2 \\ \tilde{x}_3 \end{pmatrix} \text{ where } \begin{pmatrix} \tilde{x}_2 \\ \tilde{x}_3 \end{pmatrix} = Q(\theta) \begin{pmatrix} x_2 \\ x_3 \end{pmatrix},$$

$$G_2(\theta) = \begin{pmatrix} \cos(\theta) & 0 & -\sin(\theta) \\ 0 & 1 & 0 \\ \sin(\theta) & 0 & \cos(\theta) \end{pmatrix}, \; L_{G_2(\theta)} : \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} \tilde{x}_1 \\ x_2 \\ \tilde{x}_3 \end{pmatrix} \text{ where } \begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_3 \end{pmatrix} = Q(\theta) \begin{pmatrix} x_1 \\ x_3 \end{pmatrix},$$

$$G_3(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}, \; L_{G_3(\theta)} : \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ x_3 \end{pmatrix} \text{ where } \begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{pmatrix} = Q(\theta) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Note that the matrices $G_i(\theta)$, $i \in \{1, 2, 3\}$, are orthogonal.

$$A = \begin{pmatrix} -2 & -1 & 1 \\ 3 & 2 & -1 \\ 4 & 1 & 4 \end{pmatrix}.$$

Step 1: Let us eliminate the entry $a_{31} = 4$ by using the entry $a_{21} = 3$, thus leaving the first row of A unchanged.
$\Longrightarrow$ use the Givens rotation $G_1(\theta)$ with θ such that

$$Q(\theta) \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} * \\ 0 \end{pmatrix}.$$

Take $\theta \in [0, 2\pi)$ such that $\cos(\theta) = \frac{3}{5}$ and $\sin(\theta) = -\frac{4}{5}$ (recall $\cos(\theta) = \frac{x_1}{\|x\|_2}$, $\sin(\theta) = -\frac{x_2}{\|x\|_2}$). Then,

$$G_1 := \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{3}{5} & \frac{4}{5} \\ 0 & -\frac{4}{5} & \frac{3}{5} \end{pmatrix}, \qquad G_1 A = \begin{pmatrix} -2 & -1 & 1 \\ 5 & 2 & \frac{13}{5} \\ 0 & -1 & \frac{16}{5} \end{pmatrix}.$$

$$G_1A = \begin{pmatrix} -2 & -1 & 1 \\ 5 & 2 & \frac{13}{5} \\ 0 & -1 & \frac{16}{5} \end{pmatrix}.$$

Step 2: Let us eliminate the (2,1)-entry using the (1,1)-entry, thus leaving the third row of A unchanged.

$\Longrightarrow$ use the Givens rotation $G_3(\theta)$ with θ such that

$$Q(\theta)\begin{pmatrix} -2 \\ 5 \end{pmatrix} = \begin{pmatrix} * \\ 0 \end{pmatrix}.$$

Take $\theta \in [0, 2\pi)$ such that $\cos(\theta) = \frac{-2}{\sqrt{29}}$ and $\sin(\theta) = -\frac{5}{\sqrt{29}}$. Then,

$$G_3 := \begin{pmatrix} -\frac{2}{\sqrt{29}} & \frac{5}{\sqrt{29}} & 0 \\ -\frac{5}{\sqrt{29}} & -\frac{2}{\sqrt{29}} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad G_3G_1A = \begin{pmatrix} \sqrt{29} & \frac{12}{\sqrt{29}} & \frac{11}{\sqrt{29}} \\ 0 & \frac{1}{\sqrt{29}} & -\frac{51}{5\sqrt{29}} \\ 0 & -1 & \frac{16}{5} \end{pmatrix}.$$

$$G_3G_1A = \begin{pmatrix} \sqrt{29} & \frac{12}{\sqrt{29}} & \frac{11}{\sqrt{29}} \\ 0 & \frac{1}{\sqrt{29}} & -\frac{51}{5\sqrt{29}} \\ 0 & -1 & \frac{16}{5} \end{pmatrix}.$$

Step 3: Let us eliminate the (3,2)-entry using the (2,2)-entry, thus leaving the first row of A unchanged.

$\Longrightarrow$ use the Givens rotation $G_1(\theta)$ with θ such that

$$Q(\theta)\begin{pmatrix} \frac{1}{\sqrt{29}} \\ -1 \end{pmatrix} = \begin{pmatrix} * \\ 0 \end{pmatrix}.$$

Take $\theta \in [0, 2\pi)$ s.t. $\cos(\theta) = \frac{1/\sqrt{29}}{\sqrt{30}/\sqrt{29}} = \frac{1}{\sqrt{30}}$, $\sin(\theta) = -\frac{-1}{\sqrt{30}/\sqrt{29}} = \sqrt{\frac{29}{30}}$.
Then,

$$\tilde{G}_1 := \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{30}} & -\sqrt{\frac{29}{30}} \\ 0 & \sqrt{\frac{29}{30}} & \frac{1}{\sqrt{30}} \end{pmatrix}, \quad \tilde{G}_1G_3G_1A = \begin{pmatrix} \sqrt{29} & \frac{12}{\sqrt{29}} & \frac{11}{\sqrt{29}} \\ 0 & \sqrt{\frac{30}{29}} & -\frac{103}{\sqrt{870}} \\ 0 & 0 & -\frac{7}{\sqrt{30}} \end{pmatrix} =: R.$$

$\Longrightarrow$ We have

$$\tilde{G}_1 G_3 G_1 A = \begin{pmatrix} \sqrt{29} & \frac{12}{\sqrt{29}} & \frac{11}{\sqrt{29}} \\ 0 & \sqrt{\frac{30}{29}} & -\frac{103}{\sqrt{870}} \\ 0 & 0 & -\frac{7}{\sqrt{30}} \end{pmatrix} =: R.$$

Noting that $G_1, G_3, \tilde{G}_1 \in \mathbb{R}^{3\times 3}$ are orthogonal, we have obtained the following QR factorization: $A = QR$ with R as above and

$$Q := G_1^{\mathrm{T}} G_3^{\mathrm{T}} \tilde{G}_1^{\mathrm{T}} = \begin{pmatrix} -\frac{2}{\sqrt{29}} & -\frac{\sqrt{5}}{\sqrt{174}} & -\frac{\sqrt{5}}{\sqrt{6}} \\ \frac{3}{\sqrt{29}} & \frac{11\sqrt{2}}{\sqrt{435}} & -\frac{\sqrt{2}}{\sqrt{15}} \\ \frac{4}{\sqrt{29}} & -\frac{19}{\sqrt{870}} & -\frac{1}{\sqrt{30}} \end{pmatrix}.$$

End of “Chapter 3: QR Factorization”.