1 Projectors & Reflectors (3.2)

In this section, we'll talk about projectors and reflectors, something that's important for QR decomposition.

1.1 Projectors

Definition 1.1: Projector

A **projector** is a matrix P with
$$P^2 = P.$$

Definition 1.2: Orthoprojector

If P is a projector and also symmetric (i.e., $P = P^T$), then P is called an **orthoprojector**.

(Example.) Suppose $\mathbf{u} \in \mathbb{R}^n$ is a unit vector (i.e., $||\mathbf{u}||_2 = 1$). Then, $P = \mathbf{u} \cdot \mathbf{u}^T$ is an orthoprojector. That is,
$$P = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix} = \begin{bmatrix} u_1^2 & u_1u_2 & \dots & u_1u_n \\ u_2u_1 & u_2^2 & \dots & u_2u_n \\ \vdots & \vdots & \ddots & \vdots \\ u_nu_1 & u_nu_1 & \dots & u_n^2 \end{bmatrix}.$$
To see why P here is an orthoprojector, we'll show that it satisfies some properties.

1. Definition of a projector.
$$P^2 = P \cdot P = (\mathbf{u} \cdot \mathbf{u}^T)(\mathbf{u} \cdot \mathbf{u}^T) = \mathbf{u}(\underbrace{\mathbf{u}^T\mathbf{u}}_{1})\mathbf{u}^T = \mathbf{u}\mathbf{u}^T = P.$$

2. Definition of an orthoprojector.
$$P^T = (\mathbf{u}\mathbf{u}^T)^T = (\mathbf{u}^T)^T\mathbf{u}^T = \mathbf{u}\mathbf{u}^T = P.$$

There are some additional properties to know for this case.

- $P\mathbf{u} = \mathbf{u}$:
$$P\mathbf{u} = (\mathbf{u}\mathbf{u}^T)\mathbf{u} = \mathbf{u}(\underbrace{\mathbf{u}^T\mathbf{u}}_{1}) = \mathbf{u}.$$
- If $\mathbf{v} \perp \mathbf{u}$ (i.e., $\langle \mathbf{v}, \mathbf{u} \rangle = 0$), then $P\mathbf{v} = \mathbf{0}$.
$$P\mathbf{v} = (\mathbf{u}\mathbf{u}^T)\mathbf{v} = \mathbf{u}(\underbrace{\mathbf{u}^T\mathbf{v}}_{0}) = \mathbf{0}.$$

Remarks:

- Note that if $\mathbf{u} \in \mathbb{R}^{n\times 1}$, then $\mathbf{u}^T \in \mathbb{R}^{1\times n}$ and so P will be an $n \times n$ matrix.
- Note that $\mathbf{u}\mathbf{u}^T \neq \mathbf{u}^T\mathbf{u}$. In particular, $\mathbf{u}\mathbf{u}^T$ is an $n \times n$ matrix while $\mathbf{u}^T\mathbf{u} = \langle \mathbf{u}, \mathbf{u} \rangle = ||\mathbf{u}||_2^2$.

1.2 Reflectors

Reflectors are built by *projectors*.

Definition 1.3: Reflector

For a unit vector $\mathbf{u} \in \mathbb{R}^n$ (i.e., $||\mathbf{u}||_2 = 1$), $Q = I - 2\mathbf{u}\mathbf{u}^T$ is called a (householder) **reflector**.

Remarks:

- We can rewrite the above with $Q = I - 2P$, where $P = \mathbf{u}\mathbf{u}^T$ is a projector.
- If $\mathbf{u}$ doesn't have unit norm, we can normalize it,

$$\frac{\mathbf{u}}{||\mathbf{u}||_2},$$

  so that $\left|\left|\frac{\mathbf{u}}{||\mathbf{u}||_2}\right|\right|_2 = \frac{1}{||\mathbf{u}||_2}||\mathbf{u}||_2 = 1$ (note that $||\mathbf{u}||_2$ is a scalar.) In this sense, we can write

$$Q = I - 2\frac{\mathbf{u}}{||\mathbf{u}||_2}\frac{\mathbf{u}^T}{||\mathbf{u}||_2} = I - 2\frac{\mathbf{u}\mathbf{u}^T}{||\mathbf{u}||_2^2}.$$

There are some properties of $Q = I - 2\mathbf{u}\mathbf{u}^T$ (where $\mathbf{u}$ is a unit vector) to know.

1. $Q\mathbf{u} = -\mathbf{u}$.
$$Q\mathbf{u} = (I - 2\mathbf{u}\mathbf{u}^T)\mathbf{u} = \mathbf{u} - 2\mathbf{u}\mathbf{u}^T\mathbf{u} = \mathbf{u} - 2\mathbf{u} = -\mathbf{u}.$$
2. $Q\mathbf{v} = \mathbf{v}$ such that $\mathbf{v} \perp \mathbf{u}$.
$$Q\mathbf{v} = (I - 2\mathbf{u}\mathbf{u}^T)\mathbf{v} = \mathbf{v} - 2\mathbf{u}\underbrace{\mathbf{u}^T\mathbf{v}}_{\mathbf{0}} = \mathbf{v}.$$
3. $Q^T = Q$.
$$Q^T = (I - 2\mathbf{u}\mathbf{u}^T)^T = (I - 2P)^T = I - 2P^T = I - 2P = Q.$$
   Here, note that $I^T = I$. Additionally, note that $P^T = P$.
4. $\underbrace{Q^T = Q^{-1}}_{\text{Orthogonal}}$ and $Q = Q^{-1}$ and $Q^TQ = Q^2 = I$.
$$Q^2 = QQ = (I - 2P)(I - 2P) = I - 2P - 2P + 4P^2 = I - 4P - 4P^2 = I - 4P + 4P = I.$$

Lemma 1.1

For any $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^n$ such that

$$\mathbf{y} = \begin{bmatrix}||\mathbf{x}||_2 & 0 & 0 & \dots & 0\end{bmatrix}^T,$$

define $\mathbf{v} = \mathbf{x} - \mathbf{y}$ and $\mathbf{u} = \frac{\mathbf{v}}{||\mathbf{v}||_2}$. Then,

$$Q = I - 2\mathbf{u}\mathbf{u}^T$$

is a reflector satisfying $Q\mathbf{x} = \mathbf{y}$.

Remarks:

- If $\mathbf{x} = \mathbf{y}$, then $Q = I$.
- Alternatively, if $\mathbf{e}_1 = \begin{bmatrix}1 & 0 & 0 & \dots & 0\end{bmatrix}^T$, then

$$\mathbf{y} = ||\mathbf{x}||_2\mathbf{e}_1.$$

  It should be noted that $\mathbf{e}_2 = \begin{bmatrix}0 & 1 & 0 & \dots & 0\end{bmatrix}$ and $\mathbf{e}_n = \begin{bmatrix}0 & 0 & 0 & \dots & 1\end{bmatrix}$.

1.3 QR Decomposition (For the 3rd Time)

We will talk about reduced QR later; for now, we will focus on full QR. The idea is that, with QR, we'll do something like

$$Q_n \dots Q_2 Q_1 A \mapsto R.$$

The idea is that, starting from A, we can multiply the reflectors multiple times until we end up with R, which is an upper-triangular matrix. This is analogous to LU decomposition, where we did

$$L_n \dots L_2 L_1 A \mapsto U.$$

Now, for QR decomposition, given $A \in \mathbb{R}^{n\times m}$ (our "tall" matrix), we want to find QR. We can rewrite A in column form,

$$A = \begin{bmatrix} c_1 & c_2 & c_3 & \dots & c_i & \dots & c_m \end{bmatrix},$$

where c_i is the ith column for $i = 1, 2, \dots, m$. Recall that we want to derive R; that is, we want an upper-triangular matrix. So, starting from the first column, we want to make all the entries under a_{11} 0. We can use a reflector mapping Q_1 to map the column,

$$c_1 \mapsto ||c_1||\mathbf{e}_1$$

where $\mathbf{e}_1 \in \mathbb{R}^n$, so that we end up with

$$Q_1 A = \begin{bmatrix} ||c_1|| & * & * & \dots & * \\ 0 & * & * & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & * & * & \dots & * \end{bmatrix} = \begin{bmatrix} ||c_1|| & * & * & \dots & * \\ 0 & \underline{*} & \underline{*} & \dots & \underline{*} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \underline{*} & \underline{*} & \dots & \underline{*} \end{bmatrix}.$$

From the above matrix, we can represent the underlined stars as a new matrix:

$$\tilde{A} = \begin{bmatrix} \underline{*} & \underline{*} & \dots & \underline{*} \\ \vdots & \vdots & \ddots & \vdots \\ \underline{*} & \underline{*} & \dots & \underline{*} \end{bmatrix} \in \mathbb{R}^{(n-1)\times(m-1)}.$$

So, if we have

$$\tilde{A} = \begin{bmatrix} \tilde{c_2} & \tilde{c_3} & \dots & \tilde{c_m} \end{bmatrix},$$

we want to define a reflector mapping

$$\tilde{Q_2} : \tilde{c_2} \mapsto ||\tilde{c_2}||\tilde{e_1}$$

where $\tilde{e_1} \in \mathbb{R}^{n-1}$. Now, define

$$Q_2 = \begin{bmatrix} 1 & 0 \\ 0 & \tilde{Q_2} \end{bmatrix}$$

so that

$$Q_2 Q_1 A = \begin{bmatrix} ||c_1|| & * & * & \dots & * \\ 0 & ||\tilde{c_2}|| & * & \dots & * \\ 0 & 0 & * & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & * & \dots & * \end{bmatrix} = \begin{bmatrix} ||c_1|| & * & * & \dots & * \\ 0 & ||\tilde{c_2}|| & * & \dots & * \\ 0 & 0 & \underline{*} & \dots & \underline{*} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \underline{*} & \dots & \underline{*} \end{bmatrix}.$$

From this, we can define

$$B = \begin{bmatrix} \underline{*} & \dots & \underline{*} \\ \vdots & \ddots & \vdots \\ \underline{*} & \dots & \underline{*} \end{bmatrix}.$$

Continuing this process, we should eventually end up with

$$Q_m \dots Q_1 A = \begin{bmatrix} ||c_1|| & * & * & \dots & * \\ 0 & ||\tilde{c_2}|| & * & \dots & * \\ 0 & 0 & ||\tilde{c_3}|| & \dots & * \\ \vdots & \vdots & \vdots & \ddots & * \\ 0 & 0 & 0 & \dots & ||\tilde{c_m}|| \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = R.$$

Note that $\tilde{Q}A = R \implies A = QR$. Then, the question becomes: how do we define Q? We can define Q as[1]

$$Q = \tilde{Q}^{-1} = \tilde{Q}^T.$$

Remarks:

- The product of orthogonal matrices is **orthogonal**.
- The inverse of orthogonal matrices is **orthogonal**.
- Note that full QR is not unique.

Now, if A has full rank and $r_{ii} > 0$ (the diagonal on the R), then the QR decomposition is unique. Note that

- If A has full rank, then A has m linearly independent columns and $\text{rank}(A) = \min\{n, m\} = m$.

[1] Recall that $\tilde{Q}$ is orthogonal.