39. Orthogonal Matrices

Definition 39.1. A matrix $Q \in M_{n\times n}(\mathbf{R})$ is called Q orthogonal if $Q^tQ = I$.

Suppose that

$$Q = \left[V_1 \middle| V_2 \middle| \cdots \middle| V_n \right] \qquad \text{where} \qquad V_i \in \mathbf{R}^n.$$

Then

$$Q^tQ = \begin{bmatrix} V_1^t \\ \hline V_2^t \\ \hline \vdots \\ \hline V_n^t \end{bmatrix} \left[V_1 \middle| V_2 \middle| \cdots \middle| V_n \right] = \begin{bmatrix} V_1^tV_1 & V_1^tV_2 & \cdots & V_1^tV_n \\ V_2^tV_1 & V_2^tV_2 & & V_2^tV_n \\ \vdots & & \ddots & \vdots \\ V_n^tV_1 & V_n^tV_2 & \cdots & V_n^tV_n \end{bmatrix} = [V_i^tV_j] = I$$

implies that

$$V_i^tV_j = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases} \tag{39.1}$$

Definition 39.2. Let $V_i \in \mathbf{R}^m$ for $i = 1, \ldots, n$. The set $\{V_1, \ldots, V_n\}$ is said to form an orthonormal set if the vectors V_i satisfy (39.1).

Theorem 39.3. *An orthonormal set of vectors is a linearly independent set.*

Proof: Part of homework assignment 7. □

Theorem 39.4. *Given a set of linearly independent vectors $\{X_1, \ldots, X_n\}$ there exists an orthonormal set of vectors $\{V_1, \ldots, V_n\}$ such that*

$$\langle X_1, \ldots, X_k \rangle = \langle V_1, \ldots, V_k \rangle \tag{39.2}$$

for every $k = 1, \ldots, n$.

Before proving Theorem 39.4 here is some notation.

Notation 39.5. The dot product of two vectors $v, w \in \mathbf{R}^n$ is given by $v \cdot w = v^tw$.

Notation 39.6. The norm of a vector $w \in \mathbf{R}^n$ is given by $\|w\| = \sqrt{w \cdot w}$.

Fact 39.7. From vector calculus we know that $v \cdot w = \|v\|\|w\| \cos\theta$ where θ is the angle between the vectors v and w.

Algorithm 39.8. [The Gram–Schmidt Algorithm] Let $X_i \in \mathbf{R}^m$ for $i = 1, \ldots, n$ be an independent set of vectors. Construct the vectors V_i as follows

$$\begin{aligned}
&Y_1 = X_1 && V_1 = Y_1/\|Y_1\| \\
&Y_2 = X_2 - V_1(V_1 \cdot X_2) && V_2 = Y_2/\|Y_2\| \\
&Y_3 = X_3 - V_1(V_1 \cdot X_3) - V_2(V_2 \cdot X_3) && V_3 = Y_3/\|Y_3\| \\
&\quad\vdots && \quad\vdots \\
&Y_n = X_n - V_1(V_1 \cdot X_n) - V_2(V_2 \cdot X_n) - \cdots - V_{n-1}(V_{m-1} \cdot X_n) && V_n = Y_n/\|Y_n\|.
\end{aligned}$$

Note that since X_1 is linearly independent, then $\|Y_1\| \neq 0$ and so the division in the first line by $\|Y_1\|$ is well defined. Since X_2 is linearly independent from X_1 and consequently V_1 then $\|Y_2\| \neq 0$. Similarly $\|Y_3\| \neq 0$ and so forth.

Proof of Theorem 39.4: We prove that the vectors V_i constructed by Algorithm 39.8 satisfy the necessary properties.

First, we show that if $i = j$ then $V_i \cdot V_j = 1$. Calculate

$$V_i \cdot V_i = \frac{Y_i}{\|Y_i\|} \cdot \frac{Y_i}{\|Y_i\|} = \frac{1}{\|Y_i\|^2} Y_i \cdot Y_i = \frac{1}{Y_i \cdot Y_i} Y_i \cdot Y_i = 1.$$

Next, we show that if $i \neq j$ then $V_i \cdot V_j = 0$. The proof is by induction on k where $i, j \leq k$. If $k = 2$ then the only two different indices are 1 and 2. Since $V_i \cdot V_j = V_j \cdot V_i$ we may assume $i = 1$ and $j = 2$. Then

$$\begin{aligned} V_1 \cdot V_2 = V_1 \cdot \frac{Y_2}{\|Y_2\|} &= \frac{1}{\|Y_2\|} V_1 \cdot \big(X_2 - V_1(V_1 \cdot X_2)\big) \\ &= \frac{1}{\|Y_2\|}\big(V_1 \cdot X_2 - (V_1 \cdot V_1)(V_1 \cdot X_2)\big) = 0 \end{aligned}$$

since $V_1 \cdot V_1 = 1$.

For induction on k suppose $V_i \cdot V_j = 0$ for all $i \neq j$ with $i, j \leq k$.

Claim that $V_i \cdot V_j = 0$ for all $i \neq j$ with $i, j \leq k + 1$.

Since the induction hypothesis already implies $V_i \cdot V_j = 0$ for all $i \neq j$ with $i, j \leq k$ we may assume that $i < j$ and $j = k + 1$. Then

$$\begin{aligned} V_i \cdot V_{k+1} = V_i \cdot \frac{Y_{k+1}}{\|Y_{k+1}\|} &= \frac{1}{\|Y_{k+1}\|} V_i \cdot \big(X_{k+1} - V_1(V_1 \cdot X_{k+1}) \cdots - V_k(V_k \cdot X_{k+1})\big) \\ &= \frac{1}{\|Y_{k+1}\|} V_i \cdot \Big(X_{k+1} - \sum_{l=1}^{k} V_l(V_l \cdot X_{k+1})\Big) \\ &= \frac{1}{\|Y_{k+1}\|}\Big(V_i \cdot X_{k+1} - \sum_{l=1}^{k}(V_i \cdot V_l)(V_l \cdot X_{k+1})\Big) \\ &= \frac{1}{\|Y_{k+1}\|}\big(V_i \cdot X_{k+1} - (V_i \cdot V_i)(V_i \cdot X_{k+1})\big) = 0 \end{aligned}$$

since $i, l \leq k$ implies $V_i \cdot V_l = 0$ for $l \neq i$ from the induction hypothesis.

Finally, we show that (39.2) holds. Clearly $V_i \in \langle X_1, \ldots, X_k \rangle$ for all $i = 1, \ldots, k$. Therefore $\langle V_1, \ldots, V_k \rangle \subseteq \langle X_1, \ldots, X_k \rangle$. Since $\{V_1, \ldots, V_k\}$ are orthogonal Theorem 39.3 implies they are linearly independent. It follows that $\dim\langle V_1, \ldots, V_k \rangle = k$. Therefore $\langle V_1, \ldots, V_k \rangle$ is a subspace that has the same dimension as $\langle X_1, \ldots X_k \rangle$. Problem 17 in chapter 3 of Mathews now implies that $\langle V_1, \ldots, V_k \rangle = \langle X_1, \ldots, X_k \rangle$. □

Remark 39.9. If one attempts to apply Algorithm 39.8 to a set of n linearly dependent vectors then at some point the algorithm must break down. Otherwise one would obtain

$$n = \dim\langle V_1, \ldots, V_n \rangle = \dim\langle X_1, \ldots, X_n \rangle < n$$

which is a contradiction. Since the only way the algorithm can break down is for $\|Y_i\| = 0$ for some $i = 1, \ldots, n$. It follows that if the set of vectors $\{X_1, \ldots, X_n\}$ is linearly dependent then $Y_i = 0$ for some $i = 1, \ldots, n$.