Chapter 6. Orthogonality

6.2 The Gram Schmidt Process

Note. In Section 6.1 we expressed a fondness (for computational simplicity) for bases of vector spaces made of pairwise orthogonal unit vectors. In this section we give a process (the Gram-Schmidt Process) which allows us to convert a given basis into a “nice” basis. The process can be computationally lengthy and is heavily based on the projections introduced in Section 6.1.

Definition. A set \( \{ \vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k \} \) of nonzero vectors in \( \mathbb{R}^n \) is orthogonal if the vectors \( \vec{v}_j \) are mutually perpendicular; that is, if \( \vec{v}_i \cdot \vec{v}_j = 0 \) for \( i \neq j \).

Theorem 6.2. Orthogonal Bases.
Let \( \{ \vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k \} \) be an orthogonal set of nonzero vectors in \( \mathbb{R}^n \). Then this set is independent and consequently is a basis for the subspace \( \text{sp}(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k) \).

Note. To find the projection of vector \( \vec{b} \) on to subspace \( W \) in Section 6.1 we were required to find a coordinate vector relative to a certain ordered basis. This can be simplified if we have an orthogonal basis for subspace \( W \), as given in the following theorem.
Theorem 6.3. Projection Using an Orthogonal Basis.

Let \( \{v_1, v_2, \ldots, v_k\} \) be an orthogonal basis for a subspace \( W \) of \( \mathbb{R}^n \), and let \( \vec{b} \in \mathbb{R}^n \). The projection of \( \vec{b} \) on \( W \) is

\[
\vec{b}_W = \text{proj}_W(\vec{b}) = \frac{\vec{b} \cdot v_1}{v_1 \cdot v_1} v_1 + \frac{\vec{b} \cdot v_2}{v_2 \cdot v_2} v_2 + \cdots + \frac{\vec{b} \cdot v_k}{v_k \cdot v_k} v_k
\]

\[
= \text{proj}_{\vec{v}_1}(\vec{b}) + \text{proj}_{\vec{v}_2}(\vec{b}) + \cdots + \text{proj}_{\vec{v}_k}(\vec{b}).
\]

Example. Page 347 Number 4.

Definition 6.3. Let \( W \) be a subspace of \( \mathbb{R}^n \). A basis \( \{q_1, q_2, \ldots, q_k\} \) for \( W \) is orthonormal if

1. \( q_i \cdot q_j = 0 \) for \( i \neq j \), and
2. \( q_i \cdot q_i = 1 \).

That is, each vector of the basis is a unit vector and the vectors are pairwise orthogonal.

Note. If \( \{q_1, q_2, \ldots, q_k\} \) is an orthonormal basis for \( W \), then

\[
\vec{b}_W = \text{proj}_W((\vec{b}) = (\vec{b} \cdot q_1)q_1 + (\vec{b} \cdot q_2)q_2 + \cdots + (\vec{b} \cdot q_k)q_k.
\]

Note. The previous note shows why it is computationally desirable to have an orthonormal basis. Notice that it only requires the computation of some dot products; recall that if we are given an arbitrary basis \( \{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k\} \) then to write \( \vec{b} \) as a linear combination of these basis elements we must solve the system of equations \( A\vec{x} = \vec{b} \) where \( A \) is a matrix with columns \( \vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k \) (see Note 3.3.A, for example).
Theorem 6.4. Orthonormal Basis (Gram-Schmidt) Theorem.

Let $W$ be a subspace of $\mathbb{R}^n$, let $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k\}$ be any basis for $W$, and let

$$W_j = \text{sp}(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_j) \text{ for } j = 1, 2, \ldots, k.$$ 

Then there is an orthonormal basis $\{\vec{q}_1, \vec{q}_2, \ldots, \vec{q}_k\}$ for $W$ such that $W_j = \text{sp}(\vec{q}_1, \vec{q}_2, \ldots, \vec{q}_j)$.

Note. The proof of Theorem 6.4 is computational. We summarize the proof in the following procedure:

**Gram-Schmidt Process.**

To find an orthonormal basis for a subspace $W$ of $\mathbb{R}^n$:

1. Find a basis $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k\}$ for $W$.

2. Let $\vec{v}_1 = \vec{a}_1$. For $j = 1, 2, \ldots, k$, compute in succession the vector $\vec{v}_j$ given by subtracting from $\vec{a}_j$ its projection on the subspace generated by its predecessors.

3. The $\vec{v}_j$ so obtained form an orthogonal basis for $W$, and they may be normalized to yield an orthonormal basis.

Note. We can recursively describe the way to find $\vec{v}_j$ as:

$$\vec{v}_j = \vec{a}_j - \left( \frac{\vec{a}_j \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1} \vec{v}_1 + \frac{\vec{a}_j \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2} \vec{v}_2 + \cdots + \frac{\vec{a}_j \cdot \vec{v}_{j-1}}{\vec{v}_{j-1} \cdot \vec{v}_{j-1}} \vec{v}_{j-1} \right) = \vec{a}_j - \text{proj}_{\text{sp}(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{j-1})}(\vec{a}_j).$$

If we normalize the $\vec{v}_j$ as we go by letting $\vec{q}_j = (1/|\vec{v}_j|)|\vec{v}_j|$, then we have

$$\vec{v}_j = \vec{a}_j - ((\vec{a}_j \cdot \vec{q}_1)\vec{q}_1 + (\vec{a}_j \cdot \vec{q}_2)\vec{q}_2 + \cdots + (\vec{a}_j \cdot \vec{q}_{j-1})\vec{q}_{j-1}).$$
Note. We can geometrically illustrate the Gram-Schmidt Process in this notation in a special case as follows:

Example. Page 348 Number 10.

Note. A matrix “factorization” involves writing a given matrix $A$ as a product $A = BC$ where matrices $B$ and $C$ have some desired property. The desired properties are usually related to simplifying computations. Based on the proof of Theorem 6.4, “Orthonormal Basis (Gram-Schmidt) Theorem,” we can deduce the following.

**Corollary 1. QR-Factorization.**

Let $A$ be an $n \times k$ matrix with independent column vectors in $\mathbb{R}^n$. There exists an $n \times k$ matrix $Q$ with orthonormal column vectors and an upper-triangular invertible $k \times k$ matrix $R$ such that $A = QR$.


Note. We saw in Theorem 2.3(2), “Existence and Determination of Basis,” that every independent set of vectors in $\mathbb{R}^n$ can be enlarged to a basis for $\mathbb{R}^n$. The next corollary to Theorem 6.4 shows that a similar result holds for every set of orthogonal vectors in $\mathbb{R}^n$. 
Corollary 2. Expansion of an Orthogonal Set to an Orthogonal Basis.
Every orthogonal set of vectors in a subspace $W$ of $\mathbb{R}^n$ can be expanded if necessary to an orthogonal basis of $W$.

Examples. Page 348 Number 20, Page 340 Number 30, Page 340 Number 32, and Page 349 Number 34.

Revised: 4/23/2019