9.5: The Gram-Schmidt Orthogonalization procedure

We now come to a fundamentally important algorithm, which is called the **Gram-Schmidt orthogonalization procedure**. This algorithm makes it possible to construct, for each list of linearly independent vectors (resp. basis), a corresponding orthonormal list (resp. orthonormal basis).

**Theorem 9.5.1**

If $((v_1,\ldots,v_m))$ is a list of linearly independent vectors in $(V)$, then there exists an orthonormal list $((e_1,\ldots,e_m))$ such that

$$\Span(v_1,\ldots,v_k) = \Span(e_1,\ldots,e_k), \text{ for all } (k=1,\ldots,m).$$ \(\text{label} \{9.5.1\}\)

**Proof**

The proof is constructive, that is, we will actually construct vectors $((e_1,\ldots,e_m))$ having the desired properties. Since $((v_1,\ldots,v_m))$ is linearly independent, $(v_k \neq 0)$ for each $(k=1,\ldots,m)$. Set $e_1 = \frac{v_1}{\norm{v_1}}$. Then $e_1$ is a vector of norm 1 and satisfies Equation (9.5.1) for $(k=1)$. Next, set

$$e_2 = \frac{v_2 - \inner{v_2}{e_1}e_1}{\norm{v_2 - \inner{v_2}{e_1}e_1}}.$$ \(\text{equation}\{9.3.1\}\)

This is, in fact, the normalized version of the orthogonal decomposition Equation(9.3.1)~\eqref{eq:orthogonal decomp}. I.e.,

$$w = v_2 - \inner{v_2}{e_1}e_1,$$
where \(w \bot e_1\). Note that \(\|w\|=1\) and \(\langle w, e_2 \rangle =\langle w, v_2 \rangle\). 

Now, suppose that \((e_1, \ldots, e_{k-1})\) have been constructed such that \((e_1, \ldots, e_{k-1})\) is an orthonormal list and \(\text{Span}(v_1, \ldots, v_{k-1}) = \text{Span}(e_1, \ldots, e_{k-1})\). Then define
\[
\begin{equation*}
  e_k = \frac{v_k - \langle v_k, e_1 \rangle e_1 - \langle v_k, e_2 \rangle e_2 - \cdots - \langle v_k, e_{k-1} \rangle e_{k-1}}{\|v_k - \langle v_k, e_1 \rangle e_1 - \langle v_k, e_2 \rangle e_2 - \cdots - \langle v_k, e_{k-1} \rangle e_{k-1}\|}.
\end{equation*}
\]
Since \((v_1, \ldots, v_k)\) is linearly independent, we know that \(v_k \not\in \text{Span}(v_1, \ldots, v_{k-1})\). Hence, we also know that \(v_k \not\in \text{Span}(e_1, \ldots, e_{k-1})\). It follows that the norm in the definition of \(e_k\) is not zero, and so \(e_k\) is well-defined (i.e., we are not dividing by zero). Note that a vector divided by its norm has norm 1 so that \(\|e_k\|=1\). Furthermore,
\[
\begin{equation*}
  \langle e_k, e_i \rangle = \frac{\langle v_k, e_i \rangle - \langle v_k, e_i \rangle}{\|v_k - \langle v_k, e_1 \rangle e_1 - \langle v_k, e_2 \rangle e_2 - \cdots - \langle v_k, e_{k-1} \rangle e_{k-1}\|} = 0,
\end{equation*}
\]
for each \(1 \leq i < k\). Hence, \((e_1, \ldots, e_k)\) is orthonormal.

From the definition of \(e_k\), we see that \(v_k \in \text{Span}(e_1, \ldots, e_k)\) so that \(\text{Span}(v_1, \ldots, v_k) \subset \text{Span}(e_1, \ldots, e_k)\). Since both lists \((e_1, \ldots, e_k)\) and \((v_1, \ldots, v_k)\) are linearly independent, they must span subspaces of the same dimension and therefore are the same subspace. Hence Equation (9.5.1) holds.

Example \((PageIndex{2})\)

Take \((v_1=(1,1,0))\) and \((v_2=(2,1,1))\) in \((\mathbb{R}^3)\). The list \((v_1, v_2)\) is linearly independent (as you should verify!). To illustrate the Gram-Schmidt procedure, we begin by setting
\[
\begin{equation*}
  e_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{2}} (1,1,0).
\end{equation*}
\]
Next, set
The inner product 
\( \langle v_2, e_1 \rangle = \frac{1}{\sqrt{2}} \langle (1,1,0), (2,1,1) \rangle = \frac{3}{\sqrt{2}} \), so

\[
\begin{equation*}
\begin{align*}
\mathbf{u}_2 &= v_2 - \langle v_2, e_1 \rangle e_1 = (2,1,1) - \frac{3}{2}(1,1,0) = \frac{1}{2}(1,-1,2).
\end{align*}
\end{equation*}
\]

Calculating the norm of \( \mathbf{u}_2 \), we obtain \( \| \mathbf{u}_2 \| = \sqrt{\frac{1}{4}(1+1+4)} = \frac{\sqrt{6}}{2} \). Hence, normalizing this vector, we obtain

\[
\begin{equation*}
\begin{align*}
e_2 &= \frac{\mathbf{u}_2}{\| \mathbf{u}_2 \|} = \frac{1}{\sqrt{6}}(1,-1,2).
\end{align*}
\end{equation*}
\]

The list \( (e_1, e_2) \) is therefore orthonormal and has the same span as \( (v_1, v_2) \).

Corollary 9.5.3.

Every finite-dimensional inner product space has an orthonormal basis.

Proof

Let \( (v_1, \ldots, v_n) \) be any basis for \( \langle V \rangle \). This list is linearly independent and spans \( \langle V \rangle \). Apply the Gram-Schmidt procedure to this list to obtain an orthonormal list \( (e_1, \ldots, e_n) \), which still spans \( \langle V \rangle \) by construction. By Proposition 9.4.2, this list is linearly independent and hence a basis of \( \langle V \rangle \).

Corollary 9.5.4.

Every orthonormal list of vectors in \( \langle V \rangle \) can be extended to an orthonormal basis of \( \langle V \rangle \).

Proof

Let \( (e_1, \ldots, e_m) \) be an orthonormal list of vectors in \( \langle V \rangle \). By Proposition 9.4.2, this list is linearly independent and hence can be extended to a basis \( ((e_1, \ldots, e_m, v_1, \ldots, v_k) \) of \( \langle V \rangle \) by the Basis Extension Theorem. Now apply the Gram-Schmidt procedure to obtain a new orthonormal basis \( ((e_1, \ldots, e_m, f_1, \ldots, f_k) \) of \( \langle V \rangle \). The first \( m \) vectors do not change since they already are orthonormal. The list still spans \( \langle V \rangle \) and is linearly independent by Proposition 9.4.2 and therefore forms a basis.

Recall Theorem 7.5.3: given an operator \( T \) on a complex vector space \( \langle V \rangle \), there exists a basis \( (B) \) for \( \langle V \rangle \) such that the matrix \( M(T) \) with respect to \( (B) \) is upper triangular. We would like to extend this result to require the additional property of orthonormality.

Corollary 9.5.5
Let \((V)\) be an inner product space over \((\mathbb{F})\) and \((\mathcal{L}(V,V))\). If \((T)\) is upper-triangular with respect to some basis, then \((T)\) is upper-triangular with respect to some orthonormal basis.

**Proof**

Let \(((v_1,\ldots,v_n))\) be a basis of \((V)\) with respect to which \((T)\) is upper-triangular. Apply the Gram-Schmidt procedure to obtain an orthonormal basis \(((e_1,\ldots,e_n))\), and note that

\[
\text{\(\text{Span}(e_1,\ldots,e_k) = \text{Span}(v_1,\ldots,v_k), \quad \text{for all } 1 \leq k \leq n.\)}
\]

We proved before that \((T)\) is upper-triangular with respect to a basis \(((v_1,\ldots,v_n))\) if and only if \(((\text{Span}(v_1,\ldots,v_k))\) is invariant under \((T)\) for each \((1 \leq k \leq n)\). Since these spans are unchanged by the Gram-Schmidt procedure, \((T)\) is still upper triangular for the corresponding orthonormal basis.

**Contributors**

- Isaiah Lankham, Mathematics Department at UC Davis
- Bruno Nachtergaele, Mathematics Department at UC Davis
- Anne Schilling, Mathematics Department at UC Davis

Both hardbound and softbound versions of this textbook are available online at WorldScientific.com.