These notes correspond to Section 5.2 in the text.

**Gram-Schmidt Orthogonalization**

We have seen that it can be very convenient to have an orthonormal basis for a given vector space, in order to compute expansions of arbitrary vectors within that space. Therefore, given a non-orthonormal basis, it is desirable to have a process for obtaining an orthonormal basis from it. Fortunately, we have such a process, known as *Gram-Schmidt orthogonalization*. Suppose that we have a linearly independent, but not orthonormal, set of functions \( \{\chi_1, \chi_2, \ldots\} \) that span a given vector space \( V \). To construct an orthonormal set \( \{\varphi_1, \varphi_2, \ldots\} \) from this set, we proceed as follows. First, to obtain \( \varphi_1 \), we simply normalize \( \chi_1 \):

\[
\varphi_1 = \frac{\chi_1}{\|\chi_1\|}.
\]

Next, to obtain \( \varphi_2 \), we need to ensure that it is orthogonal to \( \varphi_1 \), and then normalize it.

As an intermediate step, we seek a function \( \psi_2 \) of the form

\[
\psi_2 = \chi_2 + c_{12} \varphi_1
\]

such that \( \langle \varphi_1 | \psi_2 \rangle = 0 \). Then, we can set \( \varphi_2 = \psi_2/\|\psi_2\| \). Taking the scalar product of both sides of the above equation with \( \varphi_1 \), we obtain

\[
0 = \langle \varphi_1 | \psi_2 \rangle = \langle \varphi_1 | \chi_2 \rangle + c_{12} \langle \varphi_1 | \varphi_1 \rangle.
\]

Because the \( \varphi_j \) are orthonormal, it follows that

\[
c_{12} = -\langle \varphi_1 | \chi_2 \rangle.
\]

We conclude that \( \varphi_2 \) can be obtained as follows:

\[
\begin{align*}
\psi_2 &= \chi_2 - \langle \varphi_1 | \chi_2 \rangle \varphi_1 \\
\varphi_2 &= \frac{\psi_2}{\|\psi_2\|}.
\end{align*}
\]

We now have a set of two functions that is orthonormal.

Now, to obtain \( \varphi_3 \), we must ensure that it is orthogonal to \( \varphi_1 \) and \( \varphi_2 \), and then normalized. To that end, we seek a function \( \psi_3 \) of the form

\[
\psi_3 = \chi_3 + c_{13} \varphi_1 + c_{23} \varphi_2
\]

such that \( \langle \varphi_1 | \psi_3 \rangle = \langle \varphi_2 | \psi_3 \rangle = 0 \). Then, we can set \( \varphi_3 = \psi_3/\|\psi_3\| \). Taking the scalar product of both sides of the above equation with \( \varphi_1 \), and then separately, \( \varphi_2 \), we obtain

\[
\begin{align*}
0 &= \langle \varphi_1 | \psi_3 \rangle = \langle \varphi_1 | \chi_3 \rangle + c_{13} \langle \varphi_1 | \varphi_1 \rangle + c_{23} \langle \varphi_1 | \varphi_2 \rangle \\
0 &= \langle \varphi_2 | \psi_3 \rangle = \langle \varphi_2 | \chi_3 \rangle + c_{13} \langle \varphi_2 | \varphi_1 \rangle + c_{23} \langle \varphi_2 | \varphi_2 \rangle.
\end{align*}
\]
Because the $\varphi_j$ are orthonormal, it follows that
\[ c_{13} = -\langle \varphi_1 | \chi_3 \rangle, \quad c_{23} = -\langle \varphi_2 | \chi_3 \rangle. \]

We conclude that $\varphi_3$ can be obtained as follows:
\[ \psi_3 = \chi_3 - \langle \varphi_1 | \chi_3 \rangle \varphi_1 - \langle \varphi_2 | \chi_3 \rangle \varphi_2 \]
\[ \varphi_3 = \frac{\psi_3}{\| \psi_3 \|}. \]

We now have a set of three functions that are orthonormal.

Continuing this process, we see that we can obtain each function $\varphi_j$ as follows:
\[ \psi_j = \chi_j - \sum_{k=0}^{j-1} \langle \varphi_k | \chi_j \rangle \varphi_k \]
\[ \varphi_j = \frac{\psi_j}{\| \psi_j \|}. \]

This yields a set of functions $\{ \varphi_1, \varphi_2, \ldots \}$ that is an orthonormal basis of the space spanned by $\{ \chi_1, \chi_2, \ldots \}$, with respect to the scalar product that is used.

**Example** We wish to obtain a set of orthonormal polynomials with respect to the scalar product
\[ \langle f | g \rangle = \int_{-1}^{1} f^*(s) g(s) \, ds. \]

This will be accomplished by applying Gram-Schmidt orthogonalization to the set $\{1, x, x^2, x^3, \ldots \}$. Setting $\chi_j(x) = x^j$ for $j = 0, 1, 2, \ldots$, our orthogonal set $\{ \varphi_j \}, j = 0, 1, 2, \ldots,$ is obtained as follows:
\[ \psi_0(x) = \chi_0(x) = 1, \]
\[ \varphi_0(x) = \frac{\psi_0(x)}{\| \psi_0 \|} = \frac{1}{\langle 1 | 1 \rangle^{1/2}} = \frac{1}{\sqrt{\int_{-1}^{1} 1 \, ds^{1/2}}} = \frac{1}{\sqrt{2}}, \]
\[ \psi_1(x) = \chi_1(x) - \langle \varphi_0 | \chi_1 \rangle \varphi_0(x) = x - \left\langle \frac{1}{\sqrt{2}} \right| x \right\rangle \frac{1}{\sqrt{2}} = x - \frac{1}{2} \int_{-1}^{1} s \, ds = x - \frac{1}{2} 0 = x. \]
\[ \varphi_1(x) = \frac{\psi_1(x)}{\|\psi_1\|} \]
\[ = \frac{x}{\langle x|x \rangle^{1/2}} \]
\[ = \sqrt{\frac{3}{x^2}} \]
\[ \psi_2(x) = \chi_2(x) - \langle \varphi_0|\chi_2 \rangle \varphi_0(x) - \langle \varphi_1|\chi_2 \rangle \varphi_1(x) \]
\[ = x^2 - \left( \frac{1}{\sqrt{2}} \right) x^2 \frac{1}{\sqrt{2}} - \left( \frac{3}{2}x \right) \sqrt{\frac{3}{2}}x \]
\[ = x^2 - \frac{1}{2} \int_{-1}^{1} s^2 ds - \frac{3}{2}x \int_{-1}^{1} s^3 ds \]
\[ = x^2 - \frac{1}{2} x^3 - \frac{3}{2}x^0 \]
\[ = x^2 - \frac{1}{3} \]
\[ \varphi_2(x) = \frac{\psi_2(x)}{\|\psi_2\|} \]
\[ = \frac{x^2 - \frac{1}{3}}{\langle x^2 - \frac{1}{3} | x^2 - \frac{1}{3} \rangle^{1/2}} \]
\[ = \frac{\int_{-1}^{1} (x - \frac{1}{3})^2 ds}{\int_{-1}^{1} x^2 - \frac{1}{3}} \]
\[ = \sqrt{\frac{5}{2} \left( \frac{3}{2}x^2 - \frac{1}{2} \right)} \]

Continuing this process, we obtain
\[ \varphi_3(x) = \sqrt{\frac{7}{2} \left( \frac{5}{2}x^2 - \frac{3}{2}x \right)} \]

In general,
\[ \varphi_n(x) = \sqrt{\frac{2n+1}{2}} P_n(x), \]

where \( P_n(x) \) is the Legendre polynomial of \( n \)th degree. We will learn more about these orthogonal (but not orthonormal) polynomials later in this course.

While Gram-Schmidt orthogonalization can be applied to the monomial basis \( \{1, x, x^2, x^3, \ldots \} \) to obtain an orthonormal sequence of polynomials, it can be quite cumbersome, as can be seen from the preceding example. However, a modification of this procedure can yield a much more efficient approach.

Suppose that we have already generated a sequence of \( n \) orthonormal polynomials \( \varphi_0, \varphi_1, \varphi_2, \ldots, \varphi_{n-1} \) with respect to some scalar product
\[ \langle f|g \rangle = \int_a^b f^*(x)g(x)w(x) \, dx, \]
where \( \varphi_j \) is of degree \( j \) for \( j = 0, 1, 2, \ldots, n - 1 \). Then, to obtain \( \varphi_n \), which of degree \( n \), we orthogonalize \( x\varphi_{n-1}(x) \), which is of degree \( n \), against \( \varphi_0, \varphi_1, \ldots, \varphi_{n-1} \) using the same approach as in Gram-Schmidt orthogonalization. That is, we compute

\[
\psi_n(x) = x\varphi_{n-1}(x) - \sum_{j=0}^{n-1} \langle \varphi_j | x\varphi_{n-1} \rangle \varphi_j(x),
\]

\[
\varphi_n(x) = \frac{\psi_n(x)}{\|\psi_n\|}.
\]

Now, consider the scalar product \( \langle \varphi_j | x\varphi_{n-1} \rangle \). Using the properties of the scalar product, we have

\[
\langle \varphi_j | x\varphi_{n-1} \rangle = \langle x\varphi_j | \varphi_{n-1} \rangle.
\]

However, because \( \varphi_0, \varphi_1, \ldots, \varphi_{n-1} \) are orthonormal, \( \langle p | \varphi_{n-1} \rangle = 0 \) if \( p(x) \) is any polynomial of degree less than \( n - 1 \). Because \( x\varphi_j(x) \) is of degree \( j + 1 \), it follows that \( \langle \varphi_j | x\varphi_{n-1} \rangle = 0 \) whenever \( j + 1 < n - 1 \), or \( j < n - 2 \). Therefore, our orthogonalization procedure simplifies to

\[
\psi_n(x) = x\varphi_{n-1}(x) - \langle \varphi_{n-2} | x\varphi_{n-1} \rangle \varphi_{n-2}(x) - \langle \varphi_{n-1} | x\varphi_{n-1} \rangle \varphi_{n-1}(x),
\]

\[
\varphi_n(x) = \frac{\psi_n(x)}{\|\psi_n\|}.
\]

That is, any family of orthogonal polynomials satisfies a *three-term recurrence relation*, in which each polynomial depends on the previous two. Table lists several families of orthogonal polynomials that can be generated from such a recurrence relation; we will see some of these families later in the course.

<table>
<thead>
<tr>
<th>Polynomials</th>
<th>Scalar Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Legendre</td>
<td>[ \int_{-1}^{1} P_n(x)P_m(x) , dx = 2\delta_{mn}/(2n + 1) ]</td>
</tr>
<tr>
<td>Shifted Legendre</td>
<td>[ \int_{0}^{1} P_n^<em>(x)P_m^</em>(x) , dx = \delta_{mn}/(2n + 1) ]</td>
</tr>
<tr>
<td>Chebyshev, first kind</td>
<td>[ \int_{-1}^{1} T_n(x)T_m(x)(1 - x^2)^{-1/2} , dx = \delta_{mn}/(2 - \delta_{n0}) ]</td>
</tr>
<tr>
<td>Shifted Chebyshev, first kind</td>
<td>[ \int_{0}^{1} T_n^<em>(x)T_m^</em>(x)(1 - x)^{-1/2} , dx = \delta_{mn}/(2 - \delta_{n0}) ]</td>
</tr>
<tr>
<td>Chebyshev, second kind</td>
<td>[ \int_{-1}^{1} U_n(x)U_m(x)(1 - x^2)^{1/2} , dx = \delta_{mn}/2 ]</td>
</tr>
<tr>
<td>Leguerre</td>
<td>[ \int_{0}^{\infty} L_n(x)L_m(x)e^{-x} , dx = \delta_{mn} ]</td>
</tr>
<tr>
<td>Associated Laguerre</td>
<td>[ \int_{0}^{\infty} L_n^k(x)L_m^k(x)e^{-x} , dx = \delta_{mn}(n + k)!/n! ]</td>
</tr>
<tr>
<td>Hermite</td>
<td>[ \int_{-\infty}^{\infty} H_n(x)H_m(x)e^{-x^2} , dx = 2^m\delta_{mn}\sqrt{\pi}n! ]</td>
</tr>
</tbody>
</table>

As can be seen in the following example, Gram-Schmidt orthogonalization can be applied to vectors in *any* inner product space, such as vectors in \( \mathbb{R}^n \).

**Example** Given the vectors in \( \mathbb{R}^3 \),

\[
|a_1\rangle = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \quad |a_2\rangle = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}, \quad |a_3\rangle = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix},
\]

we will use Gram-Schmidt orthogonalization to obtain an orthonormal set of vectors, \{\(|b_1\rangle, |b_2\rangle, |b_3\rangle\)\}. We have

\[
|b_1\rangle = \frac{|a_1\rangle}{\langle a_1 | a_1 \rangle^{1/2}}
\]
\[
|b'\rangle = |a_2\rangle - \langle b_1|a_2\rangle|b_1\rangle
= \left( \begin{array}{c} 1 \\ 2 \\ -3 \end{array} \right) - \frac{9}{\sqrt{6}} \frac{1}{\sqrt{6}} \left( \begin{array}{c} 1 \\ 1 \\ -2 \end{array} \right)
= \left( \begin{array}{c} -1/2 \\ 1/2 \\ 0 \end{array} \right),
\]
\[
|b_2\rangle = \frac{|b'\rangle}{\langle b'\rangle|b'\rangle}^{1/2}
= \frac{1}{\langle \frac{1}{2} \rangle^{1/2}}|b'\rangle
= \frac{1}{\sqrt{2}} \left( \begin{array}{c} -1 \\ 1 \\ 0 \end{array} \right),
\]
\[
|b'_3\rangle = |a_3\rangle - \langle b_1|a_3\rangle|b_1\rangle - \langle b_2|a_3\rangle|b_2\rangle
= \left( \begin{array}{c} 0 \\ 1 \\ 1 \end{array} \right) - \left( -\frac{1}{\sqrt{6}} \right) \frac{1}{\sqrt{6}} \left( \begin{array}{c} 1 \\ 1 \\ -2 \end{array} \right) - \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} \left( \begin{array}{c} -1 \\ 1 \\ 0 \end{array} \right)
= \left( \begin{array}{c} 0 \\ 1 \\ 1 \end{array} \right) + \frac{1}{6} \left( \begin{array}{c} 1 \\ 1 \\ -2 \end{array} \right) - \frac{1}{2} \left( \begin{array}{c} -1 \\ 1 \\ 0 \end{array} \right)
= \frac{2}{3} \left( \begin{array}{c} 1 \\ 1 \\ 1 \end{array} \right),
\]
\[
|b_3\rangle = \frac{|b'_3\rangle}{\langle b'_3\rangle|b'_3\rangle}^{1/2}
= \frac{1}{\langle \frac{2}{3} \rangle^{1/2}}|b'_3\rangle
= \frac{1}{\sqrt{3}} \left( \begin{array}{c} 1 \\ 1 \\ 1 \end{array} \right).
\]

We conclude that our orthonormal set of vectors is
\[
|b_1\rangle = \frac{1}{\sqrt{6}} \left( \begin{array}{c} 1 \\ 1 \\ -2 \end{array} \right), \quad |b_2\rangle = \frac{1}{\sqrt{2}} \left( \begin{array}{c} -1 \\ 1 \\ 0 \end{array} \right), \quad |b_3\rangle = \frac{1}{\sqrt{3}} \left( \begin{array}{c} 1 \\ 1 \\ 1 \end{array} \right).
\]

△