7.9. Orthonormal basis and the Gram-Schmidt Process

We can find an orthonormal basis for any vector space using Gram-Schmidt process. Such bases are very useful. Orthogonal projections can be computed using dot products.

Fourier series, wavelets, and so on from these.
Orthogonal basis. Orthonormal basis

- **Orthogonal basis**: A basis that is an orthogonal set.
- **Orthonormal basis**: A basis that is an orthonormal set.
- **Example 1**: \{(0,1,0), (1,0,1), (-1,0,1)\}
- **Example 2**: \{(3/7,-6/7,2/7),(2/7,3/7,6/7), (6/7,2/7,-3/7)\}
- **Example 3**: The standard basis of \(\mathbb{R}^n\).
Theorem 7.9.1  An orthogonal set of nonzero vectors in $\mathbb{R}^n$ is linearly independent.

- Proof: $v_1, v_2, \ldots, v_k$ Orthogonal set.
  - Suppose $c_1v_1 + c_2v_2 + \ldots + c_kv_k = 0$.
  - Dot with $v_1$. $c_1v_1.v_1 = 0$. Since $v_1$ has nonzero length, $c_1 = 0$.
  - Do for each $v_j$s. Thus all $c_j = 0$.

- Thus an orthogonal (orthonormal) set of $n$ nonzero vectors is a basis always.

How to find these?
Orthogonal projections using orthonormal projections

- Proj$_W$x $= M(M^T M)^{-1}M^T(x)$.
- Recall M has columns that form a basis of W.
- Suppose we chose the orthonormal basis of W.
- $M^T M = I$ by orthonormality.
- Thus Proj$_w(x) = MM^Tx$.
- $P = MM^T$.
- Example 4.
Theorem 7.9.2

(a) If \( \{v_1, v_2, \ldots, v_k\} \) is an orthonormal basis for a subspace \( W \) of \( \mathbb{R}^n \), then the orthogonal projection of a vector \( x \) in \( \mathbb{R}^n \) onto \( W \) can be expressed as

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

(7)

(b) If \( \{v_1, v_2, \ldots, v_k\} \) is an orthogonal basis for a subspace \( W \) of \( \mathbb{R}^n \), then the orthogonal projection of a vector \( x \) in \( \mathbb{R}^n \) onto \( W \) can be expressed as

\[
\text{proj}_W x = \frac{x \cdot v_1}{\|v_1\|^2}v_1 + \frac{x \cdot v_2}{\|v_2\|^2}v_2 + \cdots + \frac{x \cdot v_k}{\|v_k\|^2}v_k
\]  

(8)

• Proof: (a) \( M = [v_1, v_2, \ldots, v_k] \).

\[
\text{proj}_W x = M(M^T x) = [v_1, v_2, \ldots, v_k] \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_k^T \end{bmatrix} x
\]

\[
= [v_1, v_2, \ldots, v_k] \begin{bmatrix} v_1^T x \\ v_2^T x \\ \vdots \\ v_k^T x \end{bmatrix} = (x \cdot v_1)v_1 + (x \cdot v_2)v_2 + \ldots + (x \cdot v_k)v_k
\]
• Proof (b): Divide by the lengths to obtain an orthonormal basis of W. Apply (a).

• Note: Even if \( W = \mathbb{R}^n \), one can use the same formula.

**Theorem 7.9.3** If \( P \) is the standard matrix for an orthogonal projection of \( \mathbb{R}^n \) onto a subspace of \( \mathbb{R}^n \), then \( \text{tr}(P) = \text{rank}(P) \).

• Proof: \( P = MM^T = v_1v_1^T + \ldots + v_kv_k^T \).
  • \( \text{tr}P = \text{tr}(v_1v_1^T) + \ldots + \text{tr}(v_kv_k^T) = v_1.v_1 + \ldots + v_k.v_k = k \)
  • This by Formula 27 in Sec 3.1.

• Example 7: \( \frac{13}{49} + \frac{45}{49} + \frac{40}{49} = 2 \) (Example 4)
Linear combinations of orthonormal basis vectors.

- If \( w \) is in \( W \), then \( \text{proj}_W(w) = w \). In particular, if \( W = \mathbb{R}^n \), and \( w \) any vector, we have

\[
\text{Theorem 7.9.4} \\
(a) \text{ If } \{v_1, v_2, \ldots, v_k\} \text{ is an orthonormal basis for a subspace } W \text{ of } \mathbb{R}^n, \text{ and if } w \text{ is a vector in } W, \text{ then} \\
w = (w \cdot v_1)v_1 + (w \cdot v_2)v_2 + \cdots + (w \cdot v_k)v_k \quad (11)
(b) \text{ If } \{v_1, v_2, \ldots, v_k\} \text{ is an orthogonal basis for a subspace } W \text{ of } \mathbb{R}^n, \text{ and if } w \text{ is a vector in } W, \text{ then} \\
w = \frac{w \cdot v_1}{\|v_1\|^2}v_1 + \frac{w \cdot v_2}{\|v_2\|^2}v_2 + \cdots + \frac{w \cdot v_k}{\|v_k\|^2}v_k \quad (12)
The above formula is very useful to find “coordinates” given an orthonormal basis.

Example 8:
Gram-Schmidt orthogonalization process

Let \( v_1 = w_1 \).

We will produce nonzero subspace \( \{w_1, w_2, \ldots, w_k\} \) and basis \( \{v_1, v_2, \ldots, v_k\} \) such that \( \{v_1, v_2, \ldots, v_k\} \) is an orthonormal basis for \( \{w_1, w_2, \ldots, w_k\} \).

\[ v_2 = w_2 - \text{proj}_{v_1}(w_2) = w_2 - \frac{w_2 \cdot v_1}{||v_1||^2} v_1. \]

\( v_2 \) is not zero. (Otherwise, \( w_2 = \text{proj}_{v_1}(w_2). \) \)

\( \{v_1, v_2\} \) orthogonal set. Let \( W_2 = \text{Span}\{v_1, v_2\} \).

\[ v_3 = w_3 - \text{proj}_{W_2}(w_3) = w_3 - \frac{w_3 \cdot v_1}{||v_1||^2} v_1 - \frac{w_3 \cdot v_2}{||v_2||^2} v_2. \]

\( v_3 \) is nonzero since \( w_3 \) is not in \( W_2 \) by independence of \( \{w_1, w_2, w_3\} \). \( v_3 \) is orthogonal to \( v_1 \) and \( v_2 \).
• We obtained orthogonal set of $v_1, v_2, ..., v_l$. Let $W_l = \text{Span}\{v_1, ..., v_l\}$.

• $v_{l+1} = w_{l+1} - \text{proj}_{W_l}(w_{l+1}) = w_{l+1} - v_1(w_{l+1}.v_1)/||v_1||^2 - ... - v_l(w_{l+1}.v_l)/||v_l||^2$

• Then $v_{l+1}$ is not 0 since $w_{l+1}$ is not in $W_l$.

• $v_{l+1}$ is orthogonal to $v_1, ..., v_l$.
  • $v_i.(w_{l+1} - v_1(w_{l+1}.v_1)/||v_1||^2 - ... - v_l(w_{l+1}.v_l)/||v_l||^2 = v_i.w_{l+1} - v_i.v_i (w_{l+1}.v_i)/||v_i||^2=0$ for $i=1, ..., l$.

• Finally, we achieve $v_1, v_2, ..., v_k$.

• We can normalize to obtain an orthonormal basis.
Example 9: \((0,0,0,1),(0,0,1,1),(0,1,1,1),(1,1,1,1)\).

Example 10: \(x+y+z+2t = 0, \ 2x+y+z+t=0\).

Properties:

**Theorem 7.9.6** If \(S = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}\) is a basis for a nonzero subspace of \(\mathbb{R}^n\), and if \(S' = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}\) is the corresponding orthogonal basis produced by the Gram–Schmidt process, then:

(a) \(\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_j\}\) is an orthogonal basis for \(\text{span}\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_j\}\) at the \(j\)th step.

(b) \(\mathbf{v}_j\) is orthogonal to \(\text{span}\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{j-1}\}\) at the \(j\)th step (\(j \geq 2\)).
Extending the orthonormal set to orthonormal basis.

Theorem 7.9.7  If $W$ is a nonzero subspace of $\mathbb{R}^n$, then:

(a) Every orthogonal set of nonzero vectors in $W$ can be enlarged to an orthogonal basis for $W$.

(b) Every orthonormal set in $W$ can be enlarged to an orthonormal basis for $W$.

• Proof (a): Given $v_1, \ldots, v_k$. Add $v_{k+1}$ orthogonal to $\text{Span}\{v_1, \ldots, v_k\}$. Add $v_{k+2}$ orthogonal to $\text{Span}\{v_1, v_2, \ldots, v_k, v_{k+1}\}$. By induction....

• Proof (b): see book