Orthogonal and Orthonormal Bases With Homework #2

This handout discusses orthogonal and orthonormal bases of a finite-dimensional real vector space. (Later, we will have to consider the case of vector spaces over the complex numbers.) To be definite, we will consider the vector space $\mathbb{R}^n$, where n is a positive integer. You can consider the elements of this space to be signals and transformations such as the k-level Haar Transform H_k to be linear transformations from $\mathbb{R}^n$ to itself.

Some review of linear algebra: Recall that a set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ of elements of $\mathbb{R}^n$ is said to be **linearly independent** if whenever $\sum_{i=1}^{k} c_i\mathbf{x}_i = \mathbf{0}$ for some $c_1, c_2, \ldots, c_k \in \mathbb{R}$, it follows that $c_1 = c_2 = \cdots = c_k = 0$. That is, the only way for a linear combination of $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ to be zero is for all the coefficients in the linear combination to be zero. If $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ is a linearly independent set of vectors in $\mathbb{R}^n$, then $k \le n$, and $k = n$ if and only if $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ is a **basis** of $\mathbb{R}^n$.

Given a basis $B = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ of $\mathbb{R}^n$, any element $\mathbf{v} \in \mathbb{R}^n$ can be written *uniquely* as a linear combination $\mathbf{v} = \sum_{i=1}^{n} c_i\mathbf{x}_i$, where the c_i are real numbers. The numbers c_i are called the **components** of $\mathbf{v}$ relative to the basis B.

Finally, we need some facts about scalar products: If $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$, then the **scalar product** of $\mathbf{x}$ and $\mathbf{y}$ is defined to be $\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} x_iy_i$. An important fact about scalar products is that they are linear in both arguments. Thus, for example, if $s, t \in \mathbb{R}$ and $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$, then $\mathbf{x} \cdot (s\mathbf{y} + t\mathbf{z}) = s(\mathbf{x} \cdot \mathbf{y}) + t(\mathbf{x} \cdot \mathbf{z})$. The scalar product is also known as the "dot product," because of the usual notation. However, there is another common notation for scalar product that we might use sometimes: $\langle \mathbf{x}, \mathbf{y} \rangle$. A third common term for the scalar product is "inner product."

Definition: Two vectors $\mathbf{x}$ and $\mathbf{y}$ are said to be **orthogonal** if $\mathbf{x} \cdot \mathbf{y} = 0$, that is, if their scalar product is zero.

Theorem: Suppose $\mathbf{x}_1$, $\mathbf{x}_2$, …, $\mathbf{x}_k$ are non-zero vectors in $\mathbb{R}^n$ that are pairwise orthogonal (that is, $\mathbf{x}_i \cdot \mathbf{x}_j = 0$ for all $i \neq j$). Then the set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ is a lineary independent set of vectors.

Proof: Suppose that we have $c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_k\mathbf{x}_k = \mathbf{0}$ for some scalers $c_1, c_2, \ldots, c_k \in \mathbb{R}$. Let i be any integer in the range $1 \le i \le k$. We must show that $c_i = 0$. Now, consider the dot product of c_i with $c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_k\mathbf{x}_k$. Since $c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_k\mathbf{x}_k = \mathbf{0}$, we have:

$$\begin{aligned} 0 &= \mathbf{x}_i \cdot \mathbf{0} \\ &= \mathbf{x}_i \cdot (c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_k\mathbf{x}_k) \\ &= c_1(\mathbf{x}_i \cdot \mathbf{x}_1) + c_2(\mathbf{x}_i \cdot \mathbf{x}_2) + \cdots + c_2(\mathbf{x}_i \cdot \mathbf{x}_k) \\ &= c_i(\mathbf{x}_i \cdot \mathbf{x}_i) \end{aligned}$$

where the last equality follows because $\mathbf{x}_i \cdot \mathbf{x}_j = 0$ for $i \neq j$. Now, since $\mathbf{x}_i$ is not the zero vector, we know that $\mathbf{x}_i \cdot \mathbf{x}_i \neq 0$. So the fact that $0 = c_i(\mathbf{x}_i \cdot \mathbf{x}_i)$ implies $c_i = 0$, as we wanted to show.

Corollary: Suppose that $B = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ is a set of n vectors in $\mathbb{R}^n$ that are pairwise orthogonal. Then A is a basis of $\mathbb{R}^n$.

Proof: This follows simply because any set of n linearly independent vectors in $\mathbb{R}^n$ is a basis.

Definition: The **length** or **norm** of a vector $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ is defined to be $|\mathbf{x}| = \sqrt{\mathbf{x} \cdot \mathbf{x}}$. (Note then than $\mathbf{x} \cdot \mathbf{x} = |\mathbf{x}|^2$.)

Definition: A basis $B = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ of $\mathbb{R}^n$ is said to be an **orthogonal basis** if the elements of B are pairwise orthogonal, that is $\mathbf{x}_i \cdot \mathbf{x}_j$ whenever $i \neq j$. If in addition $\mathbf{x}_i \cdot \mathbf{x}_i = 1$ for all i, then the basis is said to be an **orthonormal basis**. Thus, an orthonormal basis is a basis consisting of unit-length, mutually orthogonal vectors.

We introduce the notation δ_{ij} for integers i and j, defined by $\delta_{ij} = 0$ if $i \neq j$ and $\delta_{ii} = 1$. Thus, a basis $B = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ is orthonormal if and only if $\mathbf{x}_i \cdot \mathbf{x}_j = \delta_{ij}$ for all i, j.

Given a vector $\mathbf{v} \in \mathbb{R}^n$ and an orthonormal basis $B = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ of $\mathbb{R}^n$, it is very easy to find the components of $\mathbf{v}$ with respect to the basis B. In fact, the i^{th} component of $\mathbf{v}$ is simply $\mathbf{v} \cdot \mathbf{x}_i$. This is the content of the following theorem:

Theorem: Suppose that $B = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ is an orthonormal basis of $\mathbb{R}^n$ and $\mathbf{v}$ is any vector in $\mathbb{R}^n$. Then

$$\mathbf{v} = \sum_{i=1}^{n} (\mathbf{v} \cdot \mathbf{x}_i)\mathbf{x}_i$$

Proof: Since B is a basis, we can write $\mathbf{v} = \sum_{i=1}^{n} c_i \mathbf{x}_i$ for some unique $c_1, c_2, \ldots, c_n \in \mathbb{R}$. We just have to show that $c_i = \mathbf{v} \cdot \mathbf{x}_i$ for each i. But

$$\begin{aligned}
\mathbf{v} \cdot \mathbf{x}_i &= \left(\sum_{j=1}^{n} c_j \mathbf{x}_j \right) \cdot \mathbf{x}_i \\
&= \sum_{j=1}^{n} c_j (\mathbf{x}_j \cdot \mathbf{x}_i) \\
&= \sum_{j=1}^{n} c_j \delta_{ji} \\
&= c_i
\end{aligned}$$

We can look at this in terms of linear transformations. Given any basis $B = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ of $\mathbb{R}^n$, we can always define a linear transformation $T_B \colon \mathbb{R}^n \to \mathbb{R}^n$ by $T(\mathbf{v}) = (\mathbf{v} \cdot \mathbf{x}_1, \mathbf{v} \cdot \mathbf{x}_2, \ldots, \mathbf{v} \cdot \mathbf{x}_n)$. Now, suppose we are given $(t_1, t_2, \ldots, t_n) \in \mathbb{R}^n$ and we would like to find a vector $\mathbf{v}$ such that $T(\mathbf{v}) = (t_1, t_2, \ldots, t_n)$. That is, we are trying to compute $T^{-1}(t_1, t_2, \ldots, t_n)$. We know from the definition of T that $\mathbf{v} \cdot \mathbf{x}_i = t_i$ for all i. For a general basis this information does not make it easy to recover $\mathbf{v}$. However, for an *orthonormal* basis, $\mathbf{v} \cdot \mathbf{x}_i$ is just the i^{th} component of $\mathbf{v}$ relative to the basis. That is, we can write $\mathbf{v} = \sum_{i=1}^{n} (\mathbf{v} \cdot \mathbf{x}_i)\mathbf{x}_i = \sum_{i=1}^{n} t_i \mathbf{x}_i$. This could also be written as $T^{-1}(t_1, t_2, \ldots, t_n) = \sum_{i=1}^{n} t_i \mathbf{x}_i$. Another way of looking at this is to say that if we know the transformed vector $T(\mathbf{v})$ of an unknown vector $\mathbf{v}$, we have a simple explicit formula for reconstructing $\mathbf{v}$.

If we apply this to the k-level Haar Transform $H_k \colon \mathbb{R}^N \to \mathbb{R}^N$, we know that the components of $H_k(\mathbf{f})$ can be computed as scalar products of $\mathbf{f}$ with the vectors $\mathbf{V}_i^k$ for $i = 1, 2, \ldots, \frac{N}{2^k}$ and $\mathbf{W}_i^j$ for $j = 1, 2, \ldots, k$ and $i = 1, 2, \ldots, \frac{N}{2^j}$:

$$\begin{aligned}
H_k(\mathbf{f}) &= \left(a_1^k, a_2^k, \ldots, a_{N/2^k}^k \mid d_1^k, d_2^k, \ldots, d_{N/2^k}^k \mid d_1^{k-1}, d_2^{k-1}, \ldots, d_{N/2^{k-1}}^{k-1} \mid \cdots \mid d_1^1, d_2^1, \ldots, d_{N/2}^1\right) \\
a_i^k &= \mathbf{f} \cdot \mathbf{V}_i^k, \text{ for } i = 1, 2, \ldots, N/2^k \\
d_i^j &= \mathbf{f} \cdot \mathbf{W}_i^j, \text{ for } j = 1, 2, \ldots, k \text{ and } i = 1, 2, \ldots, N/2^j
\end{aligned}$$

Now, it turns out that this set of vectors $\mathbf{V}_i^k$ and $\mathbf{W}_i^j$ forms an orthonormal basis of $\mathbb{R}^N$. Knowing this, we know automatically how to reconstruct a signal $\mathbf{f}$ form its k-level Haar transform. Namely, if $\mathbf{f}$ is a signal such that

$$H_k(\mathbf{f}) = \left(a_1^k, a_2^k, \ldots, a_{N/2^k}^k \,|\, d_1^k, d_2^k, \ldots, d_{N/2^k}^k \,|\, d_1^{k-1}, d_2^{k-1}, \ldots, d_{N/2^{k-1}}^{k-1} \,|\, \cdots \,|\, d_1^1, d_2^1, \ldots, d_{N/2}^1\right)$$

then

$$\mathbf{f} \;=\; \sum_{i=1}^{N/2^k} a_i^k \mathbf{V}_i^k \;+\; \sum_{j=1}^{k} \sum_{i=1}^{N/2^j} d_i^j \mathbf{W}_i^j$$

(Admittedly, the naming and indexing of the basis vectors is sort of strange, but you shouldn't let this obscure the fundamental simplicity of the result.)

The fact that we can so easily reconstruct a signal given its transform shows why it has been considered so useful to construct orthonormal bases of wavelets (and scaling functions). The case of Haar wavelets is relatively straightforward, and its orthonormality has been known and understood for a century. Finding other systems of wavelets that have the orthonormality property has not been so easy. In the 1980s, a general method for constructing orthonormal bases of wavelets was discovered. This is one of the breakthroughs that has incited a lot of the recent interest in wavelet theory.

Homework Exercises — due Friday, February 3

1. Suppose that $B = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ is an orthogonal basis of $\mathbb{R}^n$, but not necessarily an orthonormal basis. Suppose that $\mathbf{v} \in \mathbb{R}^n$. Find a simple formula for the components of $\mathbf{v}$ relative to the basis B. That is, suppose that $\mathbf{v} = \sum_{i=1}^n c_i \mathbf{x}_i$. Find a formula for c_i. (The formula should use scalar products.) Justify your answer.

2. Show that the vectors $\mathbf{V}_i^k$ for $i = 1, 2, \ldots, \frac{N}{2^k}$ and $\mathbf{W}_i^j$ for $j = 1, 2, \ldots, k$ and $i = 1, 2, \ldots, \frac{N}{2^j}$ form an orthonormal basis of $\mathbb{R}^N$. This will take some work. A useful concept is the **support** of a vector $\mathbf{v} \in \mathbb{R}^N$. The support of $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ is defined to be the set of indices for which v_i is non-zero. That is, support$(\mathbf{v}) = \{i \,|\, v_i \neq 0\}$. Note that given any two vectors in the set that you are considering, it is the case that *either* their supports do not overlap at all *or* the support of one of the two vectors is contained inside the support of other vector. (What can you say about the dot product of two vectors whose supports do not overlap?)

The remaining exercises use the Java application HaarTransformDemo.jar. You can find it in the directory */classes/s06/math371* on the computers in the Math/CS lab, or you can download it from a link on the course web page. The program can draw Haar k-level transforms of an input signal of length 2^n. The number of points can be selected using a pop-up menu at the bottom of the program window. The program can also do full or partial reconstructions of the input signal from the full n-level Haar transform, like those found in Figure 1.3 in the textbook. (A Level 0 reconstruction uses only the average a_1^n of the input signal; a level 1 transform adds in information from the level-n difference d_1^n; level 2 adds in the level-$(n-1)$ differences d_1^{n-1} and d_2^{n-1}, and so on. The Level n reconstruction should be equal to the original signal.) A pop-up menu at the bottom of the window determines which of the possible output signals is displayed.

You can change the Input Signal by dragging points. When you use the left mouse button, nearby points are dragged along. If you drag with the right-mouse button, only a single point is moved. You can also specify the input signal as a function $f(n)$.

3. Start the program. Drag the center point of the input signal upwards to create an input signal with one "hump" (or use the formula `3/( ((n-64)/16)^2 + 1 )^2` for input signal). Describe and explain in detail the 1-level and 2-level Haar Transforms of this signal.

4. Still using 128 points, use the formula `sin(pi*n/32)` for the input signal. Describe and explain the Level 0 through Level 7 Reconstructions of this signal.

5. Switch to 1024 points. Right-click a point near the center of the input signal. Discuss the Haar Transforms and Reconstructions of this signal.