Orthogonal bases

A collection of vectors $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_N\}$ in a finite dimensional vector space $\mathcal{S}$ is called an **orthogonal basis** if

1. $\text{span}(\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_N\}) = \mathcal{S}$,
2. $\boldsymbol{v}_j \perp \boldsymbol{v}_k$ (i.e. $\langle \boldsymbol{v}_j, \boldsymbol{v}_k \rangle = 0$) for all $j \neq k$.

If in addition the vectors are normalized (under the induced norm),

$$\|\boldsymbol{v}_n\| = 1, \quad \text{for} \;\; n = 1, \ldots, N,$$

we will call it an **orthonormal basis** or **orthobasis**.

A note on infinite dimensions

In infinite dimensions, we need to be a little more careful with what we mean by "span". Traditionally, the span is defined as the set of all possible linear combinations of *finitely many* elements of $\mathcal{S}$. Thus, if $\mathcal{B} = \{\boldsymbol{v}_n\}_{n \in \mathbb{Z}}$ is an infinite sequence of orthogonal vectors in a Hilbert space $\mathcal{S}$, it is an orthobasis if the *closure* of $\text{span}(\mathcal{B})$ is $\mathcal{S}$; this is written

$$\text{cl}\,\text{Span}\left(\{\boldsymbol{v}_n\}_n\right) = \mathcal{S}.$$

We don't need to get into too much, but basically this means that every vector in $\mathcal{S}$ can be approximated arbitrarily well by a finite linear combination of vectors in $\mathcal{B}$.

Here is an example which illustrates the point: Let $x(t)$ be any function on $[0, 1]$ which is not a polynomial — say $x(t) = \sin(2\pi t)$. Let $\mathcal{B} = \{1, t, t^2, t^3, \ldots\}$; the span (set of a finite linear combinations of elements) of $\mathcal{B}$ is all polynomials on $[0, 1]$. So $\boldsymbol{x} \notin \text{span}(\mathcal{B})$. But $x(t)$ can be approximated arbitrarily well by elements in $\mathcal{B}$ (using higher and higher order polynomials) so $\boldsymbol{x} \in \text{cl}\,\text{Span}(\mathcal{B})$).

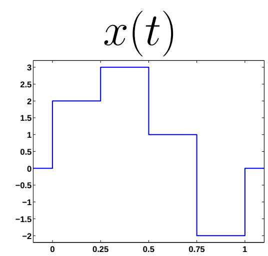
Examples.

1. $\mathcal{S} = \mathbb{R}^2$, equipped with the standard inner product

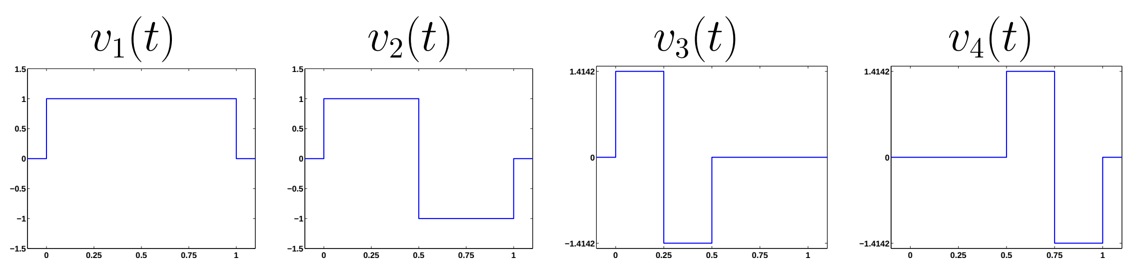
$$\boldsymbol{v}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}, \qquad \boldsymbol{v}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}$$

2. $\mathcal{S}$ = space of piecewise constant functions on $[0, 1/4),\ [1/4, 1/2),\ [1/2, 3/4),\ [3/4, 1]$

   Example signal:



   The following four functions form an orthobasis for this space



3. **Fourier series**

$$\left\{ v_k(t) = \frac{1}{\sqrt{2\pi}} e^{jkt} \ , \ k \in \mathbb{Z} \right\} \quad \text{is an orthobasis for } L_2([0, 2\pi])$$

(with the standard inner product).

Let's quickly check the orthogonality:

$$\begin{aligned} \left\langle \frac{1}{\sqrt{2\pi}} e^{jk_1 t}, \frac{1}{\sqrt{2\pi}} e^{jk_2 t} \right\rangle &= \frac{1}{2\pi} \int_0^{2\pi} e^{j(k_1 - k_2)t} \ \mathrm{d}t \\ &= \begin{cases} 1, & k_1 = k_2 \\ 0, & k_1 \neq k_2 \end{cases}. \end{aligned}$$

It is also true that the closure of $\operatorname{span}(\{(2\pi)^{-1/2} e^{jkt}\}_{k=-\infty}^{\infty})$ is $L_2([0, 2\pi])$. The proof of this is a bit involved; if you are interested, see Chapter 5 of Young's *Introduction to Hilbert Space*.

4. **Sampling**

$B_{\pi/T}(\mathbb{R})$ = real-valued functions which are bandlimited to π/T, equipped with the standard inner product. The set of functions

$$\left\{v_n(t) = \sqrt{T}\,\frac{\sin(\pi(t-nT)/T)}{\pi(t-nT)},\ n \in \mathbb{Z}\right\}$$

is an orthobasis for $B_{\pi/T}(\mathbb{R})$. (Notice that we have a slightly different normalization than when we looked at the sampling theorem — we have a $\sqrt{T}$ out front instead of T.)

Check the orthogonality:

$$\left\langle \sqrt{T}\frac{\sin(\pi(t-n_1T)/T)}{\pi(t-n_1T)}\ ,\ \sqrt{T}\frac{\sin(\pi(t-n_2T)/T)}{\pi(t-n_2T)} \right\rangle$$

$$= \frac{1}{2\pi}\int_{-\pi/T}^{\pi/T} T e^{-j\Omega T n_1} e^{j\Omega T n_2}\ \mathrm{d}\Omega \qquad \text{(Parseval)}$$

$$= \frac{T}{2\pi}\int_{-\pi/T}^{\pi/T} e^{j\Omega T(n_1-n_2)}\ \mathrm{d}\Omega$$

$$= \begin{cases} 1, & n_1 = n_2 \\ 0, & n_1 \neq n_2 \end{cases}.$$

That the (closure of the) span of this set is $B_{\pi/T}(\mathbb{R})$ is essentially the content of the Shannon-Nyquist sampling theorem.

Again, sampling $x(t) \in B_{\pi/T}(\mathbb{R})$ is equivalent to a Fourier Series analysis of $X(j\Omega)$ on $[-\pi/T, \pi/T]$.

5. **Legendre Polynomials**

Define

$$p_0(t) = 1, \quad p_1(t) = t,$$

and then for $n = 1, 2, \ldots$

$$p_{n+1}(t) = \frac{2n+1}{n+1}\, t\, p_n(t) - \frac{n}{n+1} p_{n-1}(t),$$

and so

$$\begin{aligned} p_2(t) &= \frac{1}{2}(3t^2 - 1) \\ p_3(t) &= \frac{1}{2}(5t^3 - 3t) \\ p_4(t) &= \frac{1}{8}(35t^4 - 30x^2 + 3) \\ &\vdots \qquad \text{etc.} \end{aligned}$$

These $p_n(t)$ are called *Legendre polynomials*, and if we renormalize them, taking

$$v_n(t) = \sqrt{\frac{2n+1}{2}}\, p_n(t),$$

then $v_0(t), \ldots, v_N(t)$ are an orthobasis for polynomials of degree N on $[-1, 1]$.

Computing approximations with the Legendre basis is far more stable than computing the approximation in the standard basis.

Linear approximation and orthobases

Let's return to our linear approximation problem:
Given $\boldsymbol{x} \in \mathcal{S}$, we want to find the closest point in a subspace $\mathcal{T}$.

Suppose we have an orthobasis $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N\}$ for $\mathcal{T}$. Then solving this problem is easy. Here's why: we know the solution is

$$\hat{\boldsymbol{x}} = a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \cdots + a_N\boldsymbol{v}_N \tag{1}$$

where the a_n are given by

$$\begin{bmatrix} a_1 \\ \vdots \\ a_N \end{bmatrix} = \boldsymbol{G}^{-1}\boldsymbol{b}, \quad \text{with } \boldsymbol{G} = \begin{bmatrix} \langle \boldsymbol{v}_1, \boldsymbol{v}_1 \rangle & \cdots & \langle \boldsymbol{v}_N, \boldsymbol{v}_1 \rangle \\ \vdots & \ddots & \vdots \\ \langle \boldsymbol{v}_1, \boldsymbol{v}_N \rangle & \cdots & \langle \boldsymbol{v}_N, \boldsymbol{v}_N \rangle \end{bmatrix}, \ \boldsymbol{b} = \begin{bmatrix} \langle \boldsymbol{x}, \boldsymbol{v}_1 \rangle \\ \vdots \\ \langle \boldsymbol{x}, \boldsymbol{v}_N \rangle \end{bmatrix}$$

Now since $\langle \boldsymbol{v}_n, \boldsymbol{v}_k \rangle = 1$ if $n = k$ and 0 otherwise, $\boldsymbol{G} = \mathbf{I}$ (the identity matrix), and so $\boldsymbol{G}^{-1} = \mathbf{I}$ as well, and

$$\begin{bmatrix} a_1 \\ \vdots \\ a_N \end{bmatrix} = \begin{bmatrix} \langle \boldsymbol{x}, \boldsymbol{v}_1 \rangle \\ \vdots \\ \langle \boldsymbol{x}, \boldsymbol{v}_N \rangle \end{bmatrix}. \tag{2}$$

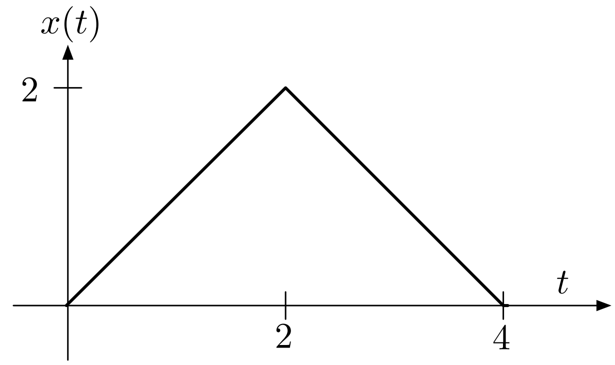
So calculating the closest point is as easy as computing N inner products — no matrix inversion necessary.

Combining the expressions (1) and (2) gives us the compact expression

$$\hat{\boldsymbol{x}} = \sum_{n=1}^{N} \langle \boldsymbol{x}, \boldsymbol{v}_n \rangle \, \boldsymbol{v}_n.$$

Example. Suppose $x(t) \in L_2([0,4])$ is

$$x(t) = \begin{cases} t & 0 \le t \le 2 \\ 4 - t & 2 \le t \le 4 \end{cases}$$



Let $\mathcal{T}$ = piecewise constant functions on $[0,1),\ [1,2),\ [2,3),\ [3,4]$.

Find the closest point in $\mathcal{T}$ to $\boldsymbol{x}$. A good orthobasis to use is

$$v_n(t) = \begin{cases} 1 & (n-1) \le t \le n \\ 0 & \text{otherwise} \end{cases}, \quad n = 1, 2, 3, 4.$$

Orthobasis expansions

The orthogonality principle (easily) gives us an expression for the **expansion coefficients** of a vector in an orthobasis.

Suppose a finite dimensional space $\mathcal{S}$ has an orthobasis $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$. Given any $\boldsymbol{x} \in \mathcal{S}$, the closest point in $\mathcal{S}$ to $\boldsymbol{x}$ is $\boldsymbol{x}$ itself (of course). This gives us the following **reproducing formula**:

$$\boldsymbol{x} = \sum_{n=1}^{N} \langle \boldsymbol{x}, \boldsymbol{v}_n \rangle \, \boldsymbol{v}_n, \quad \text{for all } \boldsymbol{x} \in \mathcal{S}.$$

In infinite dimensions, if $\mathcal{S}$ has an orthobasis $\{\boldsymbol{v}_n\}_{n=-\infty}^{\infty}$ and $\boldsymbol{x} \in \mathcal{S}$ obeys

$$\sum_{n=-\infty}^{\infty} |\langle \boldsymbol{x}, \boldsymbol{v}_n \rangle|^2 \; < \; \infty,$$

then we can write

$$\boldsymbol{x} = \sum_{n=-\infty}^{\infty} \langle \boldsymbol{x}, \boldsymbol{v}_n \rangle \, \boldsymbol{v}_n.$$

(We need the sequence of expansion coefficients to be square-summable to make sure the sum of vectors above converges to something.)

In other words, $\boldsymbol{x} \in \mathcal{S}$ is captured without loss by the discrete list of numbers

$$\ldots, \; \langle \boldsymbol{x}, \boldsymbol{v}_{-1} \rangle, \; \langle \boldsymbol{x}, \boldsymbol{v}_0 \rangle, \; \langle \boldsymbol{x}, \boldsymbol{v}_1 \rangle, \; \ldots$$

An orthobasis gives us a natural way to discretize vectors in $\mathcal{S}$ through a set of expansion coefficients. Moreover, there is a straightforward and explicit way to compute these expansion coefficients — you simply take an inner product with the corresponding basis vector.

Example: Sampling a bandlimited function.

$B_{\pi/T}$ = space of bandlimited signals equipped with the standard inner product. We have seen already that

$$v_n(t) = \sqrt{T}\,\frac{\sin(\pi(t-nT)/T)}{\pi(t-nT)}, \quad n \in \mathbb{Z}$$

is an orthobasis for $B_{\pi/T}$. This means that any $\boldsymbol{x} \in B_{\pi/T}$ can be written

$$\boldsymbol{x} = \sum_{n=-\infty}^{\infty} \langle \boldsymbol{x}, \boldsymbol{v}_n \rangle\, \boldsymbol{v}_n.$$

What are the $\langle \boldsymbol{x}, \boldsymbol{v}_n \rangle$?

$$\begin{aligned}
\langle \boldsymbol{x}, \boldsymbol{v}_n \rangle &= \left\langle x(t)\ ,\ \sqrt{T}\,\frac{\sin(\pi(t-nT)/T)}{\pi(t-nT)} \right\rangle \\
&= \frac{1}{2\pi}\int_{-\pi/T}^{\pi/T} X(j\Omega)\,\sqrt{T}e^{jn\Omega T}\ \mathrm{d}\Omega \\
&= \sqrt{T}x(nT),
\end{aligned}$$

which is simply a sample scaled by $\sqrt{T}$. So the reproducing formula is just a restatement of the sampling theorem:

$$\begin{aligned}
x(t) &= \sum_{n=-\infty}^{\infty} \langle \boldsymbol{x}, \boldsymbol{v}_n \rangle\, \boldsymbol{v}_n \\
&= \sum_{n=-\infty}^{\infty} \sqrt{T}\,x(nT)\,\frac{\sqrt{T}\sin(\pi(t-nT)/T)}{\pi(t-nT)} \\
&= \sum_{n=-\infty}^{\infty} x(nT) g_T(t-nT).
\end{aligned}$$

The moral of the story is that we can recreate a vector in a Hilbert space from the sequence of numbers $\{\langle \boldsymbol{x}, \boldsymbol{v}_n \rangle\}$. We can think of every different orthobasis for $\mathcal{S}$ as a different **transform**, and the $\{\langle \boldsymbol{x}, \boldsymbol{v}_n \rangle\}$ as **transform coefficients**.

Next we will see that our notions of **distance** and **angle** also carry over to this discrete space.

Parseval's Theorem

One handy fact (and a fact we have used many times in this course already) about the Fourier transform is that it is **energy preserving**,

$$\|x(t)\|_2^2 = \int_{-\infty}^{\infty} |x(t)|^2 \, \mathrm{d}t = \frac{1}{2\pi}\int_{-\infty}^{\infty} |X(j\Omega)|^2 \, \mathrm{d}\Omega = \frac{1}{2\pi}\|X(j\Omega)\|_2^2,$$

and more generally, it preserves the L_2 inner product:

$$\begin{aligned}\langle x(t), y(t)\rangle &= \int_{-\infty}^{\infty} x(t)\overline{y(t)}\,\mathrm{d}t = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(j\Omega)\overline{Y(j\Omega)}\,\mathrm{d}\Omega \\ &= \frac{1}{2\pi}\langle X(j\Omega), Y(j\Omega)\rangle.\end{aligned}$$

It is not not too hard to show that something very similar is true for any orthobasis expansion. Let $\mathcal{S}$ be a Hilbert space with inner product $\langle\cdot,\cdot\rangle_S$ which induces norm $\|\cdot\|_S$. Let $\{v_k\}_{k\in\Gamma}$ be an orthobasis[1] for $\mathcal{S}$. Then for every $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{S}$,

$$\langle \boldsymbol{x}, \boldsymbol{y}\rangle_S = \sum_{k\in\Gamma} \alpha_k \overline{\beta_k},$$

where

$$\alpha_k = \langle \boldsymbol{x}, \boldsymbol{v}_k\rangle_S, \qquad \beta_k = \langle \boldsymbol{y}, \boldsymbol{v}_k\rangle_S.$$

You can think of the $\{\alpha_k\}$ as the transform coefficients of $\boldsymbol{x}$ and the $\{\beta_k\}$ as the transform coefficients of $\boldsymbol{y}$. So we have

$$\begin{aligned}\langle \boldsymbol{x}, \boldsymbol{y}\rangle_S &= \langle \boldsymbol{\alpha}, \boldsymbol{\beta}\rangle_{\ell_2}, \\ \|\boldsymbol{x}\|_S^2 &= \|\boldsymbol{\alpha}\|_2^2.\end{aligned}$$

[1] We are using Γ to be an arbitrary index set here; it can be either finite, e.g. $\Gamma = 1, 2, \ldots, N$, or infinite, e.g. $\Gamma = \mathbb{Z}$.

$\Rightarrow$ An orthobasis makes every Hilbert space **equivalent** to ℓ_2.

All of the geometry (lengths, angles) maps into standard Euclidean geometry in coefficient space. As you can imagine, this is a pretty useful fact.

Proof of Parseval. With $\alpha_k = \langle \boldsymbol{x}, \boldsymbol{v}_k \rangle_S$ and $\beta_k = \langle \boldsymbol{y}, \boldsymbol{v}_k \rangle_S$, we can write

$$\boldsymbol{x} = \sum_{k \in \Gamma} \alpha_k \, \boldsymbol{v}_k, \quad \text{and} \quad \boldsymbol{y} = \sum_{k \in \Gamma} \beta_k \, \boldsymbol{v}_k,$$

and so

$$\begin{aligned} \langle \boldsymbol{x}, \boldsymbol{y} \rangle_S &= \left\langle \sum_{k \in \Gamma} \alpha_k \boldsymbol{v}_k, \sum_{\ell \in \Gamma} \beta_\ell \boldsymbol{v}_\ell \right\rangle_S \\ &= \sum_{k \in \Gamma} \alpha_k \left\langle \boldsymbol{v}_k, \sum_{\ell \in \Gamma} \beta_\ell \boldsymbol{v}_\ell \right\rangle_S \\ &= \sum_{k \in \Gamma} \sum_{\ell \in \Gamma} \alpha_k \overline{\beta_\ell} \langle \boldsymbol{v}_k, \boldsymbol{v}_\ell \rangle_S . \end{aligned}$$

For a fixed value of k, only one term in the inner sum above will be nonzero, as $\langle \boldsymbol{v}_k, \boldsymbol{v}_\ell \rangle = 0$ unless $\ell = k$. Thus

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle_S = \sum_{k \in \Gamma} \alpha_k \overline{\beta_k}.$$

A straightforward consequence of the result above is that distances in $\mathcal{S}$ under the induced norm are equivalent to Euclidean (ℓ_2) distances in coefficient space:

$$\|\boldsymbol{x} - \boldsymbol{y}\|_S = \|\boldsymbol{\alpha} - \boldsymbol{\beta}\|_2 = \left(\sum_{k\in\Gamma}(\alpha_k - \beta_k)^2\right)^{1/2}.$$

Thus changing the value of an orthobasis expansion coefficient by an amount ϵ will change the signal by an amount (as measured in $\|\cdot\|_S$) ϵ.

To be more precise about this, suppose $\boldsymbol{x}$ has transform coefficients $\{\alpha_k = \langle \boldsymbol{x}, \boldsymbol{v}_k\rangle_S\}$. If I perturb one of them, say at location k_0, by setting

$$\tilde{\alpha}_k = \begin{cases} \alpha_{k_0} + \epsilon & k = k_0 \\ \alpha_k & k \neq k_0 \end{cases},$$

and then synthesizing

$$\tilde{\boldsymbol{x}} = \sum_{k\in\Gamma} \tilde{\alpha}_k \boldsymbol{v}_k,$$

we will have

$$\|\boldsymbol{x} - \tilde{\boldsymbol{x}}\|_S = \epsilon.$$

Notice that while the error is localized to one expansion coefficient, it could effect the entire reconstruction, but its net effect will still be ϵ.

Here is another example. Suppose I sample a signal $x_c(t)$ which is bandlimited to π/T at a rate T, producing the sample sequence $x[n] = x_c(nT)$. Each of these samples gets perturbed by a (possibly different) amount $\epsilon[n]$:

$$\tilde{x}[n] = x[n] + \epsilon[n].$$

We resynthesize the signal using sinc interpolation:

$$\tilde{x}_c(t) = \sum_{n=-\infty}^{\infty} \tilde{x}[n] h_T(t - nT),$$

and the difference between this signal and the "true" signal is

$$\begin{aligned} x_c(t) - \tilde{x}_c(t) &= \sum_{n=-\infty}^{\infty} (x[n] - \tilde{x}[n]) h_T(t - nT) \\ &= \sum_{n=-\infty}^{\infty} \sqrt{T}(x[n] - \tilde{x}[n])\, h_T(t - nT)/\sqrt{T}. \end{aligned}$$

Since the $\{h_T(t-nT)/\sqrt{T}\}_{n\in\mathbb{Z}}$ are an orthobasis for $B_{\pi/T}$, we know

$$\begin{aligned} \|x_c(t) - \tilde{x}_c(t)\|_{L_2}^2 &= \int |x_c(t) - \tilde{x}_c(t)|^2 \ \mathrm{d}t \\ &= \sum_{n=-\infty}^{\infty} \left|\sqrt{T}(x[n] - \tilde{x}[n])\right|^2 \\ &= T \sum_{n=-\infty}^{\infty} |\epsilon[n]|^2 \end{aligned}$$

The upshot of this is that as we change each sample, we know exactly what the net effect will be on the reconstruction error.