

Orthogonal bases

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

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

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

$$\|\mathbf{v}_n\| = 1, \quad \text{for } n = 1, \dots, N,$$

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 $\mathcal{B} = \{\mathbf{v}_n\}_{n \in \mathbb{Z}}$ 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

$$\text{cl Span}(\{\mathbf{v}_n\}_n) = \mathcal{S}.$$

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 $\mathcal{B} = \{1, t, t^2, t^3, \dots\}$; the span (set of a finite linear combinations of elements) of \mathcal{B} is all polynomials on $[0, 1]$. So $\mathbf{x} \notin \text{span}(\mathcal{B})$. But $x(t)$ can be approximated arbitrarily well by elements in \mathcal{B} (using higher and higher order polynomials) so $\mathbf{x} \in \text{cl Span}(\mathcal{B})$.

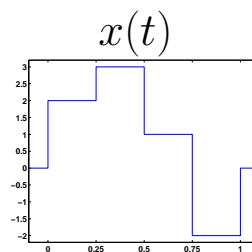
Examples.

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

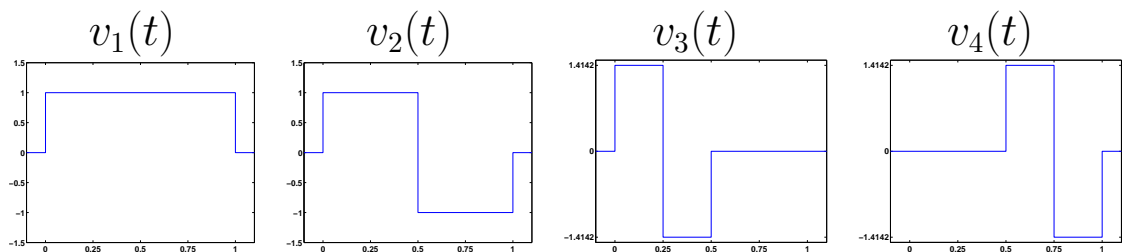
$$\mathbf{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

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

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

(with the standard inner product).

Let's quickly check the orthogonality:

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

It is also true that the closure of $\text{span}(\{(2\pi)^{-1/2} e^{jkt}\}_{k=-\infty}^{\infty})$ 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:

$$\begin{aligned} & \left\langle \sqrt{T} \frac{\sin(\pi(t - n_1T)/T)}{\pi(t - n_1T)}, \sqrt{T} \frac{\sin(\pi(t - n_2T)/T)}{\pi(t - n_2T)} \right\rangle \\ &= \frac{1}{2\pi} \int_{-\pi/T}^{\pi/T} T e^{-j\Omega T n_1} e^{j\Omega T n_2} d\Omega \quad (\text{Parseval}) \\ &= \frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} e^{j\Omega T(n_1 - n_2)} d\Omega \\ &= \begin{cases} 1, & n_1 = n_2 \\ 0, & n_1 \neq n_2 \end{cases}. \end{aligned}$$

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 $n = 1, 2, \dots$

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

and so

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

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 $v_0(t), \dots, v_N(t)$ 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 $\mathbf{x} \in \mathcal{S}$, we want to find the closest point in a subspace \mathcal{T} .

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

$$\hat{\mathbf{x}} = a_1 \mathbf{v}_1 + a_2 \mathbf{v}_2 + \dots + a_N \mathbf{v}_N \quad (1)$$

where the a_n are given by

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

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

$$\begin{bmatrix} a_1 \\ \vdots \\ a_N \end{bmatrix} = \begin{bmatrix} \langle \mathbf{x}, \mathbf{v}_1 \rangle \\ \vdots \\ \langle \mathbf{x}, \mathbf{v}_N \rangle \end{bmatrix}. \quad (2)$$

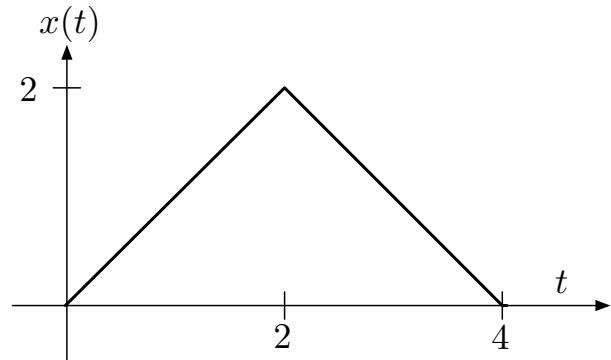
So calculating the closest point is as easy as computing N inner products — no matrix inversion necessary.

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

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

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

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



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

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

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

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 $\{\mathbf{v}_1, \dots, \mathbf{v}_N\}$. Given any $\mathbf{x} \in \mathcal{S}$, the closest point in \mathcal{S} to \mathbf{x} is \mathbf{x} itself (of course). This gives us the following **reproducing formula**:

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

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

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

then we can write

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

(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, $\mathbf{x} \in \mathcal{S}$ is captured without loss by the discrete list of numbers

$$\dots, \langle \mathbf{x}, \mathbf{v}_{-1} \rangle, \langle \mathbf{x}, \mathbf{v}_0 \rangle, \langle \mathbf{x}, \mathbf{v}_1 \rangle, \dots$$

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

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

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

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

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

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

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

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

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

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

and more generally, it preserves the L_2 inner product:

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

It is 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 orthobasis¹ for \mathcal{S} . Then for every $\mathbf{x}, \mathbf{y} \in \mathcal{S}$,

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

where

$$\alpha_k = \langle \mathbf{x}, \mathbf{v}_k \rangle_S, \quad \beta_k = \langle \mathbf{y}, \mathbf{v}_k \rangle_S.$$

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

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

¹We are using Γ to be an arbitrary index set here; it can be either finite, e.g. $\Gamma = 1, 2, \dots, N$, 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 $\alpha_k = \langle \mathbf{x}, \mathbf{v}_k \rangle_S$ and $\beta_k = \langle \mathbf{y}, \mathbf{v}_k \rangle_S$, we can write

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

and so

$$\langle \mathbf{x}, \mathbf{y} \rangle_S =$$

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:

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

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 \mathbf{x} has transform coefficients $\{\alpha_k = \langle \mathbf{x}, \mathbf{v}_k \rangle_S\}$. 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

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

we will have

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

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

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

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