Non-orthogonal bases

Although orthogonal bases have many useful properties, it is possible (and sometimes desirable) to use a set of non-orthogonal basis functions for discretization.

The main property we want from a basis is that the mapping from signal space to coefficient space is a stable bijection — each signal should have a different set of expansion coefficients, and each set of expansion coefficients should correspond to a different signal. Small changes in the expansion coefficients should not lead to large changes in the re-synthesized signal. We also want a concrete method for calculating the coefficients in a basis expansion.

We start our discussion in the familiar setting of $\mathbb{R}^N$.

Bases in $\mathbb{R}^N$

Let $\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_N \in \mathbb{R}^N$ be a basis for $\mathbb{R}^N$. We know from basic linear algebra that these vectors form a basis if and only if they are linearly independent. For any $\boldsymbol{x} \in \mathbb{R}^N$, we have

$$\boldsymbol{x} = \sum_{n=1}^{N} \alpha_n \boldsymbol{\psi}_n, \tag{1}$$

for some coefficient sequence

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{bmatrix}.$$

How do we compute the α_n?

The answer to this turns out to be straightforward as soon as we have everything written down the right way. We can write the decomposition in (1) as

$$
\begin{aligned}
\boldsymbol{x} = \alpha_1\boldsymbol{\psi}_1 + \cdots + \alpha_N\boldsymbol{\psi}_N &= \begin{bmatrix} | & | & & | \\ \boldsymbol{\psi}_1 & \boldsymbol{\psi}_2 & \cdots & \boldsymbol{\psi}_N \\ | & | & & | \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{bmatrix} \\
&= \boldsymbol{\Psi}\boldsymbol{\alpha}
\end{aligned}
$$

That is, the basis vectors are concatenated as columns in the $N \times N$ matrix $\boldsymbol{\Psi}$. Since its columns are linearly independent, $\boldsymbol{\Psi}$ is invertible, and so we have the **reproducing formula**

$$
\begin{aligned}
\boldsymbol{x} &= \boldsymbol{\Psi}\boldsymbol{\Psi}^{-1}\boldsymbol{x} \\
&= \sum_{n=1}^{N} \langle \boldsymbol{x}, \widetilde{\boldsymbol{\psi}}_n \rangle \boldsymbol{\psi}_n,
\end{aligned}
$$

where $\widetilde{\boldsymbol{\psi}}_n$ is the n^{th} **row** of the **inverse** of $\boldsymbol{\Psi}$:

$$
\boldsymbol{\Psi}^{-1} = \begin{bmatrix} - & \widetilde{\boldsymbol{\psi}}_1^{\mathrm{T}} & - \\ - & \widetilde{\boldsymbol{\psi}}_2^{\mathrm{T}} & - \\ & \vdots & \\ - & \widetilde{\boldsymbol{\psi}}_N^{\mathrm{T}} & - \end{bmatrix}
$$

We have

$$
\begin{aligned}
\textbf{Transform (analysis):} \quad & \alpha_n = \langle \boldsymbol{x}, \widetilde{\boldsymbol{\psi}}_n \rangle, \quad n = 1, \ldots, N; \\
\textbf{Inverse (synthesis):} \quad & \boldsymbol{x} = \sum_{n=1}^{N} \alpha_n \boldsymbol{\psi}_n
\end{aligned}
$$

So we compute the expansion coefficients by taking inner products against basis signals, just not the same basis signals as we are using to re-synthesize $\boldsymbol{x}$. The $\widetilde{\boldsymbol{\psi}}_1, \ldots, \widetilde{\boldsymbol{\psi}}_N$ themselves are linearly independent, and are called the **dual basis** for $\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_N$.

Also note that while the $\{\boldsymbol{\psi}_n\}$ are not orthonormal and the $\{\widetilde{\boldsymbol{\psi}}_n\}$ are not orthonormal, jointly they obey the relation

$$\langle \boldsymbol{\psi}_n, \widetilde{\boldsymbol{\psi}}_\ell \rangle = \begin{cases} 1, & n = \ell, \\ 0, & \mathrm{n} \neq \ell. \end{cases}$$

(This follows simply from the fact that $\boldsymbol{\Psi}\boldsymbol{\Psi}^{-1} = \mathbf{I}$.) For this reason, $\{\boldsymbol{\psi}_n\}$ and $\{\widetilde{\boldsymbol{\psi}}_n\}$ are called **biorthogonal bases**.

Bases for subspaces of $\mathbb{R}^N$

Suppose that $\mathcal{T}$ is a M-dimensional ($M < N$) subspace of $\mathbb{R}^N$, and $\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_M \in \mathcal{T}$ is a basis for this space. Now

$$\boldsymbol{\Psi} = \begin{bmatrix} | & & | \\ \boldsymbol{\psi}_1 & \cdots & \boldsymbol{\psi}_M \\ | & & | \end{bmatrix}$$

is $N \times M$ — it is not square, and so it is not invertible. But since the $\boldsymbol{\psi}_m$ are linearly independent, it does have a left inverse, and hence we can derive a reproducing formula.

Let $\boldsymbol{x} \in \mathcal{T}$, so there exists an $\alpha \in \mathbb{R}^M$ such that $\boldsymbol{x} = \boldsymbol{\Psi}\boldsymbol{\alpha}$. Then the reproducing formula is

$$\boldsymbol{x} = \boldsymbol{\Psi}(\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{\Psi})^{-1}\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{x}.$$

(The $M \times M$ matrix $\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{\Psi}$ is invertible by the linear independence of $\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_M$.) To see that the formula above holds, simply plug $\boldsymbol{x} = \boldsymbol{\Psi}\boldsymbol{\alpha}$ into the expression on the right hand side.

We can write

$$\boldsymbol{x} = \sum_{m=1}^{M} \langle \boldsymbol{x}, \widetilde{\boldsymbol{\psi}}_m \rangle \boldsymbol{\psi}_m,$$

where

$$\widetilde{\boldsymbol{\psi}}_m = \ m^{\text{th}} \text{ row of the } M \times N \text{ matrix } (\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{\Psi})^{-1}\boldsymbol{\Psi}^{\mathrm{T}}.$$

Notice that when $M = N$ (and so $\boldsymbol{\Psi}$ is square and invertible), this agrees with the result in the previous section, as in this case

$$(\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{\Psi})^{-1}\boldsymbol{\Psi}^{\mathrm{T}} = \boldsymbol{\Psi}^{-1}(\boldsymbol{\Psi}^{\mathrm{T}})^{-1}\boldsymbol{\Psi}^{\mathrm{T}} = \boldsymbol{\Psi}^{-1}.$$

Bases in finite dimensional spaces

The construction of the dual basis in $\mathbb{R}^N$ (which tells us how to compute the expansion coefficients for a basis) relied on manipulating matrices that contained the basis vectors. If our signals of interest are not length N vectors, but still live in a finite dimensional Hilbert space $\mathcal{S}$, then we can proceed in a similar manner.

Let $\psi_1(t), \ldots, \psi_N(t)$ be a basis for an N-dimensional inner product space $\mathcal{S}$. Let $x(t) \in \mathcal{S}$ be another arbitrary signal in this space. We know that the closest point to $\boldsymbol{x}$ in $\mathcal{S}$ is $\boldsymbol{x}$ itself, and from our work on the closest point problem, we know that we can write

$$x(t) = \sum_{n=1}^{N} \alpha_n \psi_n(t),$$

where

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{bmatrix} = \begin{bmatrix} \langle \boldsymbol{\psi}_1, \boldsymbol{\psi}_1 \rangle & \langle \boldsymbol{\psi}_2, \boldsymbol{\psi}_1 \rangle & \cdots & \langle \boldsymbol{\psi}_N, \boldsymbol{\psi}_1 \rangle \\ \langle \boldsymbol{\psi}_1, \boldsymbol{\psi}_2 \rangle & \langle \boldsymbol{\psi}_2, \boldsymbol{\psi}_2 \rangle & \cdots & \langle \boldsymbol{\psi}_N, \boldsymbol{\psi}_2 \rangle \\ \vdots & & \ddots & \\ \langle \boldsymbol{\psi}_1, \boldsymbol{\psi}_N \rangle & \langle \boldsymbol{\psi}_2, \boldsymbol{\psi}_N \rangle & \cdots & \langle \boldsymbol{\psi}_N, \boldsymbol{\psi}_N \rangle \end{bmatrix}^{-1} \begin{bmatrix} \langle \boldsymbol{x}, \boldsymbol{\psi}_1 \rangle \\ \langle \boldsymbol{x}, \boldsymbol{\psi}_2 \rangle \\ \vdots \\ \langle \boldsymbol{x}, \boldsymbol{\psi}_N \rangle \end{bmatrix}.$$

Use $\boldsymbol{H}$ to denote the inverse Gram matrix above. Then

$$\alpha_n = \sum_{\ell=1}^{N} H_{n,\ell} \langle \boldsymbol{x}, \boldsymbol{\psi}_\ell \rangle = \left\langle \boldsymbol{x}, \sum_{\ell=1}^{N} H_{n,\ell} \boldsymbol{\psi}_\ell \right\rangle.$$

Thus

$$x(t) = \sum_{n=1}^{N} \langle \boldsymbol{x}, \widetilde{\boldsymbol{\psi}}_n \rangle \psi_n(t), \quad \text{where} \quad \widetilde{\psi}_n(t) = \sum_{\ell=1}^{N} H_{n,\ell}\, \psi_\ell(t).$$

Example: Let $\mathcal{S}$ be the space of all second-order polynomials on $[0, 1]$. Set

$$\psi_1(t) = 1, \quad \psi_2(t) = t, \quad \psi_3(t) = t^2.$$

Then

$$\boldsymbol{H} = \begin{bmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{bmatrix}^{-1} = \begin{bmatrix} 9 & -36 & 30 \\ -36 & 192 & -180 \\ 30 & -180 & 180 \end{bmatrix},$$

and

$$\begin{aligned} \widetilde{\psi}_1(t) &= 30t^2 - 36t + 9, \\ \widetilde{\psi}_2(t) &= -180t^2 + 192t - 36, \\ \widetilde{\psi}_3(t) &= 180t^2 - 180t + 30. \end{aligned}$$

Non-orthogonal basis in infinite dimensions: Riesz Bases

When $\mathcal{S}$ is infinite dimensional, we have to proceed with a little more caution. It is possible that we have a infinite set of vectors which are linearly independent and span $\mathcal{S}$ (after closure), but the representation is completely unstable.

Recall that a (possibly infinite) set of vectors is called linearly independent if no finite subset is linearly dependent. Trouble can come when larger and larger sets are coming closer and closer to being linearly dependent. That is, if $\{\boldsymbol{\psi}_n,\ 1 \leq n \leq \infty\}$ is a set of vectors, there might be no $\alpha_2, \ldots, \alpha_L$ such that

$$\boldsymbol{\psi}_1 = \sum_{\ell=2}^{L} \alpha_\ell \boldsymbol{\psi}_\ell,$$

exactly, no matter how large L is. But there could be an infinite sequence $\{\alpha_\ell\}$ such that

$$\lim_{L\to\infty} \|\boldsymbol{\psi}_1 - \sum_{\ell=2}^{L} \alpha_\ell \boldsymbol{\psi}_\ell\| = 0.$$

We will see an example of this below.

Our definition of basis prevents sequences like the above from occurring.

Definition. We say[1] $\{\psi_n\}_{n=1}^{\infty}$ is a **Riesz basis** if there exists constants $A, B > 0$ such that

$$A\sum_{n=1}^{\infty} |\alpha_n|^2 \ \leq\ \left\|\sum_{n=1}^{\infty} \alpha_n \boldsymbol{\psi}_n\right\|^2 \ \leq\ B\sum_{n=1}^{\infty} |\alpha_n|^2$$

[1]This definition uses the natural numbers to index the set of basis functions, but of course it applies equally to any countably infinite sequence.

uniformly for all sequences $\{\alpha_n\}$ with $\sum_{n=1}^{\infty} |\alpha_n|^2 < \infty$.

Note that if $\{\psi_n\}$ is an orthonormal basis, then it is a Riesz basis with $A = B = 1$ (Parseval's theorem).

Our first example is something which is *not* a Riesz basis.

Example: Multiscale Tent Functions

Consider this set of continuous-time signals on $[0, 1]$.

$$\phi_0(t) = \sqrt{2}(1-t), \quad \phi_1(t) = \sqrt{2}t,$$

$$\psi_0(t) = \begin{cases} \sqrt{3}t, & 0 \le t \le 1/2, \\ \sqrt{3}(1-t), & 1/2 \le t \le 1. \end{cases}$$

$$\psi_{j,n}(t) = 2^{j/2}\psi_0(2^j t - n), \quad j \ge 1, \quad n = 0, \dots, 2^j - 1.$$

Sketch:

From your sketch above, it should be clear that

$$\mathrm{Span}(\{\boldsymbol{\phi}_0, \boldsymbol{\phi}_1, \boldsymbol{\psi}_0, \boldsymbol{\psi}_{j,n}, \quad 1 \le j \le J, \ n = 0, \dots, 2^j - 1\})$$

is the set of all continuous piecewise-linear functions on dyadic intervals of length $2^{-(J+1)}$. Since this set is dense in $L_2([0,1])$, we can write

$$x(t) = b_0\phi_0(t) + b_1\phi_1(t) + \sum_{j,n} c_{j,n}\boldsymbol{\psi}_{j,n}(t)$$

for some sequence of numbers $\{b_0, b_1, c_{j,n}\}$. The problem is that this sequence of numbers might not be well-behaved.

To see that this collection of functions cannot be a Riesz basis, notice that using the functions with $1 \leq j \leq J$, we can match the samples of any function on the grid with spacing $2^{-(J+1)}$, with linear interpolation in between these samples. From this, we see that $\phi_0(t)$ can be matched exactly on the interval $[2^{-(J+1)}, 1 - 2^{-(J+1)}]$ with a linear combination of $\boldsymbol{\psi}_{j,n},\ 0 \leq j \leq J$. This means that there is a sequence of numbers $\{\beta_{j,n}\}$ such that

$$\phi_0(t) = \sum_{j\geq 0}\sum_{n=0}^{2^j-1} \beta_{j,n}\psi_{j,n}(t).$$

This means that the non-zero sequence of numbers $\{1, 0, \beta_{j,n},\ j \geq 0,\ n = 0, \ldots, 2^j - 1\}$ synthesizes the $\mathbf{0}$ signal, thus violating the condition that $A > 0$ uniformly.

Example: Non-harmonic sinusoids

Consider the set of signals on $[0, 1]$

$$\psi_k(t) = e^{j2\pi\gamma_k t}, \quad k \in \mathbb{Z}$$

where the frequencies γ_k are a sequence of numbers obeying

$$\gamma_k < \gamma_{k+1}, \quad \gamma_k \to -\infty \text{ as } k \to -\infty, \quad \gamma_k \to +\infty \text{ as } k \to +\infty.$$

Of course, if $\gamma_k = k$, this is the classical Fourier Series basis, and the $\{\boldsymbol{\psi}_k\}$ form an orthobasis. If the γ_k are no longer equally spaced by an integer multiple, these signals are not orthogonal. However, if they are not too far from being uniformly spaced, they still form a Riesz basis. "Kadec's 1/4-Theorem" is a result from harmonic analysis that says: If there exists a $\delta < 1/4$ such that

$$|\gamma_k - k| \leq \delta \quad \text{for all } k,$$

then $\{\boldsymbol{\psi}_k\}$ is a Riesz basis with

$$A = (\cos(\pi\delta) - \sin(\pi\delta))^2, \quad B = (2 - \cos(\pi\delta) + \sin(\pi\delta))^2.$$

Example: Modulated Bumps

In a previous lecture, we saw that signals of the form

$$\psi_{n,k}(t) = g(t-n)\cos((k+1/2)\pi t),\ k \geq 0,\ n \in \mathbb{Z},$$

formed an orthobasis if the windowing function $g(t)$ was chosen carefully (this was called the Lapped Orthogonal Transform). If we are not so concerned with orthogonality, we can use many different kinds of windows. For example, the set

$$\psi_{n,k}(t) = e^{-(t-n-1/2)^2/2}\sin((k+1/2)\pi t), \quad k \geq 1,\ n \in \mathbb{Z},$$

is a Riesz basis for $L_2(\mathbb{R})$. The basis functions are essentially a bell curve centered on the half integers modulated by sinusoids of different frequencies. (Using sin in the expression above instead of cos makes the symmetries work out the way they need to.) In this case, the Riesz constants A, B and the dual basis can be computed explicitly[2].

[2] R. R. Coifman and Y. Meyer, "Gaussian Bases," *Applied and Computational Harmonic Analysis*, vol. 2, pp. 299–302, 1995.

Riesz representation

If $\{\boldsymbol{\psi}_n\}$ is a Riesz basis for a Hilbert space $\mathcal{S}$, then there is a dual basis $\{\widetilde{\boldsymbol{\psi}}_n\}$ such that for all $\boldsymbol{x} \in \mathcal{S}$,

$$\boldsymbol{x} = \sum_{n=1}^{\infty} \langle \boldsymbol{x}, \widetilde{\boldsymbol{\psi}}_n \rangle \, \boldsymbol{\psi}_n.$$

So just as in finite dimensions, we can think of the $\{\boldsymbol{\psi}_n\}$ as providing a *transform*: we calculate the transform coefficients using $\alpha_n = \langle \boldsymbol{x}, \widetilde{\boldsymbol{\psi}}_n \rangle$, and can invert the transform (re-synthesize the signal) using $\boldsymbol{x} = \sum_n \alpha_n \boldsymbol{\psi}_n$.

Linear approximation

Similarly, if $\{\boldsymbol{\psi}_n\}$ is a Riesz basis for an infinite dimensional subspace $\mathcal{T}$ of $\mathcal{S}$, then there exists a dual basis with $\widetilde{\boldsymbol{\psi}}_n \in \mathcal{T}$ such that

$$\hat{\boldsymbol{x}} = \sum_{n=1}^{\infty} \langle \boldsymbol{x}, \widetilde{\boldsymbol{\psi}}_n \rangle \, \boldsymbol{\psi}_n,$$

is the best approximation to $\boldsymbol{x}$ in $\mathcal{T}$. That is, $\hat{\boldsymbol{x}}$ above is the solution to

$$\min_{\boldsymbol{v} \in \mathcal{T}} \|\boldsymbol{x} - \boldsymbol{v}\|$$

Stable reconstruction

Suppose that we compute the basis expansion coefficients for a signal $\boldsymbol{x}$:

$$\alpha_n = \langle \boldsymbol{x}, \widetilde{\boldsymbol{\psi}}_n \rangle,$$

and then perturb them:

$$\hat{\alpha}_n = \alpha_n + \epsilon_n.$$

What effect does this have on the reconstructed signal? Set

$$\hat{\boldsymbol{x}} = \sum_{n=-\infty}^{\infty} \hat{\alpha}_n \boldsymbol{\psi}_n,$$

then

$$\boldsymbol{x} - \hat{\boldsymbol{x}} = \sum_{n=-\infty}^{\infty} \epsilon_n \boldsymbol{\psi}_n,$$

and so if $\{\boldsymbol{\psi}_n\}$ is a Riesz basis with constants A, B,

$$A\|\boldsymbol{\epsilon}\|_2^2 \;\leq\; \|\boldsymbol{x} - \hat{\boldsymbol{x}}\|^2 \;\leq\; B\|\boldsymbol{\epsilon}\|_2^2.$$

That is, the squared-error in signal space is no greater than B times the squared-error in coefficient space (and no less than A times the coefficient squared-error.)

Computing the dual basis

We have not said anything yet about how to compute the dual basis in infinite dimensions. This is because it is much less straightforward than in the finite dimensional case — instead matrix equations, we have to manipulate linear operators that act on sequences of numbers of infinite length.

But still, we can do this in certain cases, as we will see in the next section. Let's draw some parallels to the finite dimensional case to see what needs to be done. For a finite N-dimensional space, we form the $N \times N$ Gram matrix $\boldsymbol{G}$ by filling in the entries $G_{n,\ell} =$

$\langle \boldsymbol{\psi}_\ell, \boldsymbol{\psi}_n \rangle$, invert it to get another $N \times N$ matrix $\boldsymbol{H}$, and then set $\widetilde{\boldsymbol{\psi}}_n = \sum_{\ell=1}^N H_{n,\ell} \boldsymbol{\psi}_\ell$.

We can follow the same procedure in infinite dimensions, but now the Gram "matrix" has an infinite number of rows and columns. The Grammian is a linear operator $\mathscr{G} : \ell_2(\mathbb{Z}) \rightarrow \ell_2(\mathbb{Z})$; it maps infinite length sequences to infinite length sequences. Given an input $\boldsymbol{x} \in \ell_2(\mathbb{Z})$ to this operator, the output at the n^{th} index is given by

$$[\mathscr{G}(\boldsymbol{x})](n) = \sum_{\ell=-\infty}^{\infty} \langle \boldsymbol{\psi}_\ell, \boldsymbol{\psi}_n \rangle x[\ell].$$

It turns out that the conditions for being a Riesz basis ensure that $\mathscr{G}$ is invertible, that is, that there is another linear operator $\mathscr{H}$: $\ell_2(\mathbb{Z}) \rightarrow \ell_2(\mathbb{Z})$ such that

$$\mathscr{H}(\mathscr{G}(\boldsymbol{x})) = \boldsymbol{x}, \quad \text{for all } \ \boldsymbol{x} \in \ell_2(\mathbb{Z}).$$

In general, we need completely different methods to compute the inverse $\mathscr{H} = \mathscr{G}^{-1}$ than we do in the finite dimensional case. But in the end, the action of $\mathscr{H}$ will be specified by a two-dimensional array of numbers $\{H_{n,\ell},\ n, \ell \in \mathbb{Z}\}$; for any $\boldsymbol{v} \in \ell_2(\mathbb{Z})$,

$$[\mathscr{H}(\boldsymbol{v})](n) = \sum_{\ell=-\infty}^{\infty} H_{n,\ell}\ v[\ell].$$