I. Orthobases and Frames

Vector Spaces

In this course our fundamental model for signals will be that they are *vectors* in an appropriate *vector space*. Thus, we will begin by reviewing some of the basic concepts of vector spaces. For an excellent introduction to this material, see *An Introduction to Hilbert Space* by N. Young, especially Chapters 1–4.

We begin by formally defining what we mean by a vector and a vector space.

Definition 1. A *vector space* H is a set of "vectors" together with a field R of "scalars" (usually $R = \mathbb{R}$ or $\mathbb{C}$) such that

- H is closed on addition, i.e., for any $f, g \in H$, $f + g \in H$.
- H is closed on multiplication with a scalar, i.e., for any $f \in H$, $\alpha \in R$, $\alpha f \in H$.

Similarly, a subset $V \subseteq H$ is called a *subspace* if V is closed under vector addition and scalar multiplication.

Examples:

- $H = \mathbb{R}^n$ is a vector space (with scalar field $\mathbb{R}$)
- $H = \mathbb{C}^n$ is a vector space (with either $\mathbb{R}$ or $\mathbb{C}$ as a scalar field)
- $V = \mathbb{R}$ is a subspace of $H = \mathbb{R}^2$.
- The set of all finite-energy signals is a vector space.
- The set of bandlimited signals forms a subspace of the set of all finite-energy signals.

We will mostly be interested in vector spaces that have two additional types of structure: a *norm* and an *inner product*.

Norms

Norms are a generalization of the geometric concept of *length*.

Definition 2. A *norm* $\|\cdot\| : H \to \mathbb{R}$ on H satisfies

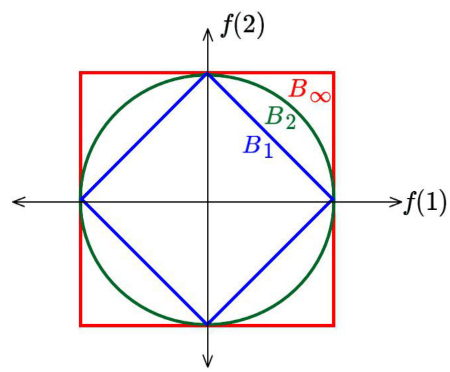
Positive-definiteness: For any $f \in H$, $\|f\| \geq 0$ and $\|f\| = 0$ if and only if $f = 0$.

Homogeneity: For any $\alpha \in \mathbb{C}$, $\|\alpha f\| = |\alpha| \, \|f\|$.

Triangle Inequality: For any $f, g \in H$, $\|f + g\| \leq \|f\| + \|g\|$.

Examples:

- For $H = \mathbb{R}^N$ (or $\mathbb{C}^N$) and for $p < \infty$, we define
$$\|f\|_{\ell_p} = \left(\sum_{k=1}^{n} |f[k]|^p \right)^{1/p}.$$
For $1 \leq p < \infty$, this satisfies the requirements of a norm. For $p = \infty$ we can extend this with
$$\|f\|_{\ell_\infty} = \max_{1 \leq k \leq n} |f[k]|,$$
which is also a valid norm. In $\mathbb{R}^2$, if we set $B_p = \{f : \|f\|_{\ell_p} = 1\}$, then we have the following picture Later in the course we will occasionally also



consider this definition where $p < 1$, but such "norms" do not actually satisfy all of the requirements to be a norm. Note that it is also common to use the notation $\|\cdot\|_p$ for $\|\cdot\|_{\ell_p}$.

- For a sequence $\{f[k]\}_{k \in \mathbb{Z}}$ (a sequence indexed by the integers), we can extend the notion of the ℓ_p-norms and define
$$\|f\|_{\ell_p(\mathbb{Z})} = \left(\sum_{k \in \mathbb{Z}} |f[k]|^p \right)^{1/p}$$
for $1 \leq p < \infty$ and
$$\|f\|_{\ell_\infty(\mathbb{Z})} = \sup_{k \in \mathbb{Z}} |f[k]|.$$
We will also often use the notation
$$\ell_p(\mathbb{Z}) = \left\{ f : \|f\|_{\ell_p(\mathbb{Z})} < \infty \right\}$$
to denote the set of finite-norm signals.

- Analogous definitions hold for continuous-time signals. Suppose that $f(t)$ is supported on $T \subseteq \mathbb{R}$ (e.g, $T = \mathbb{R}$ or $T = [0,1]$), then we can define

$$\|f\|_{L_p(T)} = \left(\int_T |f(t)|^p \, dt \right)^{1/p}$$

and

$$\|f\|_{L_\infty(T)} = \sup_{t \in T} |f(t)| .$$

As before, we will also often use the notation

$$L_p(T) = \left\{ f : \|f\|_{L_p(T)} < \infty \right\}$$

to denote the set of signals supported on T with finite-norm.

- Let $f(t)$ be a continuous-time signal on $[0,1]$ with Fourier series coefficients given by $\{c_k\}_{k \in \mathbb{Z}}$. Then

$$\|f\| = \sum_{k \in \mathbb{Z}} |c_k|$$

is a valid norm.

- Let H be the space of $m \times n$ matrices. Then

$$\|A\| = \sigma_{\max}(A),$$

where $\sigma_{\max}(A)$ denotes the maximum singular value of A, is a norm in H. This is known as the *operator norm*.

- Let H again be the space of $m \times n$ matrices. Then

$$\|A\|_F = \left(\sum_{i,j} |A_{ij}|^2 \right)^{1/2} = \left(\sum_k \sigma_k^2 \right)^{1/2}$$

is a valid norm. This is known as the *Frobenius norm* and is equivalent to treating the matrix A as an $mn \times 1$ vector and computing its ℓ_2 norm or taking the ℓ_2 norm of the vector of singular values of A.

Inner Products

In addition to norms, it is also useful to introduce the additional structure of an inner product, which will help us generalize the geometric concept of *angle*.

Definition 3. An *inner product* $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{C}$ on H satisfies

Conjugate Symmetry: For any $f, g \in H$, $\langle f, g \rangle = \langle g, f \rangle^*$.

Linearity: For any $\alpha \in \mathbb{C}$, $\langle \alpha f, g \rangle = \alpha \langle f, g \rangle$ and for any $f, g, h \in H$, $\langle f + g, h \rangle = \langle f, h \rangle + \langle g, h \rangle$.

Positive-definiteness: For any $f \in H$, $\langle f, f \rangle \geq 0$, and $\langle f, f \rangle = 0$ if and only if $f = 0$.

It is easy to verify that an inner product induces a norm on H via

$$\|f\|_H = \sqrt{\langle f, f \rangle_H}.$$

We will refer to such a norm as an *induced norm* to distinguish it from other possible norms that might be defined on H (e.g., in $\mathbb{R}^n$, we can define many norms, but only some of these norms are induced by valid inner products). By far, we will be spending most of our time in inner product spaces such as $L_2(\mathbb{R})$, $L_2([a,b])$, $\ell_2(\mathbb{Z})$, or $\mathbb{C}^n$, each equipped with the "standard" inner product given by

$$\langle f, g \rangle_{L_2(\mathbb{R})} = \int_{-\infty}^{\infty} f(t) g(t)^* dt$$

and

$$\langle f, g \rangle_{\ell_2(\mathbb{Z})} = \sum_{k=-\infty}^{\infty} f[k] g[k]^*,$$

with analogous definitions for $\langle \cdot, \cdot \rangle_{L_2[a,b]}$ and $\langle \cdot, \cdot \rangle_{\mathbb{C}^n}$. Other examples of inner products include:

Examples:

- Consider the space of real-valued, finite-variance, zero-mean random variables. Then $\langle X, Y \rangle = \mathbb{E}[XY]$ is a valid inner product and the induced norm is $\|X\| = \text{StdDev}(X)$.

- Consider the space of $m \times n$ matrices. Then

$$\langle A, B \rangle_{\text{tr}} = \text{trace}\,(A^* B) = \sum_{i,j} A_{ij} B_{ij}^*$$

  is a valid inner product and induces the Frobenius norm.

An inner product allows us to define the *angle* between two signals. To see why this is the case, recall the incredibly important *Cauchy-Schwarz inequality*:

$$|\langle f, g \rangle_H| \leq \|f\|_H \|g\|_H .$$

(Try proving this at home or check out the proof on Wikipedia.) Using this fact we can see that for a real vector space, we can define an angle θ between two (non-zero) signals in H via

$$\cos\theta = \frac{\langle f, g \rangle_H}{\|f\|_H \|g\|_H}.$$

From Cauchy-Schwarz we are guaranteed that the right-hand side of this equation is always in the range of $[-1, 1]$, so that the definition makes sense. For a complex vector space this definition is somewhat problematic since finding θ would require computing the inverse cosine of a complex number. Thus, the typical convention in this case is to essentially ignore the "direction" of f and g and define the angle via

$$\cos\theta = \frac{|\langle f, g\rangle_H|}{\|f\|_H \|g\|_H},$$

which computes the angle between the two subspaces spanned by f and g.

Another way of viewing the additional structure induced by the existence of an inner product is through the *parallelogram law.*

Theorem 1 (Parallelogram Law)**.** *For any inner product space H, for any $f, g \in H$ we have*

$$\|f+g\|_H^2 + \|f-g\|_H^2 = 2\|f\|_H^2 + 2\|g\|_H^2$$

Proof. We can write

$$\begin{aligned}
\|f+g\|_H^2 + \|f-g\|_H^2 &= \langle f+g, f+g\rangle_H + \langle f-g, f-g\rangle_H \\
&= \langle f, f\rangle_H + \langle f, g\rangle_H + \langle g, f\rangle_H + \langle g, g\rangle_H \\
&\quad + \langle f, f\rangle_H - \langle f, g\rangle_H - \langle g, f\rangle_H + \langle g, g\rangle_H \\
&= 2\langle f, f\rangle_H + 2\langle g, g\rangle_H \\
&= 2\|f\|_H^2 + 2\|g\|_H^2,
\end{aligned}$$

as desired. □

One can also show that given any normed vector space with a norm that satisfies the parallelogram law, once can define a valid inner product using that norm. Thus, in a sense, the parallelogram law captures all of the additional structure imposed on a vector space by an inner product.

Hilbert Spaces

One aspect of dealing with infinite-dimensional vector spaces that we are mostly glossing over is that of *completeness.* To state what we mean by completeness, we consider the following notions of convergence.

Definition 4. Let $\{f_k\}_{k=1}^{\infty}$ be a sequence of vectors in H. We say that $\{f_k\}_{k=1}^{\infty}$ *converges* to f^* if and only if for every $\varepsilon > 0$ there is a k_0 such that $\|f_k - f^*\|_H < \varepsilon$ for all $k > k_0$. In this case we say that $\{f_k\}_{k=1}^{\infty}$ is a *convergent sequence* with f^* as its *limit.*

Definition 5. A sequence $\{f_k\}_{k=1}^{\infty}$ is said to be a *Cauchy sequence* if for any $\varepsilon > 0$ there is a k_0 such that $\|f_k - f_j\|_H < \varepsilon$ for every $k, j > k_0$.

A normed vector space or an inner product space are said to be *complete* if every Cauchy sequence is also a convergent sequence. Informally, complete spaces are ones where sequences that look like they are converging actually do converge to something in the space. A complete normed vector space is also called a *Banach space*, and a complete inner product space is also called a *Hilbert space*. All of the vector spaces we will deal with in this course are complete (e.g., $L_p(\mathbb{R})$, $\ell_p(\mathbb{Z})$, or $\mathbb{C}^n$), so we will not focus on this technicality too much. However, the importance of completeness to signal processing applications is hard to overstate. As a simple example, consider the following consequence of being in a Hilbert space.

Definition 6. A subset A of a vector space H is *convex* if for all $f, g \in A$ and $\lambda \in (0, 1)$, $\lambda f + (1 - \lambda)g \in A$.

Theorem 2 ("The Fundamental Theorem of Approximation")**.** *Let A be a nonempty, closed (complete), convex set in a Hilbert space H. For any $f \in H$ there is a unique point in A that is closest to f, i.e., f has a unique "best approximation" in A.*

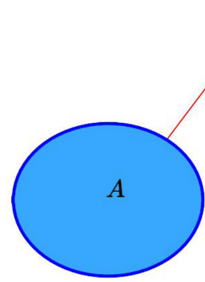


Figure 1: The best approximation to f in convex set A.

Note that in non-Hilbert spaces, this may not be true! The proof is rather technical. See Chapter 3 of Young for details. Also known as the "closest point property", this is very useful in compression and denoising.