Signal Space Concepts — Why we Care

- **Signal Space Concepts** are a powerful tool for the analysis of communication systems and for the design of optimum receivers.
- **Key Concepts:**
  - Orthonormal basis functions — tailored to signals of interest — span the signal space.
  - *Representation theorem:* allows any signal to be represented as a (usually finite dimensional) vector
    - Signals are interpreted as points in signal space.
  - For random processes, representation theorem leads to random signals being described by random vectors with uncorrelated components.
    - *Theorem of Irrelavance* allows us to disregrad nearly all components of noise in the receiver.
- We will briefly review key ideas that provide underpinning for signal spaces.

Linear Vector Spaces

- The basic structure needed by our signal spaces is the idea of linear vector space.
- **Definition:** A linear vector space $\mathcal{S}$ is a collection of elements ("vectors") with the following properties:
  - Addition of vectors is defined and satisfies the following conditions for any $x, y, z \in \mathcal{S}$:
    1. $x + y \in \mathcal{S}$ (closed under addition)
    2. $x + y = y + x$ (commutative)
    3. $(x + y) + z = x + (y + z)$ (associative)
    4. The zero vector $\vec{0}$ exists and $\vec{0} \in \mathcal{S}$. $x + \vec{0} = x$ for all $x \in \mathcal{S}$.
    5. For each $x \in \mathcal{S}$, a unique vector $(-x)$ is also in $\mathcal{S}$ and $x + (-x) = \vec{0}$.

Linear Vector Spaces — continued

- **Definition — continued:**
  - Associated with the set of vectors in $\mathcal{S}$ is a set of scalars. If a, b are scalars, then for any $x, y \in \mathcal{S}$ the following properties hold:
    1. $a \cdot x$ is defined and $a \cdot x \in \mathcal{S}$.
    2. $a \cdot (b \cdot x) = (a \cdot b) \cdot x$
    3. Let 1 and 0 denote the multiplicative and additive identies of the field of scalars, then $1 \cdot x = x$ and $0 \cdot x = \vec{0}$ for all $x \in \mathcal{S}$.
    4. Associative properties:

$$a \cdot (x + y) = a \cdot x + a \cdot y$$
$$(a + b) \cdot x = a \cdot x + b \cdot x$$

Running Examples

- The space of length-N vectors $\mathbb{R}^N$

$$\begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_N + y_N \end{pmatrix} \text{ and } a \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} = \begin{pmatrix} a \cdot x_1 \\ \vdots \\ a \cdot x_N \end{pmatrix}$$

- The collection of all square-integrable signals over $[T_a, T_b]$, i.e., all signals $x(t)$ satisfying

$$\int_{T_a}^{T_b} |x(t)|^2 \, dt < \infty.$$

  - Verifying that this is a linear vector space is easy.
  - This space is called $L^2(T_a, T_b)$ (pronounced: ell-two).

Inner Product

- To be truly useful, we need linear vector spaces to provide
  - means to measure the length of vectors and
  - to measure the distance between vectors.
- Both of these can be achieved with the help of inner products.
- **Definition:** The inner product of two vectors $x, y, \in \mathcal{S}$ is denoted by $\langle x, y \rangle$. The inner product is a *scalar* assigned to x and y so that the following conditions are satisfied:
  1. $\langle x, y \rangle = \langle y, x \rangle$ (for complex vectors $\langle x, y \rangle = \langle y, x \rangle^*$)
  2. $\langle a \cdot x, y \rangle = a \cdot \langle x, y \rangle$, with scalar a
  3. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$, with vector z
  4. $\langle x, x \rangle > 0$, except when $x = \vec{0}$; then, $\langle x, x \rangle = 0$.

Exercise: Valid Inner Products?

- $x, y \in \mathbb{R}^N$ with

$$\langle x, y \rangle = \sum_{n=1}^{N} x_n y_n$$

- **Answer:** Yes; this is the standard *dot product*.
- $x, y \in \mathbb{R}^N$ with

$$\langle x, y \rangle = \sum_{n=1}^{N} x_n \cdot \sum_{n=1}^{N} y_n$$

- **Answer:** No; last condition does not hold, which makes this inner product useless for measuring distances.

Exercise: Valid Inner Products?

- $x(t), y(t) \in L^2(a, b)$ with

$$\langle x(t), y(t) \rangle = \int_a^b x(t) y(t) \, dt$$

- **Answer:** Yes; continuous-time equivalent of the dot-product.
- $x, y \in \mathbb{C}^N$ with

$$\langle x, y \rangle = \sum_{n=1}^{N} x_n y_n^*$$

- **Answer:** Yes; the conjugate complex is critical to meet the last condition (e.g., $\langle j, j \rangle = -1 < 0$).
- $x, y \in \mathbb{R}^N$ with

$$\langle x, y \rangle = x^T K y = \sum_{n=1}^{N} \sum_{m=1}^{N} x_n K_{n,m} y_m$$

Exercise: Valid Inner Products?

- $x, y \in \mathbb{R}^N$ with

$$\langle x, y \rangle = x^T K y = \sum_{n=1}^{N} \sum_{m=1}^{N} x_n K_{n,m} y_m$$

  with K an $N \times N$-matrix
- **Answer:** Only if K is positive definite (i.e., $x^T K x > 0$ for all $x \neq \vec{0}$).

Norm of a Vector

- **Definition:** The norm of vector $x \in \mathcal{S}$ is denoted by $\|x\|$ and is defined via the inner product as

$$\|x\| = \sqrt{\langle x, x \rangle}.$$

  - Notice that $\|x\| > 0$ unless $x = \vec{0}$, then $\|x\| = 0$.
  - The norm of a vector measures the length of a vector.
  - For signals $\|x(t)\|^2$ measures the *energy* of the signal.
- **Example:** For $x \in \mathbb{R}^N$, Cartesian length of a vector

$$\|x\| = \sqrt{\sum_{n=1}^{N} |x_n|^2}$$

Norm of a Vector — continued

- **Illustration:**

$$\|a \cdot x\| = \sqrt{\langle a \cdot x, a \cdot x \rangle} = |a| \|x\|$$

  - Scaling the vector by a, scales its length by a.

Inner Product Space

- We call a linear vector space with an associated, valid inner product an inner product space.
  - **Definition:** An inner product space is a linear vector space in which a inner product is defined for all elements of the space and the norm is given by $\|x\| = \langle x, x \rangle$.
- **Standard Examples:**
  1. $\mathbb{R}^N$ with $\langle x, y \rangle = \sum_{n=1}^{N} x_n y_n$.
  2. $L^2(a, b)$ with $\langle x(t), y(t) \rangle = \int_a^b x(t)y(t)\,dt$.

Schwartz Inequality

- The following relationship between norms and inner products holds for all inner product spaces.
- **Schwartz Inequality:** For any $x, y \in \mathcal{S}$, where $\mathcal{S}$ is an inner product space,

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\|$$

with equality if and only if $x = c \cdot y$ with scalar c
  - Proof follows from $\|x + a \cdot y\|^2 \geq 0$ with $a = -\frac{\langle x, y \rangle}{\|y\|^2}$.

Orthogonality

- **Definition:** Two vectors are orthogonal if the inner product of the vectors is zero, i.e.,

$$\langle x, y \rangle = 0.$$

- **Example:** The standard basis vectors e_m in $\mathbb{R}^N$ are orthogonal; recall

$$e_m = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \quad \text{the 1 occurs on the } m\text{-th row}$$

Orthogonality

- **Example:** The basis functions for the Fourier Series expansion $w_m(t) \in L^2(0, T)$ are orthogonal; recall

$$w_m(t) = \frac{1}{\sqrt{T}} e^{j2\pi mt/T}.$$

Distance between Vectors

- **Definition:** The distance d between two vectors is defined as the norm of their difference, i.e.,

$$d(x, y) = \|x - y\|$$

- **Example:** The Cartesian (or Euclidean) distance between vectors in $\mathbb{R}^N$:

$$d(x, y) = \|x - y\| = \sqrt{\sum_{n=1}^{N} |x_n - y_n|^2}.$$

- **Example:** The root-mean-squared error (RMSE) between two signals in $L^2(a, b)$ is

$$d(x(t), y(t)) = \|x(t) - y(t)\| = \sqrt{\int_a^b |x(t) - y(t)|^2 \, dt}$$

Properties of Distances

- Distance measures defined by the norm of the difference between vectors x, y have the following properties:
  1. $d(x, y) = d(y, x)$
  2. $d(x, y) = 0$ if and only if $x = y$
  3. $d(x, y) \leq d(x, z) + d(y, z)$ for all vectors z (Triangle inequality)

Exercise: Prove the Triangle Inequality

- Begin like this:

$$\begin{aligned} d^2(x,y) &= \|x-y\|^2 \\ &= \|(x-z)+(z-y)\|^2 \\ &= \langle (x-z)+(z-y), (x-z)+(z-y) \rangle \end{aligned}$$

-

$$\begin{aligned} d^2(x,y) &= \langle x-z, x-z \rangle + 2\langle x-z, z-y \rangle + \langle z-y, z-y \rangle \\ &\leq \langle x-z, x-z \rangle + 2|\langle x-z, z-y \rangle| + \langle z-y, z-y \rangle \\ (Schwartz): &\leq \langle x-z, x-z \rangle + 2\|x-z\| \cdot \|z-y\| + \langle z-y, z- \\ &= d(x,z)^2 + 2d(x,z) \cdot d(y,z) + d(y,z)^2 \\ &= (d(x,z)+d(y,z))^2 \end{aligned}$$

GEORGE MASON UNIVERSITY

Hilbert Spaces — Why we Care

- We would like our vector spaces to have one more property.
  - We say the sequence of vectors $\{x_n\}$ converges to vector x, if
$$\lim_{n\to\infty} \|x_n - x\| = 0.$$
  - We would like the limit point x of any sequence $\{x_n\}$ to be in our vector space.
  - Integrals and derivatives are fundamentally limits; we want derivatives and integrals to stay in the vector space.
  - A vector space is said to be closed if it contains all of its limit points.
- **Definition:** A closed, inner product space is A Hilbert Space.

Hilbert Spaces — Examples

- **Examples:** Both $\mathbb{R}^N$ and $L^2(a, b)$ are Hilbert Spaces.
- **Counter Example:** The space of rational number $\mathbb{Q}$ is **not** closed (i.e., not a Hilbert space)
  - E.g.,

$$\sum_{n=0}^{\infty} \frac{1}{n!} = e \notin \mathbb{Q},$$

even though all $\frac{1}{n!} \in \mathbb{Q}$.

Subspaces

- **Definition:** Let $\mathcal{S}$ be a linear vector space. The space $\mathcal{L}$ is a subspace of $\mathcal{S}$ if
  1. $\mathcal{L}$ is a *subset* of $\mathcal{S}$ and
  2. $\mathcal{L}$ is *closed*.

  - If $x, y \in \mathcal{L}$ then also $x, y, \in \mathcal{S}$.
  - And, $a \cdot x + b \cdot y \in \mathcal{L}$ for all scalars a, b.
- **Example:** Let $\mathcal{S}$ be $L^2(T_a, T_b)$. Define $\mathcal{L}$ as the set of all sinusoids of frequency f_0, i.e., signals of the form $x(t) = A\cos(2\pi f_0 t + \phi)$, with $0 \leq A < \infty$ and $0 \leq \phi < 2\pi$
  1. All such sinusoids are square integrable.
  2. Linear combination of two sinusoids of frequency f_0 is a sinusoid of the same frequency.

Projection Theorem

- **Definition:** Let $\mathcal{L}$ be a subspace of the Hilbert Space $\mathcal{H}$. The vector $x \in \mathcal{H}$ (and $x \notin \mathcal{L}$) is orthogonal to the subspace $\mathcal{L}$ if $\langle x, y \rangle = 0$ for every $y \in \mathcal{L}$.
- **Projection Theorem:** Let $\mathcal{H}$ be a Hilbert Space and $\mathcal{L}$ is a subspace of $\mathcal{H}$.
  Every vector $x \in \mathcal{H}$ has a unique decomposition

$$x = y + z$$

  with $y \in \mathcal{L}$ and z orthogonal to $\mathcal{L}$.
  Furthermore,

$$\|z\| = \|x - y\| = \min_{\nu \in \mathcal{L}} \|x - \nu\|.$$

  - y is called the projection of x onto $\mathcal{L}$.
  - Distance from x to all elements of $\mathcal{L}$ is minimized by y.

Exercise: Fourier Series

- Let $x(t)$ be a signal in the Hilbert space $L^2(0, T)$.
- Define the subspace $\mathcal{L}$ of signals $\nu_n(t) = A_n \cos(2\pi nt/T)$ for a fixed n and T.
- Find the signal $y(t) \in \mathcal{L}$ that minimizes

$$\min_{y(t)\in\mathcal{L}} \|x(t) - y(t)\|^2.$$

- **Answer:** $y(t)$ is the sinusoid with amplitude

$$A_n = \frac{2}{T}\int_0^T x(t)\cos(2\pi nt/T)\,dt = \frac{2}{T}\langle x(t), \cos(2\pi nt/T)\rangle.$$

  - Note that this is (part of the trigonometric form of) the Fourier Series expansion.
  - Note that the inner product involves the projection of $x(t)$ onto an element of $\mathcal{L}$.

Projection Theorem

- The Projection Theorem is most useful when the subspace $\mathcal{L}$ has certain structural properties.
- In particular, we will be interested in the case when $\mathcal{L}$ is spanned by a set of orthonormal vectors.
  - Let's define what that means.

Separable Vector Spaces

- **Definition:** A Hilbert space $\mathcal{H}$ is said to be separable if there exists a set of vectors $\{\Phi_n\}$, $n = 1, 2, \ldots$ that are elements of $\mathcal{H}$ and such that every element $x \in \mathcal{H}$ can be expressed as

$$x = \sum_{n=1}^{\infty} X_n \Phi_n.$$

  - The coefficients X_n are scalars associated with vectors Φ_n.
  - *Equality* is taken to mean

$$\lim_{n \to \infty} \left\| x - \sum_{n=1}^{\infty} X_n \Phi_n \right\|^2 = 0.$$

Representation of a Vector

- The set of vectors $\{\Phi_n\}$ is said to be complete if the above is valid for every $x \in \mathcal{H}$.
- A complete set of vectors $\{\Phi_n\}$ is said to form a basis for $\mathcal{H}$.
- **Definition:** The representation of the vector x (with respect to the basis $\{\Phi_n\}$) is the sequence of coefficients $\{X_n\}$.
- **Definition:** The number of vectors Φ_n that is required to express every element x of a separable vector space is called the dimension of the space.

Example: Length-N column Vectors

- The space $\mathbb{R}^N$ is separable and has dimension N.
  - Basis vectors ($m = 1, \ldots, N$):

$$\Phi_m = e_m = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \quad \text{the 1 occurs on the } m\text{-th row}$$

  - There are N basis vectors; dimension is N.

Example: Length-N column Vectors — continued

- (con't)
  - For any vector $x \in \mathbb{R}^N$:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} = \sum_{m=1}^{N} x_m e_m$$

Examples: L^2

- **Fourier Bases:** The following is a complete basis for $L^2(0, T)$

$$\begin{aligned} \Phi_{2n}(t) &= \sqrt{\frac{2}{T}} \cos(2\pi n t / T) \\ \Phi_{2n+1}(t) &= \sqrt{\frac{2}{T}} \sin(2\pi n t / T) \end{aligned} \quad n = 0, 1, 2, \ldots$$

  - This implies that $L^2(0, T)$ is a separable vector space.
  - $L^2(0, T)$ is infinite-dimensional.

Examples: L^2

- **Piecewise Linear Signals:** The set of vectors (signals)

$$\Phi_n(t) = \begin{cases} \frac{1}{\sqrt{T}} & (n-1)T \leq t < nT \\ 0 & \text{else} \end{cases}$$

is **not** a basis for $L^2(0, \infty)$.

  - Only piecewise constant signals can be represented.
  - But, this is a basis for the subspace of L^2 consisting of piecewise constant signals.

Orthonormal Bases

- **Definition:** A basis for a separable vector space is an orthonormal basis if the elements of the vectors that constitute the basis satisfy
  1. $\langle \Phi_n, \Phi_m \rangle = 0$ for all $n \neq m$. (*ortho*gonal)
  2. $\|\Phi_n\| = 1$, for all $n = 1, 2, \ldots$ (*normal*ized)
- Note:
  - Not every basis is orthonormal.
    - We will see shortly, every basis can be turned into an orthonormal basis.
  - Not every set of orthonornal vectors constitutes a basis.
    - Example: Piecewise Linear Signals.

Representation with Orthonormal Basis

- An orthonormal basis is much prefered over an arbitrary basis because the representation of vector x is very easy to compute.
- The representation $\{X_n\}$ of a vector x

$$x = \sum_{n=1}^{\infty} X_n \Phi_n$$

with respect to an orthonormal basis $\{\Phi_n\}$ is computed using

$$X_n = \langle x, \Phi_n \rangle.$$

The representation X_n is obtained by projecting x onto the basis vector Φ_n!

  - In contrast, when bases are not orthonormal, finding the representation of x requires solving a system of linear equations.