E9 211: Adaptive Signal Processing

Lecture 2: Linear Algebra

Outline

1. Definitions
2. Subspaces
3. Projection

Vectors

- A N-dimensional vector is assumed to be a column vector:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}$$

- Complex conjugate (Hermitian) transpose

$$\mathbf{x}^{\mathrm{H}} = (\mathbf{x}^{\mathrm{T}})^* = [x_1^*, x_2^*, \ldots, x_N^*]$$

- For a discrete-time signal x(n), we will use the following vectors

$$\mathbf{x} = \begin{bmatrix} x(0) \\ x(1) \\ \vdots \\ x(N-1) \end{bmatrix} \qquad \mathbf{x}_n = \begin{bmatrix} x(n) \\ x(n-1) \\ \vdots \\ x(n-N+1) \end{bmatrix}$$

Matrices

- An $N \times M$ matrix has N rows and M columns:

$$\mathbf{A} = [a_{ij}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1M} \\ a_{21} & a_{22} & \cdots & a_{2M} \\ \vdots & \vdots & & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NM} \end{bmatrix}$$

- Complex conjugate (Hermitian) transpose

$$\mathbf{A}^{\mathrm{H}} = (\mathbf{A}^{\mathrm{T}})^* = (\mathbf{A}^*)^{\mathrm{T}}$$

- Hermitian matrix

$$\mathbf{A} = \mathbf{A}^{\mathrm{H}}$$

E.g.,

$$\mathbf{A} = \begin{bmatrix} 1 & 1+j \\ 1-j & 1 \end{bmatrix}, \quad \text{then} \quad \mathbf{A}^{\mathrm{H}} = \begin{bmatrix} 1 & 1+j \\ 1-j & 1 \end{bmatrix} = \mathbf{A}$$

Vectors

Vector norms: $\|\mathbf{x}\|_p = \left(\sum_{i=1}^{N} |x_i|^p\right)^{1/p}$, for $p = 1, 2, \ldots$.

- $\|\mathbf{x}\|_p \geq 0$ when $\mathbf{x} \neq \mathbf{0}$ and $\|\mathbf{x}\|_p = \mathbf{0}$ *iff* $\mathbf{x} = \mathbf{0}$
- $\|\alpha\mathbf{x}\|_p = \alpha\|\mathbf{x}\|_p$ for any scalar α
- $\|\mathbf{x} + \mathbf{y}\|_p \leq \|\mathbf{x}\|_p + \|\mathbf{y}\|_p$

Examples:

Euclidean (2-norm): $\|\mathbf{x}\|_2 = \left(\sum_{i=1}^{N} |x_i^* x_i\right)^{1/2} = (\mathbf{x}^{\mathrm{H}}\mathbf{x})^{1/2}$

1-norm: $\|\mathbf{x}\|_1 = \sum_{i=1}^{N} |x_i|$

∞-norm: $\|\mathbf{x}\|_\infty = \max_i |x_i|$

Inner product:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^H \mathbf{y} = \sum_{i=1}^{N} x_i^* y_i$$

- Two vectors are *orthogonal* if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$; if the vectors have unit norm, then they are *orthonormal*
- Cauchy-Schwarz: $|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \|\mathbf{y}\|$

  $$\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}\| \|\mathbf{y}\| \cos \theta$$

  θ is the angle between the two vectors. Since $|\cos \theta| \leq 1$, the above inequality follows.
- $2|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$ as

  $$\|\mathbf{x} \pm \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 \pm 2\langle \mathbf{x}, \mathbf{y} \rangle \geq 0$$

Matrices

For $\mathbf{A} \in \mathbb{C}^{M \times N}$

- **2-norm** (spectral norm, operator norm):

$$\|\mathbf{A}\| := \max_{\mathbf{x}} \frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|} \quad \text{or} \quad \|\mathbf{A}\|^2 := \max_{\mathbf{x}} \frac{\mathbf{x}^{\mathrm{H}}\mathbf{A}^{\mathrm{H}}\mathbf{A}\mathbf{x}}{\mathbf{x}^{\mathrm{H}}\mathbf{x}}$$

  Largest magnification that can be obtained by applying $\mathbf{A}$ to any vector

- **Forbenius norm**

$$\|\mathbf{A}\|_{\mathrm{F}} := \left(\sum_{i=1}^{M} \sum_{j=1}^{N} |a_{ij}|^2 \right)^{1/2}$$

  Represents energies in its entries

Linear independence, vector spaces, and basis vectors

Linear independence

- A collection of N vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$ is called *linearly independent* if

$$\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \cdots + \alpha_N \mathbf{x}_N = 0 \quad \Leftrightarrow \quad \alpha_1 = \alpha_2 = \cdots = \alpha_N = 0$$

Rank

- The rank of $\mathbf{A}$ is the number of independent columns or rows of $\mathbf{A}$

Prototype rank-1 matrix: $\mathbf{A} = \mathbf{a}\mathbf{b}^{\mathrm{H}}$
Prototype rank-2 matrix: $\mathbf{A} = \mathbf{a}\mathbf{b}^{\mathrm{H}} + \mathbf{c}\mathbf{d}^{\mathrm{H}}$

- The ranks of $\mathbf{A}, \mathbf{A}\mathbf{A}^{\mathrm{H}}$, and $\mathbf{A}^{\mathrm{H}}\mathbf{A}$ are the same

- If $\mathbf{A}$ is square and full rank, there is a unique inverse $\mathbf{A}^{-1}$ such that

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

- An $N \times N$ matrix $\mathbf{A}$ has rank N, then $\mathbf{A}$ is invertible $\Leftrightarrow \det(\mathbf{A}) \neq 0$

Subspaces

- The space $\mathcal{H}$ spanned by a collection of vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$

$$\mathcal{H} := \{\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \cdots + \alpha_N \mathbf{x}_N | \alpha_i \in \mathbb{C}, \forall i\}$$

  is called a *linear subspace*

  - If the vectors are linearly independent they are called a *basis* for the subspace
  - The number of basis vectors is called the *dimension* of the subspace
  - If the vectors are orthogonal, then we have an *orthogonal basis*
  - If the vectors are orthonormal, then we have an *orthonormal basis*

Fundamental subspaces of $\mathbf{A}$

- Range (column span) of $\mathbf{A} \in \mathbb{C}^{M \times N}$

$$\mathrm{ran}(\mathbf{A}) = \{\mathbf{A}\mathbf{x} : \mathbf{x} \in \mathbb{C}^N\} \subset \mathbb{C}^M$$

The dimension of $\mathrm{ran}(\mathbf{A})$ is rank of $\mathbf{A}$, denoted by $\rho(\mathbf{A})$

- Kernel (row null space) of $\mathbf{A} \in \mathbb{C}^{M \times N}$

$$\ker(\mathbf{A}) = \{\mathbf{x} \in \mathbb{C}^N : \mathbf{A}\mathbf{x} = \mathbf{0}\} \subset \mathbb{C}^N$$

The dimension of $\ker(\mathbf{A})$ is $N - \rho(\mathbf{A})$

- Four fundamental subspaces

$$\mathrm{ran}(\mathbf{A}) \oplus \ker(\mathbf{A}^{\mathrm{H}}) = \mathbb{C}^M$$

$$\mathrm{ran}(\mathbf{A}^{\mathrm{H}}) \oplus \ker(\mathbf{A}) = \mathbb{C}^N$$

direct sum: $\mathcal{H}_1 \oplus \mathcal{H}_2 = \{\mathbf{x}_1 + \mathbf{x}_2 | \mathbf{x}_1 \in \mathcal{H}_1, \mathbf{x}_2 \in \mathcal{H}_2\}$

Unitary and Isometry

- A square matrix $\mathbf{U}$ is called *unitary* if $\mathbf{U}^{\mathrm{H}}\mathbf{U} = \mathbf{I}$ and $\mathbf{U}\mathbf{U}^{\mathrm{H}} = \mathbf{I}$
  - Examples are rotation or reflection matrices
  - $\|\mathbf{U}\| = 1$; its rows and columns are orthonormal
- A tall rectangular matrix $\hat{\mathbf{U}}$ is called an *isometry* if $\hat{\mathbf{U}}^{\mathrm{H}}\mathbf{U} = \mathbf{I}$
  - Its columns are orthonormal basis of a subspace (not the complete space)
  - $\|\hat{\mathbf{U}}\| = 1$;
  - There is an orthogonal complement $\hat{\mathbf{U}}^{\perp}$ of $\hat{\mathbf{U}}$ such that $[\hat{\mathbf{U}} \quad \hat{\mathbf{U}}^{\perp}]$ is unitary

Projection

- A square matrix $\mathbf{P}$ is a *projection* if $\mathbf{P}\mathbf{P} = \mathbf{P}$
- It is an orthogonal projection if $\mathbf{P}^{\mathrm{H}} = \mathbf{P}$
  - The norm of an orthogonal projection is $\|\mathbf{P}\| = 1$
  - For an isometry $\hat{\mathbf{U}}$, the matrix $\mathbf{P} = \hat{\mathbf{U}}\hat{\mathbf{U}}^{\mathrm{H}}$ is an orthogonal projection onto the space spanned by the columns of $\hat{\mathbf{U}}$.
- Suppose $\mathbf{U} = [\underbrace{\hat{\mathbf{U}}}_{d} \quad \underbrace{\hat{\mathbf{U}}^{\perp}}_{N-d}]$ is unitary. Then, from $\mathbf{U}\mathbf{U}^{\mathrm{H}} = \mathbf{I}_N$:

$$\hat{\mathbf{U}}\hat{\mathbf{U}}^{\mathrm{H}} + \hat{\mathbf{U}}^{\perp}(\hat{\mathbf{U}}^{\perp})^{\mathrm{H}} = \mathbf{I}_N, \quad \hat{\mathbf{U}}\hat{\mathbf{U}}^{\mathrm{H}} = \mathbf{P}, \quad \hat{\mathbf{U}}^{\perp}(\hat{\mathbf{U}}^{\perp})^{\mathrm{H}} = \mathbf{P}^{\perp} = \mathbf{I}_N - \mathbf{P}$$

- Any vector $\mathbf{x} \in \mathbb{C}^N$ can be decomposed as $\mathbf{x} = \hat{\mathbf{x}} + \hat{\mathbf{x}}^{\perp}$ with $\hat{\mathbf{x}} \perp \hat{\mathbf{x}}^{\perp}$:

$$\hat{\mathbf{x}} = \mathbf{P}\mathbf{x} \in \mathrm{ran}(\hat{\mathbf{U}}) \quad \hat{\mathbf{x}}^{\perp} = \mathbf{P}^{\perp}\mathbf{x} \in \mathrm{ran}(\hat{\mathbf{U}}^{\perp})$$

Projection onto the column span of a tall matrix $\mathbf{A}$

- Suppose $\mathbf{A}$ has full column rank (i.e., $\mathbf{A}^{\mathrm{H}}\mathbf{A}$ is invertible). Then,

$$\mathbf{P}_{\mathbf{A}} := \mathbf{A}(\mathbf{A}^{\mathrm{H}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{H}}, \quad \mathbf{P}_{\mathbf{A}}^{\perp} := \mathbf{I} - \mathbf{A}(\mathbf{A}^{\mathrm{H}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{H}}$$

  are orthogonal projections onto $\mathrm{ran}(\mathbf{A})$ and $\ker(\mathbf{A}^{\mathrm{H}})$, respectively.

  - To prove, verify that $\mathbf{P}^2 = \mathbf{P}$ and $\mathbf{P}^{\mathrm{H}} = \mathbf{P}$, hence $\mathbf{P}$ is an orthogonal projection.

  - If $\mathbf{b} \in \mathrm{ran}(\mathbf{A})$, then $\mathbf{b} = \mathbf{A}\mathbf{x}$ for some $\mathbf{x}$:

$$\mathbf{P}_{\mathbf{A}}\mathbf{b} = \mathbf{A}(\mathbf{A}^{\mathrm{H}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{H}}\mathbf{A}\mathbf{x} = \mathbf{b}$$

    so that $\mathbf{b}$ is invariant under $\mathbf{P}_{\mathbf{A}}$.

  - If $\mathbf{b} \perp \mathrm{ran}(\mathbf{A})$, then $\mathbf{b} \in \ker(\mathbf{A}^{\mathrm{H}}) \Leftrightarrow \mathbf{A}^{\mathrm{H}}\mathbf{b} = \mathbf{0}$:

$$\mathbf{P}_{\mathbf{A}}\mathbf{b} = \mathbf{A}(\mathbf{A}^{\mathrm{H}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{H}}\mathbf{b} = \mathbf{0}.$$