SIGNAL SUBSPACES, ORTHOGONAL PROJECTIONS, AND LEAST SQUARES ESTIMATION

The Signal Subspace Model

Let \( a_1, \ldots, a_p \in \mathbb{R}^N \) (or \( \mathbb{C}^N \)) be linearly independent (so \( p \leq N \)), and consider the \( N \times p \) matrix

\[
A = [a_1 \ldots a_p].
\]

Let \( \langle A \rangle \) denote the linear span of the columns of \( A \) (equivalently, the image of \( A \)). Then

\[
\dim(\langle A \rangle) = \text{rank}(A) = p
\]

\[
\dim(\langle A \rangle^\perp) = N - p
\]
Let $b_1, \ldots, b_{N-p}$ be a basis for $\langle A \rangle^\perp$ and set

$$B = [b_1 \ldots b_{N-p}]_{N \times (N-p)}.$$

Then

$$\langle a_i, b_j \rangle = 0 \quad i = 1, \ldots, p \quad j = 1, \ldots, N-p.$$

and $\{a_1, \ldots, a_p, b_1, \ldots, b_{N-p}\}$ is a _basis_.

The subspaces $\langle A \rangle$ and $\langle B \rangle$ form an orthogonal decomposition of $\mathbb{R}^N$ (or $\mathbb{C}^N$)

$$\langle A \rangle \oplus \langle B \rangle = \mathbb{R}^N \quad \text{(or } \mathbb{C}^N)\)
Note: The vectors $a_1, \ldots, a_p$ are not necessarily orthogonal among themselves. The same goes for $b_1, \ldots, b_{n-p}$.

In the signal subspace model, we assume our observed signal $x$ has the form

$$x = s + w$$

where

$s \in \langle A \rangle$ is the signal of interest

$w$ is entirely noise.

We use the following terminology:

$$\langle A \rangle = \langle B \rangle$$

even though $w \notin \langle B \rangle$ in general.
Example: Quadratic polynomial model

\[ s(n) = \theta_2 n^2 + \theta_1 n + \theta_0 , \quad n = 1, \ldots, N \]

\[ \Rightarrow s = A \theta \quad \text{where} \]

\[ A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ \vdots & \vdots & \vdots \\ 1 & N & N^2 \end{bmatrix} \quad \theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{bmatrix} \in \mathbb{R}^3 \]

\[ \begin{array}{c}
\circ \to \text{clean signal (unobserved)} \\
\times \to \text{noisy signal (observed)}
\end{array} \]
Exercise  Consider the sinusoidal signal

\[ s(n) = D \cdot \cos(2\pi fn + \phi), \quad n = 0, \ldots, N-1 \]

where \( f \) is known but \( D, \phi \) are unknown.

Express \( s = [s(0) \ldots s(N-1)]^T \) as an element in a two-dimensional subspace.

That is, write

\[ s = A \cdot \theta \]

where \( A \) is a known \( N \times 2 \) matrix and \( \theta \) is unknown.
Solution 1

Use \( \cos(\alpha) = \frac{e^{j\alpha} + e^{-j\alpha}}{2} \). Then

\[
s(n) = \left( \frac{D}{2} e^{j\phi} \right) e^{2\pi j f n} + \left( \frac{D}{2} e^{-j\phi} \right) e^{-2\pi j f n}
\]

\[
\Rightarrow \mathbf{s} = \begin{bmatrix}
1 & 1 \\
e^{2\pi j f} & e^{-2\pi j f} \\
e^{4\pi j f} & e^{-4\pi j f} \\
\vdots & \vdots \\
e^{2(N-1)\pi j f} & e^{-2(N-1)\pi j f}
\end{bmatrix}
\begin{bmatrix}
\theta_1 \\
\theta_2
\end{bmatrix}
\]

\( a_1, a_2 \in \mathbb{C}^N, \quad \theta \in \mathbb{C}^2 \)

Solution 2

Use \( \cos(\alpha + \beta) = \cos(\alpha) \cos(\beta) - \sin(\alpha) \sin(\beta) \)

\[
s(n) = \left( D \cos(\phi) \right) \cos(2\pi f n) + \left( -D \sin(\phi) \right) \sin(2\pi f n)
\]

\[
\Rightarrow \mathbf{s} = \begin{bmatrix}
1 & 0 \\
\cos(2\pi f) & \sin(2\pi f) \\
\cos(4\pi f) & \sin(4\pi f) \\
\vdots & \vdots \\
\cos(2(N-1)\pi f) & \sin(2(N-1)\pi f)
\end{bmatrix}
\begin{bmatrix}
\theta_1 \\
\theta_2
\end{bmatrix}
\]

\( a_1, a_2 \in \mathbb{R}^N, \quad \theta \in \mathbb{R}^2 \) if \( D \in \mathbb{R} \)
Orthogonal Projection

How can we use the knowledge of $\langle A \rangle$ to estimate $s$ from $x = s + w$?

Since $\langle A \rangle$ and $\langle B \rangle$ are orthogonal complements, we can uniquely write

$$x = u + v$$

where $u \in \langle A \rangle$ and $v \in \langle B \rangle$.

Since $v$ is pure noise, it makes sense to remove it.

In general, $s \neq u$ because $w$ has some component in $\langle A \rangle$. 
The operator that maps \( x \mapsto u \) is called \( \Pi_A \) and is denoted \( \Pi_A \).

**Proposition**

\[
\Pi_A = A (A^H A)^{-1} A^H
\]

**Proof**

Let \( x = u + v \), \( u \in \langle A \rangle \), \( v \in \langle B \rangle \).

We must show \( A (A^H A)^{-1} A^H x = u \). Since the columns of \( A \) and \( B \) form a basis, we can write

\[
\begin{align*}
    u &= A \theta, \quad \theta \in \mathbb{R}^p \\
    v &= B \phi, \quad \phi \in \mathbb{R}^{(N-p)}
\end{align*}
\]

Then

\[
A (A^H A)^{-1} A^H x = A (A^H A)^{-1} A^H A \theta + A (A^H A)^{-1} A^H B \phi
\]

\[
= \underbrace{I_{p \times p}}_{\text{Identity matrix}} \theta + \underbrace{O_{p \times (N-p)}}_{\text{Zero matrix}} \phi
\]

\[
= A \theta
\]

\[
= u
\]
Properties of projections

- \( \Pi_A^H = \) "self-adjoint"
- \( \Pi_A^2 = \) "idempotent"
- \( \Pi_A + \Pi_B = \)
- \( \Pi_A \cdot \Pi_B = \)
- If \( a_1, \ldots, a_p \) are orthonormal, then
  \[ \Pi_A = A A^H \]

Filtering interpretation

The projection operator is analogous to a bandpass filter; we only retain that information which resides in the passband, which corresponds to the signal subspace.
Least Squares Estimation

To estimate $\hat{x} = A\theta$ where

$$x = \hat{x} + \omega$$

we use the projection onto $\langle A \rangle$:

$$\hat{\omega} = \Pi_A x$$

$$= A (A^H A)^{-1} A^H x.$$

What if we want to estimate $\theta$?

An estimate $\hat{\theta}$ of $\theta$ should satisfy

$$\hat{x} = A \hat{\theta}.$$

Therefore, an obvious estimate is

$$\hat{\theta} =$$

It turns out that this is the solution to the least squares problem.
Proposition 1. The unique solution of

$$\min_{\theta} \| x - A \theta \|^2 \quad (\theta \in \mathbb{R}^p \text{ or } \mathbb{C}^p)$$

is \( \hat{\theta} = (A^H A)^{-1} A^H x \).

Proof. Write \( x = u + v \) where \( u \in \langle A \rangle \)
and \( v \in \langle A \rangle^\perp \). Observe

$$\| x - A \theta \|^2 = \| u - A \theta + v \|^2$$

$$= \langle u - A \theta + v, u - A \theta + v \rangle$$

$$= \langle u - A \theta , u - A \theta \rangle + \langle v, v \rangle$$

$$+ \langle u - A \theta , v \rangle + \langle v, u - A \theta \rangle$$

$$= 0 \quad \quad \quad \quad \quad = 0$$

$$= \| u - A \theta \|^2 + \| v \|^2.$$ 

The second term is independent of \( \theta \). Therefore, to minimize the expression, the best we can do is to make the first term 0 by taking

\( \theta = \hat{\theta} = (A^H A)^{-1} A^H x \).
Then \( A\hat{\theta} = \Pi_A x = u \). To see that \( \hat{\theta} \) is unique, if \( \hat{\theta}' \) is also such that \( \|u - A\hat{\theta}'\| = 0 \), then
\[
A\hat{\theta}' = u \Rightarrow A\hat{\theta}' = A\hat{\theta} \Rightarrow \hat{\theta}' = \hat{\theta}
\]
since the columns of \( A \) are linearly independent.

**Minimum distance property**

We may conclude that \( \Pi_A x \) is the unique point in \( \langle A \rangle \) that is closest to \( x \).

**Remark**

The operator
\[
A^\# := (A^H A)^{-1} A^H
\]
is called the **pseudo-inverse** of \( A \).
Example

Recall the sinusoid

\[ s(n) = D \cdot \cos(2\pi fn + \phi), \quad n = 0, \ldots, N-1 \]

where \( f \) is known and \( D, \phi \) are unknown.

How can we estimate \( D \) and \( \phi \) from \( s \)?

Complex solution: \( s = A e^{j\theta} \) where

\[ \theta = \begin{bmatrix} \frac{D}{2} e^{j\phi} \\ \frac{D}{2} e^{-j\phi} \end{bmatrix}. \]

Use pseudo-inverse to compute \( \hat{\theta} \) and form

\[ \hat{D} = \sqrt{4 \cdot \hat{\theta} \cdot \hat{\theta}} \quad \text{and} \quad \hat{\phi} = \frac{1}{2} \tan^{-1} \left( \frac{\hat{\theta}_1}{\hat{\theta}_2} \right) \]

Real solution: \( s = A e^{\theta} \) where

\[ e = \begin{bmatrix} D \cos(\phi) \\ -D \sin(\phi) \end{bmatrix}. \]

Use pseudo-inverse to compute \( \hat{\theta} \) and form

\[ \hat{D} = \sqrt{\hat{\theta}_1^2 + \hat{\theta}_2^2} \quad \text{and} \quad \hat{\phi} = \tan^{-1} \left( \frac{\hat{\theta}_2}{\hat{\theta}_1} \right) \]

(assuming \( D \) to be real)
Summary

- If a signal lies in a subspace, it can be estimated by projection onto that subspace.
- This "filters out" any noise in the noise subspace.
- The projection satisfies the minimum distance property, and is closely related to the least squares problem.
- This approach is non-statistical because no probabilistic model is specified for the noise. Yet it turns out to be equivalent or similar to many methods we will see later.

Key:

a. basis
b. signal subspace, noise subspace
c. orthogonal projection
d. $\Pi_A$, $\Pi_A$, $I_{N \times N}$, $O_{N \times N}$
e. $(A^H A)^{-1} A^H x$