SIGNAL SUBSPACES, ORTHOGONAL PROJECTIONS, AND LEAST SQUARES ESTIMATION

The Signal Subspace Model

Let $\underline{a}_1, \ldots, \underline{a}_p \in \mathbb{R}^N$ (or $\mathbb{C}^N$) be linearly independent (so $p \leq N$), and consider the $N \times p$ matrix

$$A = [\underline{a}_1 \; \cdots \; \underline{a}_p].$$

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

$$\dim(\langle A \rangle) = \operatorname{rank}(A) = p$$

$$\dim(\langle A \rangle^\perp) = N - p$$

Let $\underline{b}_1, \dots, \underline{b}_{N-p} \in \mathbb{R}^N$ (or $\mathbb{C}^N$) be a basis for $\langle A \rangle^{\perp}$ and set

$$B = [\underline{b}_1 \ \cdots \ \underline{b}_{N-p}].$$

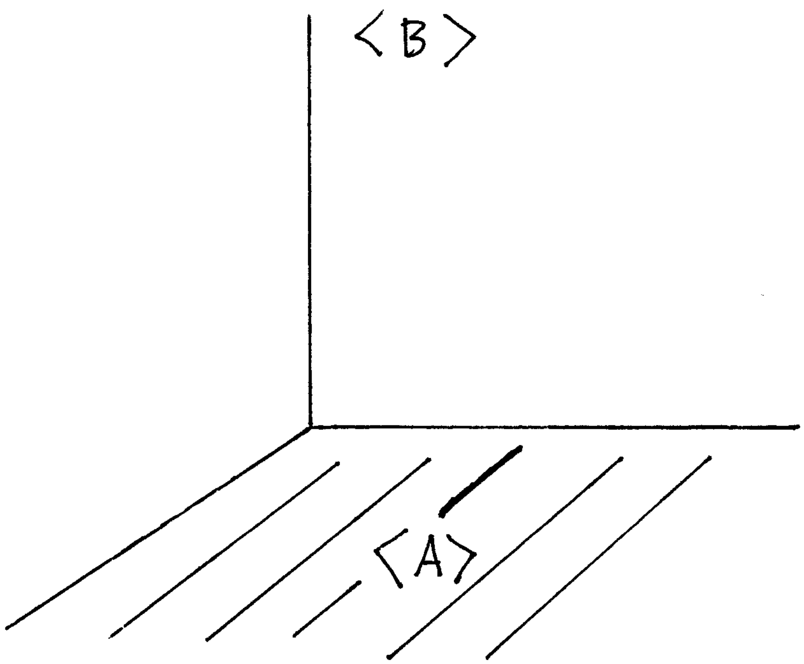
$N \times (N-p)$

Then

$$\langle \underline{a}_i, \underline{b}_j \rangle = 0 \qquad \begin{array}{l} i = 1, \dots, p \\ j = 1, \dots, N-p. \end{array}$$

ⓐ and $\{\underline{a}_1, \dots, \underline{a}_p, \underline{b}_1, \dots, \underline{b}_{N-p}\}$ is a __________.

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



$$\langle A \rangle \oplus \langle B \rangle = \mathbb{R}^N$$
(or $\mathbb{C}^N$)

Note) The vectors $\underline{a}_1, \ldots, \underline{a}_p$ are not necessarily orthoganal among themselves. The same goes for $\underline{b}_1, \ldots, \underline{b}_{N-p}$.

In the signal subspace model, we assume our observed signal $\underline{x}$ has the form

$$\underline{x} = \underline{s} + \underline{w}$$

where

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

$\underline{w}$ is entirely noise.

We use the following terminology:

$\langle A \rangle =$

(b) $\langle B \rangle =$

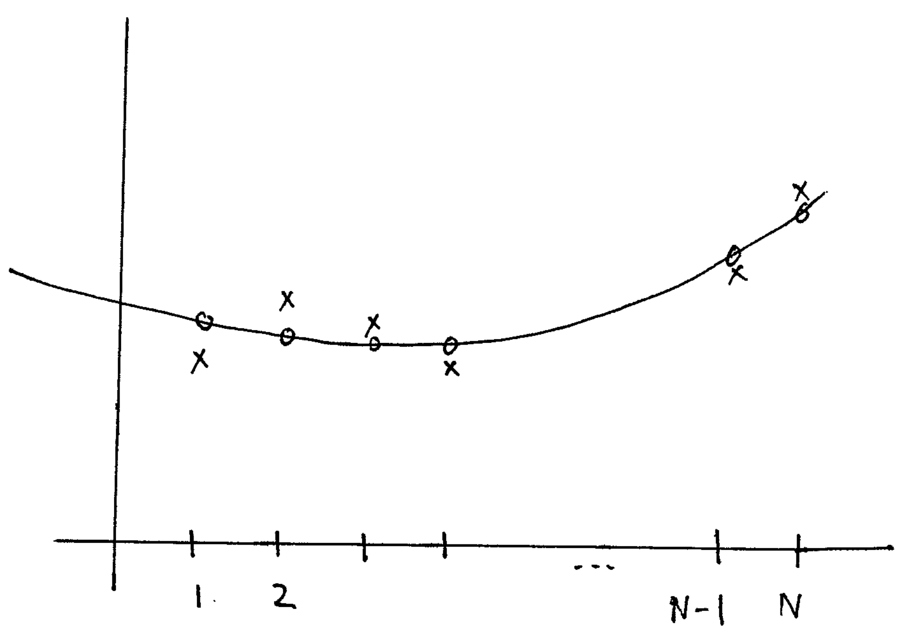
even though $w \notin \langle B \rangle$ in general.

Example | Quadratic polynomial model

$$s(n) = \theta_2 n^2 + \theta_1 n + \theta_0 , \quad n = 1, \dots, N$$

$$\Rightarrow \quad \underline{s} = A \underline{\theta} \quad \text{where}$$

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ \vdots & \vdots & \vdots \\ 1 & N & N^2 \end{bmatrix} \qquad \underline{\theta} = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{bmatrix} \in \mathbb{R}^3$$



o $\longrightarrow$ clean signal (unobserved)

x $\longrightarrow$ noisy signal (observed)

Exercise Consider the sinusoidal signal

$$s(n) = D \cdot \cos(2\pi f n + \phi), \quad n = 0, \dots, N-1$$

where f is known but D, ϕ are unknown.

Express $\underline{s} = [s(0) \ \cdots \ s(N-1)]^T$ as an element in a two-dimensional subspace.

That is, write

$$\underline{s} = A \cdot \underline{\theta}$$

where A is a <u>known</u> $N \times 2$ matrix and $\underline{\theta}$ is <u>unknown</u>.

Solution 1 Use $\cos(\alpha) = \dfrac{e^{j\alpha} + e^{-j\alpha}}{2}$. Then

$$s(n) = \underbrace{\left(\tfrac{D}{2} e^{j\phi}\right)}_{\theta_1} e^{2\pi j f n} + \underbrace{\left(\tfrac{D}{2} e^{-j\phi}\right)}_{\theta_2} e^{-2\pi j f n}$$

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

$$\underline{g}_1, \underline{g}_2 \in \mathbb{C}^N, \quad \underline{\theta} \in \mathbb{C}^2$$

Solution 2 Use $\cos(\alpha+\beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)$

$$s(n) = \underbrace{\left(D\cos(\phi)\right)}_{\theta_1} \cos(2\pi f n) + \underbrace{\left(-D\sin(\phi)\right)}_{\theta_2} \sin(2\pi f n)$$

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

$$\underline{g}_1, \underline{g}_2 \in \mathbb{R}^N, \quad \underline{\theta} \in \mathbb{R}^2 \quad \text{if } D \in \mathbb{R}$$

Orthogonal Projection

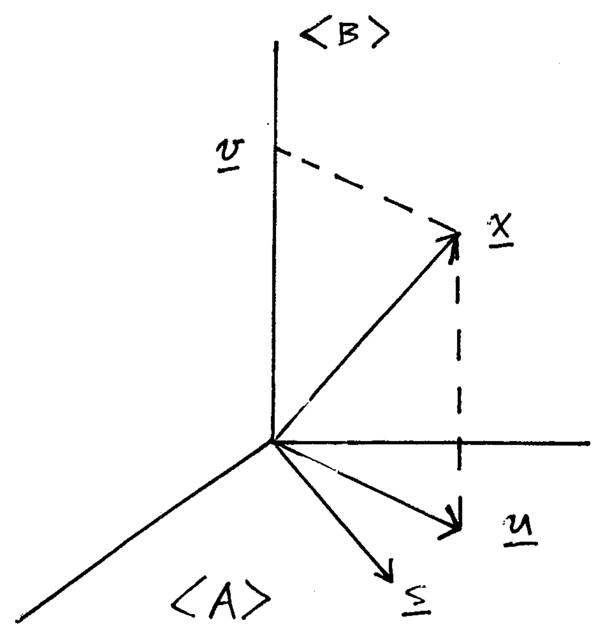
How can we use the knowledge of $\langle A \rangle$ to estimate $\underline{s}$ from $\underline{x} = \underline{s} + \underline{w}$?

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

$$\underline{x} = \underline{u} + \underline{v}$$

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

Since $\underline{v}$ is pure noise, it makes sense to remove it.



In general, $\underline{s} \neq \underline{u}$ because $\underline{w}$ has some component in $\langle A \rangle$.

The operator that maps $\underline{x} \mapsto \underline{u}$ is called

© ____________________ onto $\langle A \rangle$ and is

denoted Π_A.

Proposition

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

Proof Let $\underline{x} = \underline{u} + \underline{v}$, $\underline{u} \in \langle A \rangle$, $\underline{v} \in \langle B \rangle$.

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

$$\underline{u} = A\underline{\theta}, \qquad \underline{\theta} \quad p \times 1$$

$$\underline{v} = B\underline{\phi}, \qquad \underline{\phi} \quad (N-p) \times 1$$

Then

$$A(A^H A)^{-1} A^H \underline{x}$$

$$= A \underbrace{(A^H A)^{-1} A^H A}_{I_{p \times p}} \underline{\theta} + A(A^H A)^{-1} \underbrace{A^H B}_{0_{p \times (N-p)}} \underline{\phi}$$

$$= A\underline{\theta}$$

$$= \underline{u}$$

∎

Properties of projections

- $\Pi_A^H =$ "self-adjoint"
- $\Pi_A^2 =$ "idempotent"

(d)

- $\Pi_A + \Pi_B =$
- $\Pi_A \cdot \Pi_B =$
- If $\underline{a}_1, \ldots, \underline{a}_p$ 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 $\underline{s} = A\underline{\theta}$ where

$$\underline{x} = \underline{s} + \underline{w}$$

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

$$\hat{\underline{s}} = \Pi_A \underline{x}$$

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

What if we want to estimate $\underline{\theta}$?

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

$$\hat{\underline{s}} = A\hat{\underline{\theta}}.$$

Therefore, an obvious estimate is

e

$$\hat{\underline{\theta}} =$$

It turns out that this is the solution to the least squares problem.

Proposition The unique solution of

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

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

Proof Write $\underline{x} = \underline{u} + \underline{v}$ where $\underline{u} \in \langle A \rangle$ and $\underline{v} \in \langle A \rangle^{\perp}$. Observe

$$\begin{aligned}
\|\underline{x} - A\underline{\theta}\|^2 &= \|\underline{u} - A\underline{\theta} + \underline{v}\|^2 \\
&= \langle \underline{u} - A\underline{\theta} + \underline{v}, \underline{u} - A\underline{\theta} + \underline{v} \rangle \\
&= \langle \underline{u} - A\underline{\theta}, \underline{u} - A\underline{\theta} \rangle + \langle \underline{v}, \underline{v} \rangle \\
&\quad + \underbrace{\langle \underline{u} - A\underline{\theta}, \underline{v} \rangle}_{=0} + \underbrace{\langle \underline{v}, \underline{u} - A\underline{\theta} \rangle}_{=0} \\
&= \|\underline{u} - A\underline{\theta}\|^2 + \|\underline{v}\|^2.
\end{aligned}$$

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

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

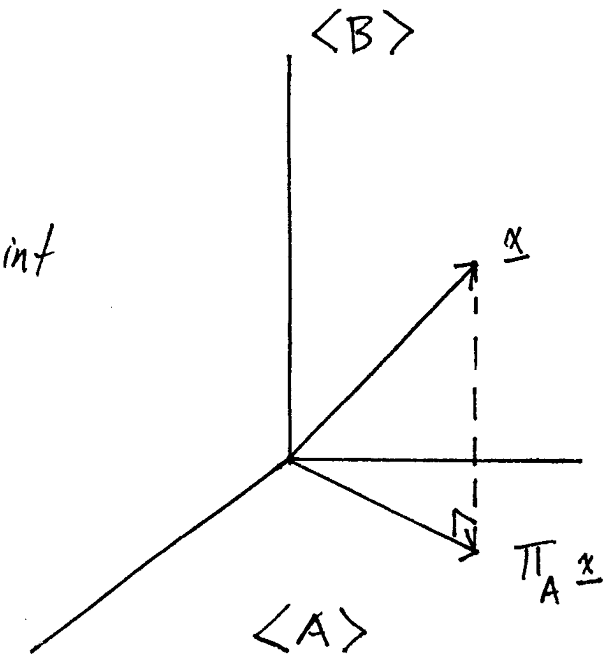
Then $A\hat{\underline{\theta}} = \Pi_A \underline{x} = \underline{u}$. To see that $\hat{\underline{\theta}}$ is unique, if $\underline{\theta}'$ is also such that $\|\underline{u} - A\underline{\theta}'\| = 0$, then

$$A\underline{\theta}' = \underline{u} \implies A\underline{\theta}' = A\hat{\underline{\theta}}$$
$$\implies \underline{\theta}' = \hat{\underline{\theta}}$$

since the columns of A are linearly independent. ∎

<u>Minimum distance property</u>

We may conclude that $\Pi_A \underline{x}$ is the unique point in $\langle A \rangle$ that is closest to $\underline{x}$.



<u>Remark</u> The operator

$$A^{\#} := (A^H A)^{-1} A^H$$

is called the <u>pseudo-inverse</u> of A.

Example Recall the sinusoid

$$s(n) = D \cdot \cos(2\pi f n + \phi) \quad , \qquad n = 0, \dots, N-1$$

where f is known and D, ϕ are unknown.

How can we estimate D and ϕ from $\underline{x}$?

Complex solution: $\underline{s} = A \underline{\theta}$ where

$$\underline{\theta} = \begin{bmatrix} \frac{D}{2} e^{j\phi} \\ \frac{D}{2} e^{-j\phi} \end{bmatrix}.$$ Use pseudo-inverse to

compute $\hat{\underline{\theta}}$ and form

$$\hat{D} = \sqrt{4 \cdot \hat{\theta}_1 \cdot \hat{\theta}_2} \quad , \quad \hat{\phi} = \frac{1}{2} \measuredangle \left(\frac{\hat{\theta}_1}{\hat{\theta}_2} \right)$$

Real solution: $\underline{s} = A \underline{\theta}$ where

$$\underline{\theta} = \begin{bmatrix} D \cos(\phi) \\ -D \sin(\phi) \end{bmatrix}.$$ Use pseudo-inverse to

compute $\hat{\underline{\theta}}$ and form

$$\hat{D} = \sqrt{\hat{\theta}_1^2 + \hat{\theta}_2^2} \quad , \quad \hat{\phi} = \tan^{-1}\left(-\frac{\hat{\theta}_2}{\hat{\theta}_1} \right)$$

(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 <u>non-statistical</u> because no probability 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. Π_A , Π_A , $I_{N \times N}$, $O_{N \times N}$

e. $(A^H A)^{-1} A^H \mathbf{x}$