Inverse Problems
Least squares solutions

Clarice Poon
University of Bath

January 27, 2020

Least square solutions

Let $A : \mathcal{U} \to \mathcal{V}$ be a bounded linear operator between Hilbert spaces $\mathcal{U}$ and $\mathcal{V}$ and consider

$$Au = f.$$

Definition 1

An element $u \in \mathcal{U}$ is called

- a least-squares solution if $\|Au - f\|_{\mathcal{V}} = \inf \left\{ \|Av - f\|_{\mathcal{V}} \ ; \ v \in \mathcal{U} \right\}$.
- a minimal-norm solution, denoted by $u^\dagger$ if

$$\|u^\dagger\|_{\mathcal{U}} \leqslant \|v\|_{\mathcal{U}}$$

  for all least-squares solutions v.

NB: If $f \notin \mathcal{R}(A)$, then a least squares solution will not satisfy $Au = f$.

Learning objectives

1. Look at characterisation of least squares solutions (properties, existence, uniqueness)
2. Review the pseudo-inverse and see that this gives the minimal norm solution.
3. Compact operators (ill-posedness, properties of pseudo-inverse and SVD decomposition).

Definitions

Domain, kernel and range

Given $A : \mathcal{U} \to \mathcal{V}$, denote

- $\mathcal{D}(A) \stackrel{\text{def.}}{=} \mathcal{U}$ the domain,
- $\mathcal{N}(A) \stackrel{\text{def.}}{=} \{u \in \mathcal{U} \ ; \ Au = 0\}$ the kernel,
- $\mathcal{R}(A) \stackrel{\text{def.}}{=} \{f \in \mathcal{V} \ ; \ f = Au, u \in \mathcal{U}\}$ the range.

Continuous linear operators

We say that A is continuous at $u \in \mathcal{U}$ if for all $\varepsilon > 0$, there exists $\delta > 0$ with

$$\|Au - Av\|_{\mathcal{V}} \leqslant \varepsilon \quad \forall v \in \mathcal{U} \quad s.t. \quad \|u - v\|_{\mathcal{U}} \leqslant \delta.$$

If A is a linear operator, then A is continuous if and only if it is bounded.

We will focus on inverse problems with bounded linear operators $A \in \mathcal{L}(\mathcal{U}, \mathcal{V})$ with $\|A\|_{\mathcal{L}(\mathcal{U},\mathcal{V})} \stackrel{\text{def.}}{=} \sup_{\|u\|_{\mathcal{U}} \leqslant 1} \|Au\|_{\mathcal{V}} < \infty$.

Let $A \in \mathcal{L}(\mathcal{U}, \mathcal{V})$ where $\mathcal{U}$ and $\mathcal{V}$ are Hilbert spaces. Every Hilbert space $\mathcal{U}$ is equipped with a scalar product $\langle \cdot, \cdot \rangle_{\mathcal{U}}$.

Adjoint

The unique **adjoint** operator of A is A^* defined by

$$\langle Au, v \rangle_{\mathcal{V}} = \langle u, A^* v \rangle_{\mathcal{U}}, \qquad \forall u \in \mathcal{U}, v \in \mathcal{V}.$$

Let $A \in \mathcal{L}(\mathcal{U}, \mathcal{V})$ where $\mathcal{U}$ and $\mathcal{V}$ are Hilbert spaces. Every Hilbert space $\mathcal{U}$ is equipped with a scalar product $\langle \cdot, \cdot \rangle_{\mathcal{U}}$.

Adjoint

The unique **adjoint** operator of A is A^* defined by

$$\langle Au, v \rangle_{\mathcal{V}} = \langle u, A^* v \rangle_{\mathcal{U}}, \qquad \forall u \in \mathcal{U}, v \in \mathcal{V}.$$

Orthogonal complement

We say that $u, v \in \mathcal{U}$ are orthogonal if $\langle u, v \rangle = 0$. Given $\mathcal{X} \subseteq \mathcal{U}$, the **orthogonal complement** of $\mathcal{X}$ in $\mathcal{U}$ is

$$\mathcal{X}^\perp \overset{\text{def.}}{=} \{u \in \mathcal{U} \,;\, \langle u, v \rangle = 0 \quad \forall v \in \mathcal{X}\}.$$

- $\mathcal{X}^\perp$ is a closed subspace of $\mathcal{U}$ and $\mathcal{U}^\perp = \{0\}$.
- $\overline{\mathcal{X}} = (\mathcal{X}^\perp)^\perp$.
- If $\mathcal{X}$ is closed, then $\mathcal{X} = (\mathcal{X}^\perp)^\perp$ and $\mathcal{U} = \mathcal{X} \oplus \mathcal{X}^\perp$.

Orthogonal projection

Let $\mathcal{X} \subset \mathcal{U}$ be a closed subspace. Then, for all $u \in \mathcal{U}$, there exists $x \in \mathcal{X}$ and $x^\perp \in \mathcal{X}^\perp$ such that $u = x + x^\perp$. The mapping $u \mapsto x$ defines a bounded linear operator $P_\mathcal{X}$ called the **orthogonal projection** onto $\mathcal{X}$.

Elementary facts about Hilbert spaces

Orthogonal projection

Let $\mathcal{X} \subset \mathcal{U}$ be a closed subspace. Then, for all $u \in \mathcal{U}$, there exists $x \in \mathcal{X}$ and $x^\perp \in \mathcal{X}^\perp$ such that $u = x + x^\perp$. The mapping $u \mapsto x$ defines a bounded linear operator $P_\mathcal{X}$ called the **orthogonal projection** onto $\mathcal{X}$.

- $P_\mathcal{X}$ is self-adjoint.
- $\|P_\mathcal{X}\| = 1$ if $\mathcal{X} \neq \{0\}$.
- $\mathrm{Id} - P_\mathcal{X} = P_{\mathcal{X}^\perp}$.
- $\|u - P_\mathcal{X} u\|_\mathcal{U} \leqslant \|u - v\|$ for all $v \in \mathcal{X}$.
- $x = P_\mathcal{X} u$ if and only if $x \in \mathcal{X}$ and $u - x \in \mathcal{X}^\perp$.

Elementary facts about Hilbert spaces

Orthogonal projection

Let $\mathcal{X} \subset \mathcal{U}$ be a closed subspace. Then, for all $u \in \mathcal{U}$, there exists $x \in \mathcal{X}$ and $x^\perp \in \mathcal{X}^\perp$ such that $u = x + x^\perp$. The mapping $u \mapsto x$ defines a bounded linear operator $P_\mathcal{X}$ called the **orthogonal projection** onto $\mathcal{X}$.

- $P_\mathcal{X}$ is self-adjoint.
- $\|P_\mathcal{X}\| = 1$ if $\mathcal{X} \neq \{0\}$.
- $\mathrm{Id} - P_\mathcal{X} = P_{\mathcal{X}^\perp}$.
- $\|u - P_\mathcal{X} u\|_\mathcal{U} \leqslant \|u - v\|$ for all $v \in \mathcal{X}$.
- $x = P_\mathcal{X} u$ if and only if $x \in \mathcal{X}$ and $u - x \in \mathcal{X}^\perp$.

Properties of range and kernel

For $A \in \mathcal{L}(\mathcal{U}, \mathcal{V})$, we have

- $\mathcal{R}(A)^\perp = \mathcal{N}(A^*)$ and thus, $\mathcal{N}(A^*)^\perp = \overline{\mathcal{R}(A)}$.
- $\mathcal{R}(A^*)^\perp = \mathcal{N}(A)$ and thus, $\mathcal{N}(A)^\perp = \overline{\mathcal{R}(A^*)}$.

So, $\mathcal{U} = \mathcal{N}(A) \oplus \overline{\mathcal{R}(A^*)}$ and $\mathcal{V} = \mathcal{N}(A^*) \oplus \overline{\mathcal{R}(A)}$. Also, $\overline{\mathcal{R}(A^*A)} = \overline{\mathcal{R}(A^*)}$.

Theorem 2

Let $f \in \mathcal{V}$ and $A \in \mathcal{L}(\mathcal{U}, \mathcal{V})$. Then, the following are equivalent.

(a) *u is a least-squares solution to $Au = f$.*

(b) *u solves the normal equation $A^*Au = A^*f$.*

(c) *u satisfies $Au = P_{\overline{\mathcal{R}(A)}}f$.*

Theorem 2

Let $f \in \mathcal{V}$ and $A \in \mathcal{L}(\mathcal{U}, \mathcal{V})$. Then, the following are equivalent.

(a) *u is a least-squares solution to $Au = f$.*

(b) *u solves the normal equation $A^*Au = A^*f$.*

(c) *u satisfies $Au = P_{\overline{\mathcal{R}(A)}}f$.*

Remark. For any solution u to the normal equation, its residue $Au - f$ is normal (orthogonal) to $\mathcal{R}(A)$: for any $v \in \mathcal{U}$,

$$0 = \langle v, A^*(Au - f)\rangle_{\mathcal{U}} = \langle Av, Au - f\rangle_{\mathcal{V}}.$$

LS solution satisfy the normal equation: (a)$\to$(b)

For any $v \in \mathcal{U}$, $F(\lambda) = \|A(u + \lambda v) - f\|_{\mathcal{V}}^2$ is smallest at $\lambda = 0$. So, $F'(0) = 2\langle Av, Au - f\rangle_{\mathcal{V}} = 0$.

LS solution satisfy the normal equation: (a)→(b)

For any $v \in \mathcal{U}$, $F(\lambda) = \|A(u + \lambda v) - f\|_{\mathcal{V}}^2$ is smallest at $\lambda = 0$. So, $F'(0) = 2\langle Av, Au - f\rangle_{\mathcal{V}} = 0$.

Normal equation implies $Au = P_{\overline{\mathcal{R}(A)}} f$: (b)→(c)

If $A^*(f - Au) = 0$, then $f - Au \in \mathcal{N}(A^*) = \mathcal{R}(A)^{\perp} = (\overline{\mathcal{R}(A)})^{\perp}$. So,

$$\forall v \in \mathcal{V}, \quad \langle P_{\overline{\mathcal{R}(A)}}(f - Au), v\rangle = \langle f - Au, P_{\overline{\mathcal{R}(A)}} v\rangle = 0$$

implies $P_{\overline{\mathcal{R}(A)}}(f - Au) = P_{\overline{\mathcal{R}(A)}}(f) - Au = 0$.

Existence and characterisation of least squares solutions

LS solution satisfy the normal equation: (a)→(b)

For any $v \in \mathcal{U}$, $F(\lambda) = \|A(u + \lambda v) - f\|_{\mathcal{V}}^2$ is smallest at $\lambda = 0$. So, $F'(0) = 2\langle Av, Au - f\rangle_{\mathcal{V}} = 0$.

Normal equation implies $Au = P_{\overline{\mathcal{R}(A)}} f$: (b)→(c)

If $A^*(f - Au) = 0$, then $f - Au \in \mathcal{N}(A^*) = \mathcal{R}(A)^\perp = (\overline{\mathcal{R}(A)})^\perp$. So,

$$\forall v \in \mathcal{V}, \quad \langle P_{\overline{\mathcal{R}(A)}}(f - Au), v\rangle = \langle f - Au, P_{\overline{\mathcal{R}(A)}} v\rangle = 0$$

implies $P_{\overline{\mathcal{R}(A)}}(f - Au) = P_{\overline{\mathcal{R}(A)}}(f) - Au = 0$.

Any solution to $Au = P_{\overline{\mathcal{R}(A)}} f$ is a LS solution: (c)→(a)

For $Au = \mathcal{P}_{\overline{\mathcal{R}(A)}} f$, we have

$$\|Au - f\|_{\mathcal{V}}^2 = \left\|(\mathrm{Id} - \mathcal{P}_{\overline{\mathcal{R}(A)}})f\right\| \leqslant \inf_{g \in \overline{\mathcal{R}(A)}} \|g - f\|_{\mathcal{V}} \leqslant \inf_{v \in \mathcal{U}} \|Av - f\|_{\mathcal{V}}^2.$$

When does a least squares solution exist?

Lemma 3

Let $f \in \mathcal{V}$ and let $\mathbb{L}$ be the set of least squares solutions to $Au = f$.

(a) $\mathbb{L} \neq \emptyset$ *if and only if* $f \in \mathcal{R}(A) \oplus \mathcal{R}(A)^{\perp}$.

(b) *If* $\mathbb{L} \neq \emptyset$*, then there exists a unique minimal norm solution* $u^{\dagger}$ *and all least squares solutions are given by* $u^{\dagger} + \mathcal{N}(A)$.

When does a least squares solution exist?

Lemma 3

Let $f \in \mathcal{V}$ and let $\mathbb{L}$ be the set of least squares solutions to $Au = f$.

(a) $\mathbb{L} \neq \emptyset$ *if and only if* $f \in \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$.

(b) *If* $\mathbb{L} \neq \emptyset$, *then there exists a unique minimal norm solution* $u^\dagger$ *and all least squares solutions are given by* $u^\dagger + \mathcal{N}(A)$.

NB: In the finite dimensional setting, $\mathcal{R}(A)$ is closed, i.e. $\mathcal{R}(A) = \overline{\mathcal{R}(A)}$ so there always exists a least-squares solution.

When does a least squares solution exist?

Lemma 3

Let $f \in \mathcal{V}$ and let $\mathbb{L}$ be the set of least squares solutions to $Au = f$.

(a) $\mathbb{L} \neq \emptyset$ *if and only if* $f \in \mathcal{R}(A) \oplus \mathcal{R}(A)^{\perp}$.

(b) *If* $\mathbb{L} \neq \emptyset$, *then there exists a unique minimal norm solution* $u^{\dagger}$ *and all least squares solutions are given by* $u^{\dagger} + \mathcal{N}(A)$.

NB: In the finite dimensional setting, $\mathcal{R}(A)$ is closed, i.e. $\mathcal{R}(A) = \overline{\mathcal{R}(A)}$ so there always exists a least-squares solution.

For (a):
If $u \in \mathbb{L}$, then $f = Au + (f - Au) \in \mathcal{R}(A) \oplus \mathcal{R}(A)^{\perp}$
(we just saw that the normal equation implies $f - Au \in \mathcal{R}(A)^{\perp}$).

When does a least squares solution exist?

Lemma 3

Let $f \in \mathcal{V}$ and let $\mathbb{L}$ be the set of least squares solutions to $Au = f$.

(a) $\mathbb{L} \neq \emptyset$ *if and only if* $f \in \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$.

(b) *If* $\mathbb{L} \neq \emptyset$*, then there exists a unique minimal norm solution* $u^\dagger$ *and all least squares solutions are given by* $u^\dagger + \mathcal{N}(A)$.

NB: In the finite dimensional setting, $\mathcal{R}(A)$ is closed, i.e. $\mathcal{R}(A) = \overline{\mathcal{R}(A)}$ so there always exists a least-squares solution.

For (a):
If $u \in \mathbb{L}$, then $f = Au + (f - Au) \in \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$
(we just saw that the normal equation implies $f - Au \in \mathcal{R}(A)^\perp$).

Conversely, if $f \in \mathcal{R}(A) \oplus (\mathcal{R}(A))^\perp$, then there exists $u \in \mathcal{U}$ and $g \in \mathcal{R}(A)^\perp$ such that $f = Au + g$. Therefore, $P_{\overline{\mathcal{R}(A)}} f = Au$.

When does a least squares solution exist?

Lemma 3

Let $f \in \mathcal{V}$ and let $\mathbb{L}$ be the set of least squares solutions to $Au = f$.

(a) *$\mathbb{L} \neq \emptyset$ if and only if $f \in \mathcal{R}(A) \oplus \mathcal{R}(A)^{\perp}$.*

(b) *If $\mathbb{L} \neq \emptyset$, then there exists a unique minimal norm solution $u^{\dagger}$ and all least squares solutions are given by $u^{\dagger} + \mathcal{N}(A)$.*

NB: In the finite dimensional setting, $\mathcal{R}(A)$ is closed, i.e. $\mathcal{R}(A) = \overline{\mathcal{R}(A)}$ so there always exists a least-squares solution.

To show (b): Letting $B = \min_u \|Au - f\|$, the minimal solution is

$$u^{\dagger} = \operatorname{argmin}_{u \in \{v \;;\; \|Av - f\| = B\}} \|u\|$$

The orthogonal projection of the zero element onto a closed convex set is unique.

When does a least squares solution exist?

Lemma 3

Let $f \in \mathcal{V}$ and let $\mathbb{L}$ be the set of least squares solutions to $Au = f$.

(a) $\mathbb{L} \neq \emptyset$ *if and only if* $f \in \mathcal{R}(A) \oplus \mathcal{R}(A)^{\perp}$.

(b) *If* $\mathbb{L} \neq \emptyset$, *then there exists a unique minimal norm solution* $u^{\dagger}$ *and all least squares solutions are given by* $u^{\dagger} + \mathcal{N}(A)$.

NB: In the finite dimensional setting, $\mathcal{R}(A)$ is closed, i.e. $\mathcal{R}(A) = \overline{\mathcal{R}(A)}$ so there always exists a least-squares solution.

To show (b): Letting $B = \min_u \|Au - f\|$, the minimal solution is

$$u^{\dagger} = \operatorname{argmin}_{u \in \{v \,;\, \|Av - f\| = B\}} \|u\|$$

The orthogonal projection of the zero element onto a closed convex set is unique.

Given any least squares solution φ, $A(\varphi - u^{\dagger}) = P_{\overline{\mathcal{R}(A)}}(f) - P_{\overline{\mathcal{R}(A)}}(f) = 0$.
Therefore, $\varphi - u^{\dagger} \in \mathcal{N}(A)$.

Let $A \in \mathcal{L}(\mathcal{U}, \mathcal{V})$, its inverse might not exist but we can define a generalised inverse by restricting its domain and range.

The Moore-Penrose generalised inverse

Let $A \in \mathcal{L}(\mathcal{U}, \mathcal{V})$, its inverse might not exist but we can define a generalised inverse by restricting its domain and range.

Consider the restriction of A to $\mathcal{N}(A)^\perp$:

$$\tilde{A} \stackrel{\text{def.}}{=} A|_{\mathcal{N}(A)^\perp} : \mathcal{N}(A)^\perp \to \mathcal{V}$$

$\tilde{A}$ has an inverse $\tilde{A}^{-1} : \mathcal{R}(A) \to \mathcal{N}(A)^\perp$ because

The Moore-Penrose generalised inverse

Let $A \in \mathcal{L}(\mathcal{U}, \mathcal{V})$, its inverse might not exist but we can define a generalised inverse by restricting its domain and range.

Consider the restriction of A to $\mathcal{N}(A)^\perp$:

$$\tilde{A} \overset{\text{def.}}{=} A|_{\mathcal{N}(A)^\perp} : \mathcal{N}(A)^\perp \to \mathcal{V}$$

$\tilde{A}$ has an inverse $\tilde{A}^{-1} : \mathcal{R}(A) \to \mathcal{N}(A)^\perp$ because

- $\tilde{A}$ is a bounded linear operator.
- **Injective**: for all $x \in \mathcal{N}(A)^\perp$, $\tilde{A}x = 0$ iff $x \in \mathcal{N}(A) \cap \mathcal{N}(A)^\perp = \{0\}$.
- **Range** is $\mathcal{R}(\tilde{A}) = \mathcal{R}(A)$: Given $y = Ax \in \mathcal{R}(A)$, write $x = x_1 + x_2 \in \mathcal{N}(A)^\perp \oplus \mathcal{N}(A)$, then $\tilde{A}x_1 = Ax_1 = Ax = y$ so $y \in \mathcal{R}(\tilde{A})$.

The Moore-Penrose generalised inverse

Let $A \in \mathcal{L}(\mathcal{U}, \mathcal{V})$, its inverse might not exist but we can define a generalised inverse by restricting its domain and range.

Consider the restriction of A to $\mathcal{N}(A)^\perp$:

$$\tilde{A} \stackrel{\text{def.}}{=} A|_{\mathcal{N}(A)^\perp} : \mathcal{N}(A)^\perp \to \mathcal{V}$$

$\tilde{A}$ has an inverse $\tilde{A}^{-1} : \mathcal{R}(A) \to \mathcal{N}(A)^\perp$ because

- $\tilde{A}$ is a bounded linear operator.
- **Injective**: for all $x \in \mathcal{N}(A)^\perp$, $\tilde{A}x = 0$ iff $x \in \mathcal{N}(A) \cap \mathcal{N}(A)^\perp = \{0\}$.
- **Range** is $\mathcal{R}(\tilde{A}) = \mathcal{R}(A)$: Given $y = Ax \in \mathcal{R}(A)$, write $x = x_1 + x_2 \in \mathcal{N}(A)^\perp \oplus \mathcal{N}(A)$, then $\tilde{A}x_1 = Ax_1 = Ax = y$ so $y \in \mathcal{R}(\tilde{A})$.

Extend $\tilde{A}^{-1}$ by zero on $\mathcal{R}(A)^\perp$ to obtain $A^\dagger : \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp \to \mathcal{N}(A)^\perp$ which satisfies $AA^\dagger x = x$ for all $x \in \mathcal{R}(A)$.

The Moore-Penrose generalised inverse

Let $A \in \mathcal{L}(\mathcal{U}, \mathcal{V})$, its inverse might not exist but we can define a generalised inverse by restricting its domain and range.

Consider the restriction of A to $\mathcal{N}(A)^\perp$:

$$\tilde{A} \stackrel{\text{def.}}{=} A|_{\mathcal{N}(A)^\perp} : \mathcal{N}(A)^\perp \to \mathcal{V}$$

$\tilde{A}$ has an inverse $\tilde{A}^{-1} : \mathcal{R}(A) \to \mathcal{N}(A)^\perp$ because

- $\tilde{A}$ is a bounded linear operator.
- **Injective**: for all $x \in \mathcal{N}(A)^\perp$, $\tilde{A}x = 0$ iff $x \in \mathcal{N}(A) \cap \mathcal{N}(A)^\perp = \{0\}$.
- **Range** is $\mathcal{R}(\tilde{A}) = \mathcal{R}(A)$: Given $y = Ax \in \mathcal{R}(A)$, write $x = x_1 + x_2 \in \mathcal{N}(A)^\perp \oplus \mathcal{N}(A)$, then $\tilde{A}x_1 = Ax_1 = Ax = y$ so $y \in \mathcal{R}(\tilde{A})$.

Extend $\tilde{A}^{-1}$ by zero on $\mathcal{R}(A)^\perp$ to obtain $A^\dagger : \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp \to \mathcal{N}(A)^\perp$ which satisfies $AA^\dagger x = x$ for all $x \in \mathcal{R}(A)$.

The Moore-Penrose generalised inverse

This extension is called the Moore-Penrose generalised inverse. It is the unique extension of $\tilde{A}^{-1}$ to $\mathcal{D}(A^\dagger) = \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$ with $\mathcal{N}(A^\dagger) = R(A)^\perp$.

- Note that $\mathcal{D}(A^\dagger)$ is dense in $\mathcal{U}$, so $A^\dagger$ is defined on the entire domain $\mathcal{U}$ only if $\mathcal{R}(A)$ is closed.
- In fact, $A^\dagger$ is bounded (continuous) if and only if $R(A)$ is closed.

We list a few additional properties of $A^\dagger$:

Lemma 4

The Moore-Penrose inverse $A^\dagger$ satisfies $\mathcal{R}(A^\dagger) = \mathcal{N}(A)^\perp$ and the Moore-Penrose equations

(a) $AA^\dagger A = A.$

(b) $A^\dagger AA^\dagger = A^\dagger$

(c) $A^\dagger A = \mathrm{Id} - P_{\mathcal{N}(A)}$

(d) $AA^\dagger = P_{\overline{\mathcal{R}(A)}}|_{\mathcal{D}(A^\dagger)}.$

Theorem 5

For $f \in \mathcal{D}(A^\dagger)$, the minimal norm solution $u^\dagger$ is given by $u^\dagger = A^\dagger f$.

Proof.

We already saw that $u^\dagger$ exists and is unique.

The Moore-Penrose generalised inverse

Theorem 5

For $f \in \mathcal{D}(A^\dagger)$, the minimal norm solution $u^\dagger$ is given by $u^\dagger = A^\dagger f$.

Proof.

We already saw that $u^\dagger$ exists and is unique.

Note that $u^\dagger \in \mathcal{N}(A)^\perp$: decompose $u^\dagger = v + w$ for some $v \in \mathcal{N}(A)$ and $w \in \mathcal{N}(A)^\perp$. Then, w is a least squares solution and $\left\|u^\dagger\right\|^2 = \|v\|^2 + \|w\|^2$, so $v = 0$ by minimality of $u^\dagger$.

Theorem 5

For $f \in \mathcal{D}(A^\dagger)$, the minimal norm solution $u^\dagger$ is given by $u^\dagger = A^\dagger f$.

Proof.

We already saw that $u^\dagger$ exists and is unique.

Note that $u^\dagger \in \mathcal{N}(A)^\perp$: decompose $u^\dagger = v + w$ for some $v \in \mathcal{N}(A)$ and $w \in \mathcal{N}(A)^\perp$. Then, w is a least squares solution and $\left\|u^\dagger\right\|^2 = \|v\|^2 + \|w\|^2$, so $v = 0$ by minimality of $u^\dagger$.

So,

$$u^\dagger = (\mathrm{Id} - P_{\mathcal{N}(A)})u^\dagger = A^\dagger A u^\dagger = A^\dagger P_{\overline{\mathcal{R}(A)}} f = A^\dagger A A^\dagger f = A^\dagger f.$$

□

Theorem 5

For $f \in \mathcal{D}(A^\dagger)$, the minimal norm solution $u^\dagger$ is given by $u^\dagger = A^\dagger f$.

Proof.

We already saw that $u^\dagger$ exists and is unique.

Note that $u^\dagger \in \mathcal{N}(A)^\perp$: decompose $u^\dagger = v + w$ for some $v \in \mathcal{N}(A)$ and $w \in \mathcal{N}(A)^\perp$. Then, w is a least squares solution and $\left\|u^\dagger\right\|^2 = \|v\|^2 + \|w\|^2$, so $v = 0$ by minimality of $u^\dagger$.

So,

$$u^\dagger = (\mathrm{Id} - P_{\mathcal{N}(A)})u^\dagger = A^\dagger A u^\dagger = A^\dagger P_{\overline{\mathcal{R}(A)}} f = A^\dagger A A^\dagger f = A^\dagger f.$$

□

Note also that since $\mathcal{N}(A^*A)^\perp = \overline{\mathcal{R}(A^*A)} = \overline{\mathcal{R}(A^*)} = \mathcal{N}(A)^\perp$, the normal equation $(A^*A)u^\dagger - A^*f = 0$ implies that

$$u^\dagger = P_{\overline{\mathcal{R}(A^*)}} u^\dagger = (A^*A)^\dagger (A^*A) u^\dagger = (A^*A)^\dagger A^* f.$$

Definition 6 (Compact operators)

Let $A \in \mathcal{L}(\mathcal{U}, \mathcal{V})$. Then A is compact if for any bounded set $\mathcal{B} \subset \mathcal{U}$, the closure of its image $\overline{A(\mathcal{B})}$ is compact in $\mathcal{V}$. We denote the space of compact operators by $\mathcal{K}(\mathcal{U}, \mathcal{V})$.

Equivalently, if $\{u_j\} \subset \mathcal{U}$ is bounded, then $\{Au_j\}$ has a convergent subsequence in $\mathcal{V}$.

Examples 1 If A has finite range, then A is compact (by Bolzano-Weierstrass).

Examples 1 If A has finite range, then A is compact (by Bolzano-Weierstrass).

Examples 2 Let $g \in C([0,1];\mathbb{R})$ and define $A : C([0,1];\mathbb{R}) \to C([0,1];\mathbb{R})$ by

$$(Af)(x) \overset{\text{def.}}{=} \int_0^x f(t)g(t)\mathrm{d}t$$

This is compact by Arzela-Ascoli (given $\{f_n\}_n$ uniformly bounded, check that $\{Af_n\}_n$ is also uniformly bounded and equi-continuous).

Examples 1 If A has finite range, then A is compact (by Bolzano-Weierstrass).

Examples 2 Let $g \in C([0,1];\mathbb{R})$ and define $A : C([0,1];\mathbb{R}) \to C([0,1];\mathbb{R})$ by

$$(Af)(x) \stackrel{\text{def.}}{=} \int_0^x f(t)g(t)\mathrm{d}t$$

This is compact by Arzela-Ascoli (given $\{f_n\}_n$ uniformly bounded, check that $\{Af_n\}_n$ is also uniformly bounded and equi-continuous).

Examples 3 Let $k \in L^2(\Omega \times \Omega)$. The operator $A : L^2(\Omega) \to L^2(\Omega)$ defined by

$$(Af)(x) = \int k(x,y)f(y)\mathrm{d}y$$

is compact.

Compact operators are very common in inverse problems. This is a major source of ill-posedness.

Theorem 7

Let $A \in \mathcal{K}(\mathcal{U}, \mathcal{V})$ with infinite dimensional range. Then, the Moore-Penrose inverse of A is discontinuous.

Proof.

Since $\mathcal{R}(A)$ is infinite dimensional, $\mathcal{U}$ and $\mathcal{N}(A)^\perp$ are also infinite dimensional.

Define a sequence $u_k \in \mathcal{N}(A)^\perp$ such that $\|u_j\|_{\mathcal{U}} = 1$ and $\langle u_k,\, u_j \rangle = \delta_{jk}$.

Compact operators are very common in inverse problems. This is a major source of ill-posedness.

Theorem 7

Let $A \in \mathcal{K}(\mathcal{U}, \mathcal{V})$ with infinite dimensional range. Then, the Moore-Penrose inverse of A is discontinuous.

Proof.

Since $\mathcal{R}(A)$ is infinite dimensional, $\mathcal{U}$ and $\mathcal{N}(A)^{\perp}$ are also infinite dimensional.

Define a sequence $u_k \in \mathcal{N}(A)^{\perp}$ such that $\|u_j\|_{\mathcal{U}} = 1$ and $\langle u_k,\, u_j \rangle = \delta_{jk}$.

Since A is compact, $f_j = Au_j$ has a convergent subsequence. So, for all $\delta > 0$, there exists j, k such that $\|f_j - f_k\|_{\mathcal{V}} \leqslant \delta$.

Compact operators are very common in inverse problems. This is a major source of ill-posedness.

Theorem 7

Let $A \in \mathcal{K}(\mathcal{U}, \mathcal{V})$ with infinite dimensional range. Then, the Moore-Penrose inverse of A is discontinuous.

Proof.

Since $\mathcal{R}(A)$ is infinite dimensional, $\mathcal{U}$ and $\mathcal{N}(A)^\perp$ are also infinite dimensional.

Define a sequence $u_k \in \mathcal{N}(A)^\perp$ such that $\|u_j\|_{\mathcal{U}} = 1$ and $\langle u_k,\, u_j\rangle = \delta_{jk}$.

Since A is compact, $f_j = Au_j$ has a convergent subsequence. So, for all $\delta > 0$, there exists j, k such that $\|f_j - f_k\|_{\mathcal{V}} \leqslant \delta$.

However,

$$\left\|A^\dagger f_j - A^\dagger f_k\right\|_{\mathcal{U}}^2 = \left\|A^\dagger A u_j - A^\dagger A u_k\right\|_{\mathcal{U}}^2 = \|u_j - u_k\|_{\mathcal{U}} = 2.$$

□

The spectral theorem: If $A \in \mathcal{K}(\mathcal{U}, \mathcal{U})$ is a self-adjoint operator on Hilbert space $\mathcal{U}$, then there exists an orthonormal basis $\{u_j\}_{j\in\mathbb{N}}$ of $\overline{\mathcal{R}(A)}$ and a sequence of eigenvalues $\{\lambda_j\}_j$ with $|\lambda_1| \geqslant |\lambda_2| \geqslant \cdots > 0$ and $\lambda_j \to 0$ such that for all $u \in \mathcal{U}$,

$$Au = \sum_{j=1}^{\infty} \lambda_j \langle u,\, u_j\rangle_{\mathcal{U}} u_j.$$

The spectral theorem: If $A \in \mathcal{K}(\mathcal{U}, \mathcal{U})$ is a self-adjoint operator on Hilbert space $\mathcal{U}$, then there exists an orthonormal basis $\{u_j\}_{j\in\mathbb{N}}$ of $\overline{\mathcal{R}(A)}$ and a sequence of eigenvalues $\{\lambda_j\}_j$ with $|\lambda_1| \geqslant |\lambda_2| \geqslant \cdots > 0$ and $\lambda_j \to 0$ such that for all $u \in \mathcal{U}$,

$$Au = \sum_{j=1}^{\infty} \lambda_j \langle u,\, u_j\rangle_{\mathcal{U}} u_j.$$

If A is not self-adjoint, then no eigenvalues (and hence no eigenvectors) need to exist. But, we can consider the eigenvalues of A^*A (which is self-adjoint and compact) to obtain a similar decomposition.

Theorem 8 (SVD of compact operators)

Let $A \in \mathcal{K}(\mathcal{U}, \mathcal{V})$. Then there exists $\sigma_1 \geqslant \sigma_2 \geqslant \cdots > 0$ and an orthonormal basis $\{u_j\}_j$ of $\mathcal{N}(A)^\perp$ and an orthonormal basis $\{v_j\}_j$ of $\overline{\mathcal{R}(A)}$ such that

$$Au_j = \sigma_j v_j \quad \text{and} \quad A^* v_j = \sigma_j u_j, \qquad \forall j \in \mathbb{N}.$$

For all $u \in \mathcal{U}$, we have the representation

$$Au = \sum_{j=1}^{\infty} \sigma_j \langle u,\, u_j \rangle v_j.$$

$\{(\sigma_j, u_j, v_j)\}_j$ is called a singular value decomposition of A. The adjoint is

$$A^* f = \sum_{j=1}^{\infty} \sigma_j \langle f,\, v_j \rangle u_j.$$

Proof of Theorem 8

Let $B = A^*A$. This is compact, self-adjoint and positive definite, so

$$Bu = \sum_j \sigma_j^2 \langle u,\, u_j \rangle x_j$$

where $\{u_j\}_j$ is an orthonormal bases of $\overline{\mathcal{R}(A^*A)}$. Define $v_j \stackrel{\text{def.}}{=} \frac{1}{\sigma_j} A u_j$.
Recall: $\overline{\mathcal{R}(A^*A)} = \overline{\mathcal{R}(A^*)} = \mathcal{N}(A)^\perp$.

Proof of Theorem 8

Let $B = A^*A$. This is compact, self-adjoint and positive definite, so

$$Bu = \sum_j \sigma_j^2 \langle u,\, u_j \rangle x_j$$

where $\{u_j\}_j$ is an orthonormal bases of $\overline{\mathcal{R}(A^*A)}$. Define $v_j \overset{\text{def.}}{=} \frac{1}{\sigma_j} A u_j$.
Recall: $\overline{\mathcal{R}(A^*A)} = \overline{\mathcal{R}(A^*)} = \mathcal{N}(A)^{\perp}$.

Clearly,

$$A^* v_k = \frac{1}{\sigma_j} A^* A u_j = \sigma_j u_j.$$

Proof of Theorem 8

Let $B = A^*A$. This is compact, self-adjoint and positive definite, so

$$Bu = \sum_j \sigma_j^2 \langle u,\, u_j \rangle x_j$$

where $\{u_j\}_j$ is an orthonormal bases of $\overline{\mathcal{R}(A^*A)}$. Define $v_j \overset{\text{def.}}{=} \frac{1}{\sigma_j} A u_j$.
Recall: $\overline{\mathcal{R}(A^*A)} = \overline{\mathcal{R}(A^*)} = \mathcal{N}(A)^\perp$.

$\{v_j\}_j$ is an orthonormal basis:

$$\langle v_j,\, v_k \rangle = \frac{1}{\sigma^2} \langle u_j,\, A^* A u_j \rangle = \delta_{jk}$$

Proof of Theorem 8

Let $B = A^*A$. This is compact, self-adjoint and positive definite, so

$$Bu = \sum_j \sigma_j^2 \langle u,\, u_j \rangle x_j$$

where $\{u_j\}_j$ is an orthonormal bases of $\overline{\mathcal{R}(A^*A)}$. Define $v_j \overset{\text{def.}}{=} \frac{1}{\sigma_j} A u_j$.
Recall: $\overline{\mathcal{R}(A^*A)} = \overline{\mathcal{R}(A^*)} = \mathcal{N}(A)^\perp$.

Extend $\{u_j\}_j$ to a basis of $\mathcal{U}$. Given $u \in \mathcal{U} = \mathcal{N}(A)^\perp \oplus \mathcal{N}(A)$,

$$Au = \sum_j \langle u,\, u_j \rangle A u_j = \sum_j \sigma_j \langle u,\, u_j \rangle v_j.$$

This also shows that $\{v_j\}_j$ is a basis of $\overline{\mathcal{R}(A)}$.
The spectral representation of A^*f is obtained similarly by extending $\{v_j\}_j$ to a basis of $\mathcal{V}$

SVD representation of the Moore-Penrose inverse

Theorem 9

Let $A \in \mathcal{K}(\mathcal{U}, \mathcal{V})$ with singular system $\{(\sigma_j, u_j, v_j)\}_j$ and $f \in \mathcal{D}(A^\dagger)$. Then,

(a) *$f \in \mathcal{D}(A^\dagger)$ if and only if the Picard condition is satisfied:*

$$\sum_{j=1}^{\infty} \frac{|\langle f,\, v_j \rangle|^2}{\sigma_j^2} < \infty.$$

(b) *If $f \in \mathcal{D}(A^\dagger)$, then $A^\dagger f = \sum_{j=1}^{\infty} \sigma_j^{-1} \langle f,\, v_j \rangle u_j$.*

SVD representation of the Moore-Penrose inverse

Theorem 9

Let $A \in \mathcal{K}(\mathcal{U}, \mathcal{V})$ *with singular system* $\{(\sigma_j, u_j, v_j)\}_j$ *and* $f \in \mathcal{D}(A^\dagger)$. *Then,*

(a) $f \in \mathcal{D}(A^\dagger)$ *if and only if the Picard condition is satisfied:*

$$\sum_{j=1}^{\infty} \frac{|\langle f,\, v_j\rangle|^2}{\sigma_j^2} < \infty.$$

(b) *If* $f \in \mathcal{D}(A^\dagger)$, *then* $A^\dagger f = \sum_{j=1}^{\infty} \sigma_j^{-1} \langle f,\, v_j\rangle u_j$.

Remarks:

The unboundedness of the Moore-Penrose inverse can clearly be seen from the SVD representation since $\left\|A^\dagger v_j\right\| = \sigma_j^{-1} \to \infty$, even though $\|v_j\| = 1$. In general, the series may not converge for a given $f \notin \mathcal{D}(A^\dagger)$.

Theorem 9

Let $A \in \mathcal{K}(\mathcal{U}, \mathcal{V})$ with singular system $\{(\sigma_j, u_j, v_j)\}_j$ and $f \in \mathcal{D}(A^\dagger)$. Then,

(a) $f \in \mathcal{D}(A^\dagger)$ *if and only if the Picard condition is satisfied:*

$$\sum_{j=1}^{\infty} \frac{|\langle f,\, v_j \rangle|^2}{\sigma_j^2} < \infty.$$

(b) *If* $f \in \mathcal{D}(A^\dagger)$, *then* $A^\dagger f = \sum_{j=1}^{\infty} \sigma_j^{-1} \langle f,\, v_j \rangle u_j$.

Remarks:

One can additionally show that if $f \in \overline{\mathcal{R}(A)}$, then $f \in \mathcal{R}(A)$ if and only if the Picard criterion is met.

Theorem 9

Let $A \in \mathcal{K}(\mathcal{U}, \mathcal{V})$ *with singular system* $\{(\sigma_j, u_j, v_j)\}_j$ *and* $f \in \mathcal{D}(A^\dagger)$. *Then,*

(a) $f \in \mathcal{D}(A^\dagger)$ *if and only if the Picard condition is satisfied:*

$$\sum_{j=1}^{\infty} \frac{|\langle f, v_j \rangle|^2}{\sigma_j^2} < \infty.$$

(b) *If* $f \in \mathcal{D}(A^\dagger)$, *then* $A^\dagger f = \sum_{j=1}^{\infty} \sigma_j^{-1} \langle f, v_j \rangle u_j$.

Remarks:

We say that an ill-posed inverse problem is mildly ill-posed if the singular values decay at most with polynomial speed. There exists $\gamma, C > 0$ such that $\sigma_j \geqslant Cj^{-\gamma}$. We say it is severely ill-posed if its singular values decay faster than polynomial speed, for all $\gamma, C > 0$, $\sigma_j \leqslant Cj^{-\gamma}$ for all j large enough.

Proof of (a) [Picard condition for $\mathcal{D}(A^\dagger)$]

- Recall that $\mathcal{D}(A^\dagger) = \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$. Suppose that $f = Au + w$ where $w \in \mathcal{R}(A)^\perp = \overline{\mathcal{R}(A)}^\perp$. Since $\{v_j\}_j$ is an ONB for $\overline{\mathcal{R}(A)}$,

  $$\langle f, v_j \rangle_{\mathcal{V}} = \langle Au, v_j \rangle_{\mathcal{V}} = \langle u, A^* v_j \rangle_{\mathcal{U}} = \sigma_j \langle u, u_j \rangle.$$

  Therefore, $\sum_j \sigma_j^{-2} \left| \langle f, v_j \rangle \right|^2 \leqslant \|u\|^2 < \infty$.

Proof of (a) [Picard condition for $\mathcal{D}(A^\dagger)$]

- Recall that $\mathcal{D}(A^\dagger) = \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$. Suppose that $f = Au + w$ where $w \in \mathcal{R}(A)^\perp = \overline{\mathcal{R}(A)}^\perp$. Since $\{v_j\}_j$ is an ONB for $\overline{\mathcal{R}(A)}$,

$$\langle f,\, v_j \rangle_{\mathcal{V}} = \langle Au,\, v_j \rangle_{\mathcal{V}} = \langle u,\, A^* v_j \rangle_{\mathcal{U}} = \sigma_j \langle u,\, u_j \rangle.$$

  Therefore, $\sum_j \sigma_j^{-2} \left| \langle f,\, v_j \rangle \right|^2 \leqslant \|u\|^2 < \infty$.

- Conversely, first note that we can write any $f \in \mathcal{V}$ as $f = f_1 + f_2$ with $f_1 \in \overline{\mathcal{R}(A)}$ and $f_2 \in \overline{\mathcal{R}(A)}^\perp$. If the Picard criterion hold, define

$$u \overset{\text{def.}}{=} \sum_j \sigma_j^{-1} \langle f,\, v_j \rangle u_j.$$

  Then, $Au = \sum_j \langle f,\, v_j \rangle v_j = P_{\overline{\mathcal{R}(A)}} f = f_1$. So, $f_1 \in \mathcal{R}(A)$ and $f \in \mathcal{D}(A^\dagger)$.

Proof of (b) [Spectral representation of $A^\dagger$]

We know that since $f \in \mathcal{D}(A^\dagger)$, $u^\dagger = A^\dagger f$ solves $A^*Au^\dagger = A^*f$.

Proof of (b) [Spectral representation of $A^\dagger$]

We know that since $f \in \mathcal{D}(A^\dagger)$, $u^\dagger = A^\dagger f$ solves $A^* A u^\dagger = A^* f$.

We just saw that

$$A^* A u^\dagger = \sum_j \sigma_j^2 \langle u^\dagger, u_j \rangle u_j \quad \text{and} \quad A^* f = \sum_j \sigma_j \langle f, v_j \rangle u_j.$$

So, $\sigma_j \langle u^\dagger, u_j \rangle = \langle f, v_j \rangle$.

Proof of (b) [Spectral representation of $A^\dagger$]

We know that since $f \in \mathcal{D}(A^\dagger)$, $u^\dagger = A^\dagger f$ solves $A^* A u^\dagger = A^* f$.

We just saw that

$$A^* A u^\dagger = \sum_j \sigma_j^2 \langle u^\dagger, u_j \rangle u_j \quad \text{and} \quad A^* f = \sum_j \sigma_j \langle f, v_j \rangle u_j.$$

So, $\sigma_j \langle u^\dagger, u_j \rangle = \langle f, v_j \rangle$.

Therefore, since $u^\dagger \in \overline{\mathcal{R}(A^*)}$,

$$u^\dagger = \sum_j \langle u^\dagger, u_j \rangle u_j = \sum_j \sigma_j^{-1} \langle f, v_j \rangle u_j = A^\dagger f.$$

Example

Back to differentiation example.

The forward operator is $A : L^2[0,1] \to L^2[0,1]$,

$$Au(t) = \int_0^t u(s)\mathrm{d}s = \int_0^1 K(s,t)u(s)\mathrm{d}s$$

where $K : [0,1]^2 \to \mathbb{R}$ is $K(s,t) = \begin{cases} 1 & s \leqslant t \\ 0 & \text{else.} \end{cases}$

A is a compact operator. The adjoint is

$$A^* f = \int_0^1 K(t,s) f(t)\mathrm{d}t = \int_s^1 f(t)\mathrm{d}t.$$

Example

The eigenvalues and eigenvectors of A^*A are such that

$$\sigma^2 x(s) = (A^*Ax)(s) = \int_s^1 \int_0^t x(r)\mathrm{d}r\mathrm{d}t.$$

- $x(1) = 0$ and $\sigma^2 x'(s) = -\int_0^s x(r)\mathrm{d}r$ so $x'(0) = 0$ and $\sigma^2 x''(s) = -x'(s)$.
- Solutions are of the form $x(s) = c_1 \sin(\sigma^{-1}s) + c_2 \cos(\sigma^{-1}s)$ for some constants c_1, c_2.
- To enforce the boundary conditions $x(1) = 0$ and $x'(0) = 0$, we see that $c_1 = 0$, $\sigma_j = \frac{2}{(2j-1)\pi}$ for $j \in \mathbb{N}$, and choosing $c_2 = \sqrt{2}$ gives the normalised eigenvectors are

  $$u_j(s) = \sqrt{2}\cos\left(\left(j - \frac{1}{2}\right)\pi s\right),$$

  and

  $$v_j(s) = \sigma_j^{-1}(Au_j)(s) = \sqrt{2}\sin\left((j - \frac{1}{2})\pi s\right).$$

Example

The Picard condition becomes

$$2\sum_{j=1}^{\infty}\sigma_j^{-2}\left(\int_0^1 f(s)\sin\left(\sigma_j^{-1}s\right)\mathrm{d}s\right)^2<\infty.$$

Expanding f in the basis $\{v_j\}$ gives

$$f(t)=\sum_{j=1}^{\infty}\left(\int_0^1 f(s)\sin(\sigma_j^{-1}s)\mathrm{d}s\right)\sin(\sigma_j^{-1}t)$$

and formally,

$$f'(t)=\sum_{j=1}^{\infty}\left(\sigma_j^{-1}\int_0^1 f(s)\sin(\sigma_j^{-1}s)\mathrm{d}s\right)\cos(\sigma_j^{-1}t)$$

The Picard condition is the condition for legitimacy of such differentiation.

From the decay of the singular values, this inverse problem is mildly ill-posed.

Summary

We looked at properties of least squares solutions

- The minimal norm solution exists when $f \in \mathcal{R}(A) + \mathcal{R}(A)^{\perp}$ and is unique. It is $A^{\dagger} f$.
- $A^{\dagger}$ is continuous iff $\mathcal{R}(A)$ is closed. For compact operators, this occurs only if $\mathcal{R}(A)$ is finite dimensional.
- For compact operators, we looked at the SVD representation of $A^{\dagger}$ – ill posedness is related to the decay of the singular values.

We looked at properties of least squares solutions

- The minimal norm solution exists when $f \in \mathcal{R}(A) + \mathcal{R}(A)^{\perp}$ and is unique. It is $A^{\dagger} f$.
- $A^{\dagger}$ is continuous iff $\mathcal{R}(A)$ is closed. For compact operators, this occurs only if $\mathcal{R}(A)$ is finite dimensional.
- For compact operators, we looked at the SVD representation of $A^{\dagger}$ – ill posedness is related to the decay of the singular values.

Takeway: $A^{\dagger} f$ is not a good solution in general! We need to regularize...