Generalized Inverses

$X \in \mathbb{C}^{n \times m}$ is a **generalized inverse** of $A \in \mathbb{C}^{m \times n}$ if it satisfies

$$AXA = A.$$

Theorem (Thm.1)

Let $X \in \mathbb{C}^{n \times m}$ be a generalized inverse of $A \in \mathbb{C}^{m \times n}$. Then $Ax = b$ has a solution iff $AXb = b$, in which case the general solution is

$$x = Xb + (I - XA)y,$$

where $y \in \mathbb{C}^{n}$ is an arbitrary vector.

Existence of Generalized Inverses

Let

$$EAP = \begin{bmatrix} I_r & K \\ O & O \end{bmatrix}, \quad r = \operatorname{rank}(A) \leq \min(n, m),$$

be the **reduced row echelon form** of $A \in \mathbb{C}^{m \times n}$ (see first year Linear Algebra course).

Any X given by

$$X = P \begin{bmatrix} I_r & O \\ O & L \end{bmatrix} E$$

for some $L \in \mathbb{C}^{(n-r)\times(m-r)}$ satisfies $AXA = A$.

- For every A there exist one or more generalized inverses.

Example 1: Determine a generalized inverse for

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

Let us find the reduced row echelon form of A:

Example 1: Determine a generalized inverse for

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

Let us find the reduced row echelon form of A:

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & -4 \\ 0 & 2 \end{bmatrix},$$

Example 1: Determine a generalized inverse for

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

Let us find the reduced row echelon form of A:

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & -4 \\ 0 & 2 \end{bmatrix},$$

$$\begin{bmatrix} 1 & 1/2 & 0 \\ 0 & -1/4 & 0 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & -4 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix},$$

Example 1: Determine a generalized inverse for

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

Let us find the reduced row echelon form of A:

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & -4 \\ 0 & 2 \end{bmatrix},$$

$$\begin{bmatrix} 1 & 1/2 & 0 \\ 0 & -1/4 & 0 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & -4 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix},$$

so $E = \begin{bmatrix} 1 & 1/2 & 0 \\ 0 & -1/4 & 0 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1/2 & 0 \\ 1/2 & -1/4 & 0 \\ -2 & 1 & 2 \end{bmatrix},$

and $P = I_2$.

Hence,

$$X = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1/2 & 0 \\ 1/2 & -1/4 & 0 \\ -2 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 1/2 & 0 \\ 1/2 & -1/4 & 0 \end{bmatrix}$$

is a generalized inverse of A. Check that $AXA = A$.

The Moore–Penrose Generalized Inverse

Moore–Penrose generalized inverse of $A \in \mathbb{C}^{m \times n}$: unique $X \in \mathbb{C}^{n \times m}$ satisfying the four Moore–Penrose conditions

$$\text{(i)}\ \ AXA = A, \qquad \text{(ii)}\ \ XAX = X,$$
$$\text{(iii)}\ \ AX = (AX)^*, \qquad \text{(iv)}\ \ XA = (XA)^*.$$

The Moore–Penrose Generalized Inverse

Moore–Penrose generalized inverse of $A \in \mathbb{C}^{m \times n}$: unique $X \in \mathbb{C}^{n \times m}$ satisfying the four Moore–Penrose conditions

$$\text{(i)}\ \ AXA = A, \qquad \text{(ii)}\ \ XAX = X,$$
$$\text{(iii)}\ \ AX = (AX)^*, \qquad \text{(iv)}\ \ XA = (XA)^*.$$

It is

- a generalized inverse by (i),
- commonly called the **pseudoinverse** of A,
- denoted by A^+.

Show existence of A^+ via singular value decomposition of A.

The Moore–Penrose Generalized Inverse

Moore–Penrose generalized inverse of $A \in \mathbb{C}^{m\times n}$: unique $X \in \mathbb{C}^{n\times m}$ satisfying the four Moore–Penrose conditions

$$\text{(i)}\ \ AXA = A, \qquad \text{(ii)}\ \ XAX = X,$$
$$\text{(iii)}\ \ AX = (AX)^*, \qquad \text{(iv)}\ \ XA = (XA)^*.$$

It is

- a generalized inverse by (i),
- commonly called the **pseudoinverse** of A,
- denoted by A^+.

Show existence of A^+ via singular value decomposition of A.

Does A^+ always exist?

Singular Value Decomposition

Theorem (Singular value decomposition, Thm. 2)

$A \in \mathbb{C}^{m \times n}$ has a singular value decomposition (SVD)

$$A = U \Sigma V^*,$$

where $U \in \mathbb{C}^{m \times m}$, $V \in \mathbb{C}^{n \times n}$ are unitary and

$$\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_p) \in \mathbb{R}^{m \times n},$$

$p = \min(m, n)$, where

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0.$$

If A is real, U and V can be taken real orthogonal.

Proof

Let $x \in \mathbb{C}^n$, $y \in \mathbb{C}^m$ s.t. $\|x\|_2 = \|y\|_2 = 1$ & $Ax = \|A\|_2 y = \sigma y$.
Let $V = [x, V_1] \in \mathbb{C}^{n \times n}$ and $U = [y, U_1] \in \mathbb{C}^{m \times m}$ be unitary.

Proof

Let $x \in \mathbb{C}^n$, $y \in \mathbb{C}^m$ s.t. $\|x\|_2 = \|y\|_2 = 1$ & $Ax = \|A\|_2 y = \sigma y$.
Let $V = [x, V_1] \in \mathbb{C}^{n \times n}$ and $U = [y, U_1] \in \mathbb{C}^{m \times m}$ be unitary.

$$U^* A V = \begin{bmatrix} y^* \\ U_1^* \end{bmatrix} A \begin{bmatrix} x & V_1 \end{bmatrix} = \begin{bmatrix} y^* A x & y^* A V_1 \\ U_1^* A x & U_1^* A V_1 \end{bmatrix} =: \begin{bmatrix} \sigma & w^* \\ 0 & B \end{bmatrix} = A_1,$$

Proof

Let $x \in \mathbb{C}^n$, $y \in \mathbb{C}^m$ s.t. $\|x\|_2 = \|y\|_2 = 1$ & $Ax = \|A\|_2 y = \sigma y$.
Let $V = [x, V_1] \in \mathbb{C}^{n \times n}$ and $U = [y, U_1] \in \mathbb{C}^{m \times m}$ be unitary.

$$U^* A V = \begin{bmatrix} y^* \\ U_1^* \end{bmatrix} A \begin{bmatrix} x & V_1 \end{bmatrix} = \begin{bmatrix} y^* A x & y^* A V_1 \\ U_1^* A x & U_1^* A V_1 \end{bmatrix} =: \begin{bmatrix} \sigma & w^* \\ 0 & B \end{bmatrix} = A_1,$$

$$A_1 \begin{bmatrix} \sigma \\ w \end{bmatrix} = \begin{bmatrix} \sigma & w^* \\ 0 & B \end{bmatrix} \begin{bmatrix} \sigma \\ w \end{bmatrix} = \begin{bmatrix} \sigma^2 + w^* w \\ B w \end{bmatrix}$$

Proof

Let $x \in \mathbb{C}^n$, $y \in \mathbb{C}^m$ s.t. $\|x\|_2 = \|y\|_2 = 1$ & $Ax = \|A\|_2 y = \sigma y$.
Let $V = [x, V_1] \in \mathbb{C}^{n\times n}$ and $U = [y, U_1] \in \mathbb{C}^{m\times m}$ be unitary.

$$U^* AV = \begin{bmatrix} y^* \\ U_1^* \end{bmatrix} A \begin{bmatrix} x & V_1 \end{bmatrix} = \begin{bmatrix} y^* Ax & y^* AV_1 \\ U_1^* Ax & U_1^* AV_1 \end{bmatrix} =: \begin{bmatrix} \sigma & w^* \\ 0 & B \end{bmatrix} = A_1,$$

$$A_1 \begin{bmatrix} \sigma \\ w \end{bmatrix} = \begin{bmatrix} \sigma & w^* \\ 0 & B \end{bmatrix} \begin{bmatrix} \sigma \\ w \end{bmatrix} = \begin{bmatrix} \sigma^2 + w^* w \\ Bw \end{bmatrix}$$

so

$$\sigma^2 + w^* w \leq \left\| \begin{bmatrix} \sigma^2 + w^* w \\ Bw \end{bmatrix} \right\|_2 = \left\| A_1 \begin{bmatrix} \sigma \\ w \end{bmatrix} \right\|_2 \leq \|A_1\|_2 (\sigma^2 + w^* w)^{1/2}.$$

Hence, $\|A_1\|_2 \geq (\sigma^2 + w^* w)^{1/2}$.

Proof

Let $x \in \mathbb{C}^n$, $y \in \mathbb{C}^m$ s.t. $\|x\|_2 = \|y\|_2 = 1$ & $Ax = \|A\|_2 y = \sigma y$.
Let $V = [x, V_1] \in \mathbb{C}^{n\times n}$ and $U = [y, U_1] \in \mathbb{C}^{m\times m}$ be unitary.

$$U^* A V = \begin{bmatrix} y^* \\ U_1^* \end{bmatrix} A \begin{bmatrix} x & V_1 \end{bmatrix} = \begin{bmatrix} y^* A x & y^* A V_1 \\ U_1^* A x & U_1^* A V_1 \end{bmatrix} =: \begin{bmatrix} \sigma & w^* \\ 0 & B \end{bmatrix} = A_1,$$

$$A_1 \begin{bmatrix} \sigma \\ w \end{bmatrix} = \begin{bmatrix} \sigma & w^* \\ 0 & B \end{bmatrix} \begin{bmatrix} \sigma \\ w \end{bmatrix} = \begin{bmatrix} \sigma^2 + w^* w \\ Bw \end{bmatrix}$$

so

$$\sigma^2 + w^* w \le \left\| \begin{bmatrix} \sigma^2 + w^* w \\ Bw \end{bmatrix} \right\|_2 = \left\| A_1 \begin{bmatrix} \sigma \\ w \end{bmatrix} \right\|_2 \le \|A_1\|_2 (\sigma^2 + w^* w)^{1/2}.$$

Hence, $\|A_1\|_2 \ge (\sigma^2 + w^* w)^{1/2}$. But,
$\|A_1\|_2 = \|U^* A V\|_2 = \|A\|_2 = \sigma$. This implies that $w = 0$.
The proof is completed by the obvious induction.

$A = U\Sigma V^* \in \mathbb{C}^{m\times n}$, with $U \in \mathbb{C}^{m\times m}$, $V \in \mathbb{C}^{n\times n}$ unitary and $\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_p) \in \mathbb{R}^{m\times n}$, $p = \min(m, n)$,

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_p = 0.$$

- σ_i, $i = 1, \ldots, p$ are **singular values** of A.
- The u_i and v_i are **left and right singular vectors** of A, respectively.

$A = U\Sigma V^* \in \mathbb{C}^{m\times n}$, with $U \in \mathbb{C}^{m\times m}$, $V \in \mathbb{C}^{n\times n}$ unitary and $\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_p) \in \mathbb{R}^{m\times n}$, $p = \min(m, n)$,

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_p = 0.$$

- σ_i, $i = 1, \ldots, p$ are **singular values** of A.
- The u_i and v_i are **left and right singular vectors** of A, respectively.

The nonzero singular values of A are the positive square roots of the nonzero e'vals of AA^* or A^*A.

The left singular vectors u_i are e'vecs of AA^* and the right singular vectors v_i are e'vecs of A^*A.

$A = U\Sigma V^* \in \mathbb{C}^{m\times n}$, with $U \in \mathbb{C}^{m\times m}$, $V \in \mathbb{C}^{n\times n}$ unitary and $\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_p) \in \mathbb{R}^{m\times n}$, $p = \min(m, n)$,

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_p = 0.$$

- σ_i, $i = 1, \ldots, p$ are **singular values** of A.
- The u_i and v_i are **left and right singular vectors** of A, respectively.

$$\begin{aligned}
\mathrm{rank}(A) &= r, \\
\mathrm{null}(A) &= \mathrm{span}\{v_{r+1}, \ldots, v_n\}, \\
\mathrm{range}(A) &= \mathrm{span}\{u_1, u_2, \ldots, u_r\}, \\
A &= \sum_{i=1}^{r} \sigma_i u_i v_i^*, \\
Av_i = \sigma_i u_i, &\quad A^* u_i = \sigma_i v_i,\ i = 1, \ldots, r.
\end{aligned}$$

Example

Compute the singular value decomposition of

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

- The eigenvalues of $A^T A = \begin{bmatrix} 5 & 2 \\ 2 & 8 \end{bmatrix}$ are 9 and 4 so the singular values of A are 3 and 2.
- Normalized eigenvectors of $A^T A$ are $v_1 = \frac{1}{\sqrt{5}} \left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix} \right]$, $v_2 = \frac{1}{\sqrt{5}} \left[\begin{smallmatrix} 2 \\ -1 \end{smallmatrix} \right]$.
- $u_1 = \frac{1}{3} A v_1 = \frac{1}{3\sqrt{5}} \begin{bmatrix} 5 \\ 2 \\ 4 \end{bmatrix}$, $u_2 = \frac{1}{2} A v_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix}$.

- We need a third orthogonal vector u_3: Application of the Gram–Schmidt process to u_1, u_2 and e_1 produces

$$u_3 = \frac{e_1 - (\sum_{i=1}^{2} e_1^T u_i) u_i}{\| e_1 - (\sum_{i=1}^{2} e_1^T u_i) u_i \|_2} = \begin{bmatrix} 2/3 \\ -1/3 \\ -2/3 \end{bmatrix}.$$

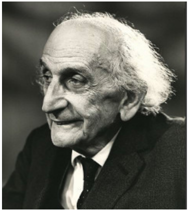
- A has the SVD $A = U \Sigma V^T$ where

$$U = \frac{1}{3\sqrt{5}} \begin{bmatrix} 5 & 0 & 2\sqrt{5} \\ 2 & 6 & -\sqrt{5} \\ 4 & -3 & -2\sqrt{5} \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 3 & 0 \\ 0 & 2 \\ 0 & 0 \end{bmatrix},$$

$$V = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}.$$

Cornelius Lanczos

- Hungarian mathematician and physicist (1893–1974)
- work in relativity theory, 1928/29 assistent of Einstein
- emigrated to the US in 1931, professor at Purdue U
- in 1938 published his first Numerical Analysis paper
- in 1944 started working for Boeing Aircraft Company (eigenvalue computations)
- in 1949 moved to National Bureau of Standards in LA
- in 1952 moved to Dublin Institute for Advance Study
- in 1972 gave guest lectures at UMIST, which now is The University of Manchester!



Existence of the Moore–Penrose Inverse

Recall that $X^+ \in \mathbb{C}^{n\times m}$ is the **pseudoinverse** of $X \in \mathbb{C}^{m\times n}$ if it satisfies the four Moore–Penrose conditions:

$$\text{(i)}\ \ XX^+X = X, \qquad \text{(ii)}\ \ X^+XX^+ = X^+,$$
$$\text{(iii)}\ \ XX^+ = (XX^+)^*, \qquad \text{(iv)}\ \ X^+X = (X^+X)^*.$$

Existence of the Moore–Penrose Inverse

Recall that $X^+ \in \mathbb{C}^{n\times m}$ is the **pseudoinverse** of $X \in \mathbb{C}^{m\times n}$ if it satisfies the four Moore–Penrose conditions:

$$\text{(i)}\ \ XX^+X = X, \qquad \text{(ii)}\ \ X^+XX^+ = X^+,$$
$$\text{(iii)}\ \ XX^+ = (XX^+)^*, \qquad \text{(iv)}\ \ X^+X = (X^+X)^*.$$

Theorem (Thm. 3)

If $A = U\Sigma V^* \in \mathbb{C}^{m\times n}$ is an SVD then $A^+ = V\Sigma^+U^*$, where

$$\Sigma^+ = \mathrm{diag}(\sigma_1^{-1}, \ldots, \sigma_r^{-1}, 0, \ldots, 0) \in \mathbb{R}^{n\times m}$$

and $r = \mathrm{rank}(A)$.

Existence of the Moore–Penrose Inverse

Recall that $X^+ \in \mathbb{C}^{n\times m}$ is the **pseudoinverse** of $X \in \mathbb{C}^{m\times n}$ if it satisfies the four Moore–Penrose conditions:

$$\text{(i)}\ \ XX^+X = X, \qquad \text{(ii)}\ \ X^+XX^+ = X^+,$$
$$\text{(iii)}\ \ XX^+ = (XX^+)^*, \qquad \text{(iv)}\ \ X^+X = (X^+X)^*.$$

Theorem (Thm. 3)

If $A = U\Sigma V^* \in \mathbb{C}^{m\times n}$ is an SVD then $A^+ = V\Sigma^+U^*$, where

$$\Sigma^+ = \mathrm{diag}(\sigma_1^{-1}, \dots, \sigma_r^{-1}, 0, \dots, 0) \in \mathbb{R}^{n\times m}$$

and $r = \mathrm{rank}(A)$.

In general, $(AB)^+ \neq B^+A^+$.

Low Rank Approximation

$A = U\Sigma V^* \in \mathbb{C}^{m\times n}$, with $U = [u_1, \dots, u_m] \in \mathbb{C}^{m\times m}$, $V = [v_1, \dots, v_n] \in \mathbb{C}^{n\times n}$ are unitary and $\Sigma = \mathrm{diag}(\sigma_1, \dots, \sigma_p) \in \mathbb{R}^{m\times n}$, $p = \min(n, m)$.

Theorem

If $\mathrm{rank}(A) > k$ *and* $A_k = \sum_{i=1}^k \sigma_i u_i v_i^*$ *then*

$$\min_{\mathrm{rank}(B)=k} \|A - B\|_2 = \|A - A_k\|_2 = \sigma_{k+1}.$$

The singular values indicate how near a given matrix is to a matrix of low rank.

Projectors

Let $\mathcal{S} \subset \mathbb{C}^m$ and let $P_{\mathcal{S}} \in \mathbb{C}^{m\times m}$.

- $P_{\mathcal{S}}$ is the **projector** onto $\mathcal{S}$ if
  - $\mathrm{range}(P_{\mathcal{S}}) = \mathcal{S}$ and,
  - $P_{\mathcal{S}}^2 = P_{\mathcal{S}}$.

  The projector is not unique.

Projectors

Let $\mathcal{S} \subset \mathbb{C}^m$ and let $P_\mathcal{S} \in \mathbb{C}^{m \times m}$.

- $P_\mathcal{S}$ is the **projector** onto $\mathcal{S}$ if
  - $\mathrm{range}(P_\mathcal{S}) = \mathcal{S}$ and,
  - $P_\mathcal{S}^2 = P_\mathcal{S}$.

  The projector is not unique.

- $P_\mathcal{S}$ is the **orthogonal projector** onto $\mathcal{S}$ if
  - $\mathrm{range}(P_\mathcal{S}) = \mathcal{S}$,
  - $P_\mathcal{S}^2 = P_\mathcal{S}$, and
  - $P_\mathcal{S}^* = P_\mathcal{S}$.

  The orthogonal projector is unique (see Exercise 6).

Projectors

Let $\mathcal{S} \subset \mathbb{C}^m$ and let $P_\mathcal{S} \in \mathbb{C}^{m \times m}$.

- $P_\mathcal{S}$ is the **projector** onto $\mathcal{S}$ if
  - $\mathrm{range}(P_\mathcal{S}) = \mathcal{S}$ and,
  - $P_\mathcal{S}^2 = P_\mathcal{S}$.

  The projector is not unique.

- $P_\mathcal{S}$ is the **orthogonal projector** onto $\mathcal{S}$ if
  - $\mathrm{range}(P_\mathcal{S}) = \mathcal{S}$,
  - $P_\mathcal{S}^2 = P_\mathcal{S}$, and
  - $P_\mathcal{S}^* = P_\mathcal{S}$.

  The orthogonal projector is unique (see Exercise 6).

Also, $P_{\mathcal{S}^\perp} = I_m - P_\mathcal{S}$ is the orthogonal projector onto the orthogonal complement of $\mathcal{S}$ in $\mathbb{C}^m$ denoted by $\mathcal{S}^\perp$.

Projectors onto Fundamental Subspaces

$$A = U\Sigma V^* \in \mathbb{C}^{m\times n} \text{ with } U = [u_1, \ldots, u_m] = [\underbrace{U_1}_{m\times r}\; U_2],$$

$$V = [v_1, \ldots, v_n] = [\underbrace{V_1}_{n\times r}\; V_2] \text{ unitary, } r := \mathrm{rank}(A).$$

Projectors onto Fundamental Subspaces

$A = U\Sigma V^* \in \mathbb{C}^{m\times n}$ with $U = [u_1, \ldots, u_m] = [\underbrace{U_1}_{m\times r}\ U_2]$,

$V = [v_1, \ldots, v_n] = [\underbrace{V_1}_{n\times r}\ V_2]$ unitary, $r := \mathrm{rank}(A)$.

Recall that

$$Av_i = \begin{cases} \sigma_i u_i, & i = 1, \ldots, r \\ 0, & i = r+1, \ldots, n, \end{cases} \quad A^* u_i = \begin{cases} \sigma_i v_i, & i = 1, \ldots, r, \\ 0, & i = r+1, \ldots, n. \end{cases}$$

Projectors onto Fundamental Subspaces

$A = U\Sigma V^* \in \mathbb{C}^{m\times n}$ with $U = [u_1, \dots, u_m] = [\underbrace{U_1}_{m\times r}\ U_2]$,

$V = [v_1, \dots, v_n] = [\underbrace{V_1}_{n\times r}\ V_2]$ unitary, $r := \operatorname{rank}(A)$.

Recall that

$$Av_i = \begin{cases} \sigma_i u_i, & i = 1, \dots, r \\ 0, & i = r+1, \dots, n, \end{cases} \quad A^* u_i = \begin{cases} \sigma_i v_i, & i = 1, \dots, r, \\ 0, & i = r+1, \dots, n. \end{cases}$$

$\operatorname{range}(A), \operatorname{null}(A), \operatorname{range}(A^*), \operatorname{null}(A^*)$ are the **four fundamental subspaces of A**:

Projectors onto Fundamental Subspaces

$A = U\Sigma V^* \in \mathbb{C}^{m\times n}$ with $U = [u_1, \ldots, u_m] = [\underbrace{U_1}_{m\times r}\ U_2]$,

$V = [v_1, \ldots, v_n] = [\underbrace{V_1}_{n\times r}\ V_2]$ unitary, $r := \mathrm{rank}(A)$.

Recall that

$$Av_i = \begin{cases} \sigma_i u_i, & i = 1, \ldots, r \\ 0, & i = r+1, \ldots, n, \end{cases} \quad A^* u_i = \begin{cases} \sigma_i v_i, & i = 1, \ldots, r, \\ 0, & i = r+1, \ldots, n. \end{cases}$$

$\mathrm{range}(A), \mathrm{null}(A), \mathrm{range}(A^*), \mathrm{null}(A^*)$ are the **four fundamental subspaces of A**:

Orthogonal projectors onto these subspaces:

$$P_{\mathrm{range}(A)} = U_1 U_1^* = AA^+, \qquad P_{\mathrm{null}(A^*)} = U_2 U_2^* = I - AA^+,$$

$$P_{\mathrm{range}(A^*)} = V_1 V_1^* = A^+ A, \qquad P_{\mathrm{null}(A)} = V_2 V_2^* = I - A^+ A.$$

Least Squares Problems

Let $Ax = b$, where $A \in \mathbb{C}^{m\times n}$ and $b \in \mathbb{C}^m$ are given.

- Theorem 1 $\Rightarrow \exists$ solution x iff $AA^+b = b$ in this case general solution is $x = A^+b + (I - A^+A)y$ for any $y \in \mathbb{C}^n$.

Note that $AA^+b = b \Leftrightarrow b \in \mathrm{range}(A)$.

Least Squares Problems

Let $Ax = b$, where $A \in \mathbb{C}^{m \times n}$ and $b \in \mathbb{C}^m$ are given.

- Theorem 1 $\Rightarrow \exists$ solution x iff $AA^+b = b$ in this case general solution is $x = A^+b + (I - A^+A)y$ for any $y \in \mathbb{C}^n$.

Note that $AA^+b = b \Leftrightarrow b \in \mathrm{range}(A)$.

Theorem (Minimum 2-norm solution, Thm. 5)

For given $A \in \mathbb{C}^{m \times n}$ and $b \in \mathbb{C}^m$ with $b \in \mathrm{range}(A)$, the vector $x = A^+b$ is the solution of minimum 2-norm amongst all the solutions to $Ax = b$.

Least Squares Problems

Let $Ax = b$, where $A \in \mathbb{C}^{m \times n}$ and $b \in \mathbb{C}^m$ are given.

- Theorem 1 $\Rightarrow \exists$ solution x iff $AA^+b = b$ in this case general solution is $x = A^+b + (I - A^+A)y$ for any $y \in \mathbb{C}^n$.

Note that $AA^+b = b \Leftrightarrow b \in \mathrm{range}(A)$.

Theorem (Minimum 2-norm solution, Thm. 5)

For given $A \in \mathbb{C}^{m \times n}$ and $b \in \mathbb{C}^m$ with $b \in \mathrm{range}(A)$, the vector $x = A^+b$ is the solution of minimum 2-norm amongst all the solutions to $Ax = b$.

Thm 1 $\Rightarrow$ no solutions when $b \notin \mathrm{range}(A)$.

Instead consider a **least squares solution**.

Least squares solutions

A **least squares solution** to $Ax = b$ is a vector $x \in \mathbb{C}^n$ s.t. $\|Ax - b\|_2$ is minimized.

Theorem (Thm. 6)

For given $A \in \mathbb{C}^{m \times n}$ and $b \in \mathbb{C}^m$ the vectors

$$x = A^+ b + (I - A^+ A)y, \quad y \in \mathbb{C}^n \text{ arbitrary},$$

minimize $\|Ax - b\|_2$. Moreover

$$x_{LS} = A^+ b$$

*is the **least squares solution of minimum 2-norm.***

Polar Decomposition

Theorem (Thm. 7)

Any $A \in \mathbb{C}^{m\times n}$ with $m \geq n$ can be represented in the form

$$A = QH,$$

where $Q \in \mathbb{C}^{m\times n}$ has orthonormal columns and $H \in \mathbb{C}^{n\times n}$ is Hermitian positive semidefinite.

Polar Decomposition

Theorem (Thm. 7)

Any $A \in \mathbb{C}^{m\times n}$ with $m \geq n$ can be represented in the form

$$A = QH,$$

where $Q \in \mathbb{C}^{m\times n}$ has orthonormal columns and $H \in \mathbb{C}^{n\times n}$ is Hermitian positive semidefinite.

Proof: (not examinable)

$$A = U\Sigma V^* = \underbrace{U \begin{bmatrix} I_r & 0 \\ 0 & I_{m-r,n-r} \end{bmatrix} V^*}_{\text{orthonormal cols}} \cdot \underbrace{V \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0_{n-r} \end{bmatrix} V^*}_{\text{Hermitian pos. semidef.}}$$

Example 6: Comparing two objects

Molecule A

Molecule B

Define $A = [\, a_1 \quad \dots \quad a_7 \,], B = [\, b_1 \quad \dots \quad b_7 \,] \in \mathbb{R}^{3\times 7}$ with a_i, b_i: coordinates of centers of sphere i.

Is molecule A the same as molecule B?

Orthogonal Procrustes problem

- Is A obtained by “rotating B", i.e., is $A = QB$ for some orthogonal matrix Q?
- Cannot expect exact equality because collected data are not exact.
- solve instead

$$\text{minimize } \|A - QB\|_F \text{ subject to } Q^T Q = I.$$

Procrustes: In the Greek myth, Procrustes was a son of Poseidon. He had an iron bed, in which he invited every passer-by to spend the night, and where he set to work on them with his smith's hammer, to stretch them to fit.

If the guest proved too tall, Procrustes would amputate the excess length. Nobody ever fit the bed exactly, because secretly Procrustes had two beds.

Procrustes continued his reign of terror until he was captured by Theseus, who "fitted" Procrustes to his own bed.