Generalized Inverse and Pseudoinverse

(Background for Robot Control Part III: Null Space Projection)

Yan-Bin Jia

Department of Computer Science

Iowa State University

Ames, IA 50011

Facts about Matrix Product

Given matrices $A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{n\times r}$,

$\text{col}(AB) \subseteq \text{col}(A)$ since the ith column of AB is $b_{1i}\boldsymbol{a}_1 + b_{2i}\boldsymbol{a}_2 + \cdots + b_{ni}\boldsymbol{a}_n$.

$$(\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n)\begin{pmatrix} & b_{1i} & \\ & b_{2i} & \\ \vdots & \vdots & \vdots \\ & b_{ni} & \end{pmatrix}$$

$\text{row}(AB) \subseteq \text{row}(B)$ since the ith row of AB is $a_{i1}\boldsymbol{b}_1 + a_{i2}\boldsymbol{b}_2 + \cdots + a_{in}\boldsymbol{b}_n$.

$$\begin{pmatrix} & & \cdots & \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ & & \cdots & \end{pmatrix}\begin{pmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \\ \vdots \\ \boldsymbol{b}_n \end{pmatrix}$$

$\text{rank}\,(AB) \leq \min\{\text{rank}\,(A), \text{rank}\,(B)\}$

$\text{rank}\,(AB) = \text{rank}\,(B)$ if $m = n$ and A nonsingular

since it also holds that
$\text{rank}\,(B) = \text{rank}\,\big(A^{-1}(AB)\big)$
$\leq \text{rank}(AB)$

$\text{rank}\,(AB) = \text{rank}\,(A)$ if $n = r$ and B nonsingular

Generalized Inverse

Given a matrix $A \in \mathbb{R}^{m\times n}$, a matrix $A^\# \in \mathbb{R}^{n\times m}$ is called a *generalized inverse* or *g-inverse* of A if $AA^\# A = A$.

- $A^\#$ always exists (by construction of one below).
  - (Special case) $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ has rank r with non-singular $A_{11} \in \mathbb{R}^{r\times r}$.

The last $m - r$ rows are linear combinations of the first r rows.

$\Downarrow$

$(A_{21}, A_{22}) = B(A_{11}, A_{12})$ for some $B \in \mathbb{R}^{(m-r)\times r}$.

$A_{21} = BA_{11}$ $\qquad$ $A_{22} = BA_{12}$

$\Downarrow$

$B = A_{21}A_{11}^{-1}$ $\Longrightarrow$ $A_{22} = A_{21}A_{11}{}^{-1}A_{12}$

$\Downarrow$

$\begin{pmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix}$ is a generalized inverse of A.

Example 1 See [1], p. 10.

Existence and Properties

- (General case) Every matrix $A' \in \mathbb{R}^{m\times n}$ with rank r can have its rows and columns permuted to become $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ such that $A_{11} \in \mathbb{R}^{r\times r}$ has rank r. Namely, $A = PA'Q$, where P and Q are some row and column permutation matrices.

  $(A')^{\#} = Q \begin{pmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix} P$ is a generalized inverses of A.

- $A^{\#}$ may not be unique.

  All the generalized inverses of $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ are in the form $(1-2a, a)$.

- That $AA^{\#}A = A$ implies $\text{rank}\left(A^{\#}\right) \geq \text{rank}\left(AA^{\#}A\right) = \text{rank}(A)$.

- If $m = n$ and A is invertible, then $A^{\#}$ is unique and $A^{\#} = A^{-1}$.

Column and Null Spaces

- $\mathbb{R}^n = \text{null}(A) \oplus \text{col}(A^\#)$.

  We also know $\mathbb{R}^n = \text{null}(A) \oplus \text{row}(A)$.

  $\Longrightarrow$ $\text{row}(A) = \text{col}(A^\#)$

- $\text{col}(A) \cap \text{null}\left(A^\#\right) = \emptyset$.

  We also know $\mathbb{R}^m = \text{null}\left(A^\#\right) \oplus \text{row}(A^\#)$.

  $\Longrightarrow$ $\text{col}(A) \subseteq \text{row}\left(A^\#\right)$

♣ These two properties are listed in Proposition 2 of [2] with proofs given in [3].

Linear System and Transpose

- We know that the linear system $A\boldsymbol{x} = \boldsymbol{b}$ has a solution if and only if $\boldsymbol{b} \in \text{col}(A)$. In this case, $\boldsymbol{x} = A^{\#}\boldsymbol{b}$ is a particular solution since

$$\begin{aligned} A\left(A^{\#}\boldsymbol{b}\right) &= AA^{\#}\boldsymbol{b} \\ &= AA^{\#}(A\boldsymbol{x}) \\ &= \left(AA^{\#}A\right)\boldsymbol{x} \\ &= A\boldsymbol{x} \\ &= \boldsymbol{b}. \end{aligned}$$

Example 2 See [1], p. 12.

- Suppose $\boldsymbol{b} \in \text{col}(A)$. The set of solutions to the linear system is

$$\left\{A^{\#}\boldsymbol{b} + \left(I_n - A^{\#}A\right)\boldsymbol{\omega} \mid \boldsymbol{\omega} \in \mathbb{R}^{\boldsymbol{n}}\right\}$$

Transpose and Projection Matrix

- ${A^{\#}}^T$ is a generalized inverse of A^T since

$$A^T {A^{\#}}^T A^T = \left(AA^{\#}A\right)^T = A^T.$$

Namely, generalized inverse commutes with transpose.

A square matrix P is a *projection matrix* if $P^2 = P$.

Any vector $\boldsymbol{v} \in \mathrm{col}(P)$ is preserved by P.

$\boldsymbol{v} = P\boldsymbol{x}$ for some $\boldsymbol{x}$ $\Longrightarrow$ $P\boldsymbol{v} = P(P\boldsymbol{x}) = P^2\boldsymbol{x} = P\boldsymbol{x} = \boldsymbol{v}$

Projector $AA^\#$

- $\text{col}(AA^\#) = \text{col}(A)$

Proof We already knew that $\text{col}(AA^\#) \subseteq \text{col}(A)$.

$$\boldsymbol{y} \in \text{col}(A) \Longrightarrow \boldsymbol{y} = A\boldsymbol{x} \text{ for some } \boldsymbol{x}$$
$$\Longrightarrow \boldsymbol{y} = AA^\# A\boldsymbol{x} = (AA^\#)\boldsymbol{y} \in \text{col}(AA^\#)$$
$$\Longrightarrow \text{col}(A) \subseteq \text{col}(AA^\#) \qquad \square$$

- $AA^\#$ is a projection matrix onto $\text{col}(A)$.

$$AA^\# = (AA^\# A)A^\# = (AA^\#)(AA^\#) \Longrightarrow AA^\# \text{ is a projection matrix.}$$
$$AA^\#\boldsymbol{x} \in \text{col}(AA^\#) = \text{col}(A)$$

- $\mathbb{R}^m = \text{col}(A) \oplus \text{null}(AA^\#) \Longrightarrow \text{col}(A) = \text{row}(AA^\#) \subseteq \text{row}(A^\#)$

♣ The last property is stated in Proposition 2 of [2] with a proof given in [3].

Projector $A^{\#}A$

- $A^{\#}A$ is a projection matrix.

$$A^{\#}A = A^{\#}(AA^{\#}A) = (A^{\#}A)(A^{\#}A)$$

- $\text{row}(A^{\#}A) = \text{row}(A)$

Proof Clearly, $\text{col}(AA^{\#}) \subseteq \text{col}(A)$.

$$\boldsymbol{y} \in \text{row}(A) \Longrightarrow \boldsymbol{y} = A^T\boldsymbol{x} \text{ for some } \boldsymbol{x}$$

$$\Longrightarrow \boldsymbol{y} = (AA^{\#}A)^T\boldsymbol{x} = (A^{\#}A)^T A^T\boldsymbol{y} \in \text{row}(A^{\#}A)$$

$$\Longrightarrow \text{row}(A) \subseteq \text{row}(A^{\#}A)$$ □

$$\mathbb{R}^n = \text{row}(A) \oplus \text{null}(A)$$
$$= \text{row}(A^{\#}A) \oplus \text{null}(A^{\#}A)$$

$\Downarrow$

$\text{null}(A^{\#}A) = \text{null}(A)$

- $I_n - A^{\#}A$ is an onto mapping from $\mathbb{R}^n$ to $\text{null}(A)$.

Proof $A(I_n - A^{\#}A) = A - AA^{\#}A = 0$

$$\Downarrow$$

$$A\left((I_n - A^{\#}A)\boldsymbol{x}\right) = 0 \text{ for any } \boldsymbol{x} \in \mathbb{R}^n$$

$$\Downarrow$$

$$(I_n - A^{\#}A)\boldsymbol{x} \in \text{null}(A)$$

That $I_n - A^{\#}A$ is an onto mapping follows from it maps $\text{null}(A)$ to itself. □

More on Projector $A^{\#}A$

- $\mathbb{R}^n = \text{null}(A) \oplus \text{col}(A^{\#}A)$

$$\Downarrow \quad \mathbb{R}^n = \text{null}(A) \oplus \text{row}(A)$$

$$\text{col}(A^{\#}A) = \text{row}(A)$$

$$\Downarrow$$

- $A^{\#}A$ projects a vector onto row(A).

♣ The first property is stated in Proposition 3 of [2] with a proof given in [3].

Reflexive Generalized Inverse

A generalized inverse $A^{\#}$ of A is called *reflexive* if $A^{\#}AA^{\#} = A^{\#}$ also holds.

- $A^{\#}$ is reflexive if and only if $\text{rank}\left(A^{\#}\right) = \text{rank}(A)$.

  (See [4], Theorem 20.3.4, page 501)

If A has full rank then it is reflexive.

Example 3 See [1], p. 22.

Generalized Inverse & Optimization

Problem

$$\min_{\boldsymbol{x}} \frac{1}{2}\boldsymbol{x}^T W \boldsymbol{x} \qquad \text{subject to } A\boldsymbol{x} = \boldsymbol{b}$$

where $W \in \mathbb{R}^{n\times n}$ is non-singular and $A \in \mathbb{R}^{m\times n}$ has rank m.

♣ For a robotic manipulator, we let W be its mass matrix, $\boldsymbol{x} = \dot{\boldsymbol{\theta}}$ its joint velocity, and A and $\boldsymbol{b}$ the Jacobian and velocity of its end effector, respectively. Then $\frac{1}{2}\boldsymbol{x}^T W \boldsymbol{x}$ is the manipulator's kinetic energy.

To solve the minimization problem, we first convert it into an unconstrained one using m Lagrange multipliers represented by a vector $\boldsymbol{\lambda}$:

$$\min_{\boldsymbol{x}} \frac{1}{2}\boldsymbol{x}^T W \boldsymbol{x} + \boldsymbol{\lambda}^T (A\boldsymbol{x} - \boldsymbol{b})$$

⇓ Vanishing of partial derivative w.r.t. $\boldsymbol{x}$

$$\boldsymbol{x}^T W + \boldsymbol{\lambda}^T A = 0$$

⇓

$$\boldsymbol{x} = -W^{-1} A^T \boldsymbol{\lambda}$$

(cont'd)

$$\boldsymbol{x} = W^{-1}A^T\boldsymbol{\lambda}$$

$\Downarrow$ $A\boldsymbol{x} = \boldsymbol{b}$

$$AW^{-1}A^T\boldsymbol{\lambda} = \boldsymbol{b}$$

rank(A) = rank$(A^T) = m$

$\Downarrow$ W non-singular

rank$(W^{-1}A^T) = m$

$\Downarrow$ $AW^{-1}A^T$ has full rank m.

$$\boldsymbol{\lambda} = (AW^{-1}A^T)^{-1}\boldsymbol{b}$$

$$\Downarrow$$

$$\boldsymbol{x} = W^{-1}A^T\boldsymbol{\lambda} = A_W^{\#}\boldsymbol{b}$$

where

$$A_W^{\#} = W^{-1}A^T(AW^{-1}A^T)^{-1}$$

Trivially, we have

$$A\boldsymbol{x} = AA_W^{\#}\boldsymbol{b} = AW^{-1}A^T(AW^{-1}A^T)^{-1}\boldsymbol{b} = \boldsymbol{b}$$

II. Pseudoinverse

A matrix $A^\dagger \in \mathbb{R}^{n\times m}$ is called the *pseudoinverse* of A if it satisfies the following four conditions (called the *Penrose conditions*):

1. $AA^\dagger A = A$ ($A^\dagger$ is a generalized inverse of A.)
2. $A^\dagger AA^\dagger = A^\dagger$ (A is a generalized inverse of $A^\dagger$.)
3. $(AA^\dagger)^T = AA^\dagger$ ($AA^\dagger$ is symmetric.)
4. $(A^\dagger A)^T = A^\dagger A$ ($A^\dagger A$ is symmetric.)

Example 4

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 5 & 7 & 7 & 9 \end{pmatrix} \qquad A^\dagger = \begin{pmatrix} 2 & -1/4 \\ 1/4 & 0 \\ 1/4 & 0 \\ -3/2 & 1/4 \end{pmatrix}$$

- A reflexive generalized inverse $A^\#$ satisfies conditions 1 and 2 but not necessarily conditions 3 and 4. It is also called an $\{1,2\}$-inverse.
  Return to Example 3 ([1], p. 23).
- A *right inverse* R (i.e., $AR = I_m$) satisfies conditions 1, 2, 3. ($\{1,2,3\}$-inverse)
- A *left inverse* L (i.e., $LA = I_n$) satisfies conditions 1, 2, 4. ($\{1,2,4\}$-inverse)

Diagonal and Full-Column-Rank Matrices

- The pseudoinverse always exists and is unique.
- The pseudoinverse is also called the *Moore-Penrose inverse*.
- It follows from the definition that $\left(A^\dagger\right)^\dagger = A$.
- The pseudoinverse of a diagonal matrix $\Sigma \in \mathbb{R}^{m\times n}$ is another diagonal matrix $\Sigma^\dagger \in \mathbb{R}^{n\times m}$.

$\sigma_1, \dots, \sigma_{\min(m,n)}$: diagonal elements of Σ

$\varrho_1, \dots, \varrho_{\min(m,n)}$: diagonal elements of $\Sigma^\dagger$

$$\varrho_i = \begin{cases} 0 \text{ if } \sigma_i = 0; \\ 1/\sigma_i \text{ if } \sigma_i \neq 0. \end{cases}$$

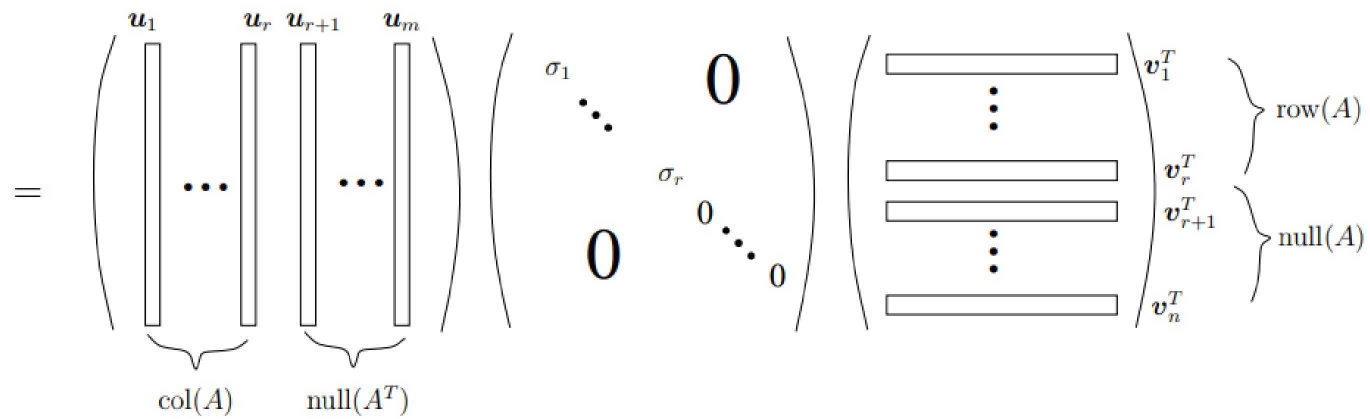
- If $\text{rank}(A) = n$, $A^\dagger = (A^T A)^{-1} A^T$ (a left inverse of A).
- If $\text{rank}(A) = m$, $A^\dagger = A^T (A A^T)^{-1}$ (a right inverse of A).

Pseudoinverse & SVD

- Every matrix $A \in \mathbb{R}^{m \times n}$ has a singular value decomposition:

$$\underset{m \times n}{A} = \underset{m \times m}{U} \; \underset{m \times n}{\Sigma} \; \underset{n \times n}{V^\top}$$

U, V: orthogonal matrices.



$$A = \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1{}^T + \cdots + \sigma_r \boldsymbol{u}_r \boldsymbol{v}_r{}^T$$

$$A^\dagger = V\Sigma^\dagger U^T = \frac{1}{\sigma_1} \boldsymbol{v}_1 \boldsymbol{u}_1{}^T + \cdots + \frac{1}{\sigma_r} \boldsymbol{v}_r \boldsymbol{u}_r{}^T$$

Example 5 See [1], p. 34.

It is easy to show that the pseudoinverse and transpose operators are commutative:

$$(A^T)^\dagger = (A^\dagger)^T$$

Pseudoinverse and Generalized Inverse

- SVD is $A = U\begin{pmatrix}\bar{\Sigma} & 0 \\ 0 & 0\end{pmatrix}V^T$ where $\bar{\Sigma} = \begin{pmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_r\end{pmatrix}$ with $\sigma_1, \dots, \sigma_r \neq 0$.

The set of all generalized inverses of A is [5]

$$\left\{V\begin{pmatrix}\bar{\Sigma}^{-1} & X \\ Y & Z\end{pmatrix}U^T \;\middle|\; X \in \mathbb{R}^{r\times(m-r)}, Y \in \mathbb{R}^{(n-r)\times r}, Z \in \mathbb{R}^{(n-r)\times(m-r)}\right\}$$

- A reflexive generalized inverse (i.e., a $\{1,2\}$-inverse) satisfies $Z = Y\bar{\Sigma}X$.
- A $\{1,3\}$-inverse satisfies $X = 0$.
- A $\{1,4\}$-inverse satisfies $Y = 0$.
- The pseudoinverse trivially has $X = 0$, $Y = 0$, and $Z = 0$.

Orthogonal Projections

A projection matrix P (i.e., $P^2 = P$) is *orthogonal* if $P^T = P$.

Four orthogonal projection matrices:

$$A = \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1{}^T + \cdots + \sigma_r \boldsymbol{u}_r \boldsymbol{v}_r{}^T$$
$$A^\dagger = \frac{1}{\sigma_1} \boldsymbol{v}_1 \boldsymbol{u}_1{}^T + \cdots + \frac{1}{\sigma_r} \boldsymbol{v}_r \boldsymbol{u}_r{}^T$$
$\Rightarrow$

- $AA^\dagger = \boldsymbol{u}_1 \boldsymbol{u}_1{}^T + \cdots + \boldsymbol{u}_r \boldsymbol{u}_r{}^T$ onto col(A).
- $A^\dagger A = \boldsymbol{v}_1 \boldsymbol{v}_1{}^T + \cdots + \boldsymbol{v}_r \boldsymbol{v}_r{}^T$ onto row(A).
- $I_n - A^\dagger A = \boldsymbol{v}_{r+1} \boldsymbol{v}_{r+1}{}^T + \cdots + \boldsymbol{v}_n \boldsymbol{v}_n{}^T$ onto null(A).

  null$(A) = \{(I_n - A^\dagger A)\boldsymbol{w} \mid \boldsymbol{w} \in \mathbb{R}^n\}$

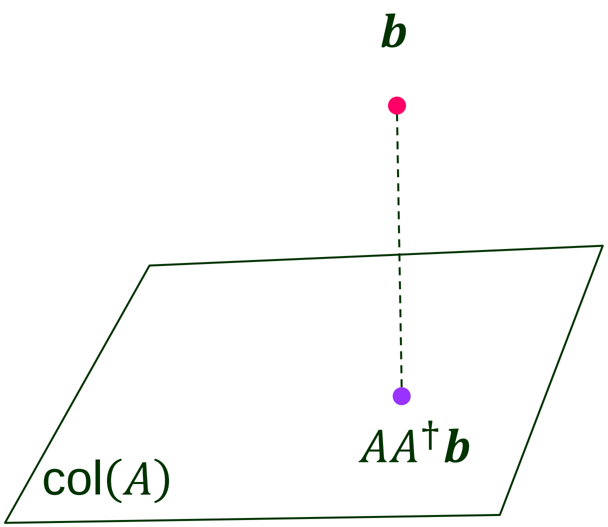
- $I_m - AA^\dagger = \boldsymbol{u}_{r+1} \boldsymbol{u}_{r+1}{}^T + \cdots + \boldsymbol{u}_m \boldsymbol{u}_m{}^T$ onto null(A^T).

Least Squares

$$\min_{\boldsymbol{x}} ||A\boldsymbol{x} - \boldsymbol{b}||$$

Since $A\boldsymbol{x} \in \text{col}(A)$, the optimizing $\boldsymbol{x}$ must make $A\boldsymbol{x}$ coincide with the orthogonal projection of $\boldsymbol{b}$ onto $\text{col}(A)$, namely,

$$A\boldsymbol{x} = AA^\dagger\boldsymbol{b}$$



- $\boldsymbol{x} = A^\dagger\boldsymbol{b}$ is an obvious solution to the minimization problem.
- $\{A^\dagger\boldsymbol{b} + (I_n - A^\dagger A)\boldsymbol{w} \mid \boldsymbol{w} \in \mathbb{R}^n\}$ is the entire solution set.
- If $b \in \text{col}(A)$, then $\boldsymbol{x} = A^\dagger\boldsymbol{b}$ is also the minimum-norm solution to $A\boldsymbol{x} = \boldsymbol{b}$.
- If A has rank n, then $\boldsymbol{x} = (A^TA)^{-1}A^T\boldsymbol{b}$.

References

1. Guangliang Chen, Lecture 6: Generalized inverse and pseudoinverse. *Math 253: Mathematical Methods for Data Visualization*. San Jose State University.

2. A. Korporal, G. Regensburger. "On the product of projectors and generalized inverses," RICAM-Report 2012-22, Johan Radon Institute for Computational and Applied Mathematics, Austrian Academy of Sciences, 2012.

3. M. Nashed, G. Vortuba, "A unified operator theory of generalized inverses," in M. Z. Nashed (Ed.), *Generalized Inverses and Applications*, Academic Press, New York, 1976, pp. 1–109.

4. David A. Harville, *Matrix Algebra From a Statistician's Perspective*, Springer, 2008.

5. Adi Ben-Israel and Thomas Nall Eden Greville. *Generalized Inverses: Theory and Applications* (2nd ed.), Springer, New York, 2003.

6. Yan-Bin Jia, "Singular Value Decomposition," *Com S 477/577: Foundations of Robotics and Computer Vision*.