

Inverses of matrices play an important role in linear algebra. They are particularly familiar as a means to express compactly the solution to a system of simultaneous linear equations. However, a matrix has an inverse only when it is square and nonsingular. The pseudo-inverse is a generalization of the inverse to the case of singular or even nonsquare matrices. In the field of robotics, pseudo-inverses are often used in the identification of manipulator dynamics and in the control of redundant manipulators. Although this appendix will deal only with matrices of real elements, the following arguments are also valid for matrices of complex elements with a little modification.

We start with the definition of the pseudo-inverse. For every finite $m \times n$ real matrix A , there is a unique $n \times m$ real matrix A^+ satisfying the following four conditions:

$$AA^+A = A, \tag{A2.1}$$

$$A^+AA^+ = A^+, \tag{A2.2}$$

$$(AA^+)^T = AA^+, \tag{A2.3}$$

$$(A^+A)^T = A^+A. \tag{A2.4}$$

This A^+ is called the *pseudo-inverse* of A .

We first show the existence of A^+ satisfying equations A2.1–A2.4. If $A = \mathbf{0}$, then the obvious solution $A^+ = \mathbf{0}$ satisfies these conditions. If we assume that $A \neq \mathbf{0}$ and let $r = \text{rank}A$, then A can be expressed as the product of an $m \times r$ matrix B and an $r \times n$ matrix C :

$$A = BC. \tag{A2.5}$$

Using the above B and C , we define an $n \times m$ matrix D by

$$D = C^T(CC^T)^{-1}(B^TB)^{-1}B^T. \tag{A2.6}$$

Then it is easy to show that $A^+ = D$ satisfies equations A2.1–A2.4. This completes the proof of the existence of A^+ .

Next we will show the uniqueness of A^+ . Let any two matrices satisfying equations A2.1–A2.4 be A_1^+ and A_2^+ . Then

$$\begin{aligned} A_1^+ - A_2^+ &= A_1^+AA_1^+ - A_2^+AA_2^+ \\ &= A^TA_1^{+T}A_1^+ - A_2^+A_2^{+T}A^T \\ &= (AA_2^+A)^TA_1^{+T}A_1^+ - A_2^+A_2^{+T}(AA_1^+A)^T \end{aligned}$$

$$\begin{aligned} &= A^TA_2^{+T}A^TA_1^{+T}A_1^+ - A_2^+A_2^{+T}A^TA_1^{+T}A^T \\ &= A^TA_2^{+T}A_1^+ - A_2^+A_1^{+T}A^T \\ &= A_2^+AA_1^+ - A_2^+AA_1^+ = \mathbf{0}, \end{aligned} \tag{A2.7}$$

proving the uniqueness of A^+ .

The pseudo-inverse has the following properties:

$$(i) (A^+)^+ = A. \tag{A2.8}$$

$$(ii) (A^T)^+ = (A^+)^T, (AA^T)^+ = (A^+)^TA^+. \tag{A2.9}$$

(iii) Let A be an $m \times n$ matrix and B be an $n \times p$ matrix. Then generally $(AB)^+$ is not equal to B^+A^+ . If, however, $\text{rank}A = \text{rank}B = n$, then

$$(AB)^+ = B^+A^+. \tag{A2.10}$$

$$(iv) A^+ = (A^TA)^+A^T = A^T(AA^T)^+. \tag{A2.11}$$

(v) If an $m \times n$ matrix A satisfies $\text{rank}A = m$, then

$$A^+ = A^T(AA^T)^{-1}. \tag{A2.12}$$

Also, if $\text{rank}A = n$, then

$$A^+ = (A^TA)^{-1}A^T. \tag{A2.13}$$

(vi) Let $r = \text{rank}A$, and let the singular-value decomposition (see appendix 3) of A be

$$A = U\Sigma V^T, \tag{A2.14}$$

where U and V are appropriate orthogonal matrices and Σ is the following $m \times n$ matrix:

$$\Sigma = \left[\begin{array}{ccc|ccc} \sigma_1 & & & & & \\ & \sigma_2 & & & & \\ & & \ddots & & & \\ & & & \sigma_r & & \\ \hline & & & & & \\ \hline & & & & & \end{array} \right] \begin{array}{l} \left. \vphantom{\begin{array}{ccc|ccc} \sigma_1 & & & & & \\ & \sigma_2 & & & & \\ & & \ddots & & & \\ & & & \sigma_r & & \\ \hline & & & & & \\ \hline & & & & & \end{array}} \right\} r \\ \left. \vphantom{\begin{array}{ccc|ccc} & & & & & \\ \hline & & & & & \\ \hline & & & & & \end{array}} \right\} m-r \end{array} \tag{A2.15}$$

with

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0.$$

Then A^+ is given by

$$A^+ = V\Sigma^+U^T, \tag{A2.16}$$

where Σ^+ is the $n \times m$ matrix defined by

$$\Sigma^+ = \left[\begin{array}{ccc|ccc} \sigma_1^{-1} & & & & & \\ & \sigma_2^{-1} & & & & \\ & & \ddots & & & \\ & & & \sigma_r^{-1} & & \\ \hline & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \end{array} \right] \left. \begin{array}{l} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{array} \right\} \begin{array}{l} r \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ n-r \end{array} \tag{A2.17}$$

Note that Σ^+ is the pseudo-inverse of Σ .

(vii) Let us consider a system of simultaneous linear equations given by

$$Ax = b, \tag{A2.18}$$

where A is a known $m \times n$ matrix, b is a known m -dimensional vector, and x is an unknown n -dimensional vector. When A is square ($m = n$) and nonsingular, the solution to equation A2.18 is, as is well known, given by

$$x = A^{-1}b. \tag{A2.19}$$

When A is not nonsingular, if A and b satisfy

$$b \in R(A), \tag{A2.20}$$

then equation A2.18 is solvable but the solution is not unique. Taking this into account, we now consider a more general problem of minimizing the Euclidean norm of error ($Ax - b$):

$$\|Ax - b\| = \sqrt{(Ax - b)^T(Ax - b)}. \tag{A2.21}$$

As is proved below, the general solution to this problem is

$$x = A^+b + (I - A^+A)k, \tag{A2.22}$$

where k is an arbitrary n -dimensional vector. If equation A2.18 has a solution, equation A2.22 is its general solution; if equation A2.18 has no solution, then equation A2.22 is the general form of its best approximate solution in the sense that it minimizes the norm (equation A2.21). Furthermore, if the best approximate solution is not unique, the one that minimizes its own norm $\|x\|$ is given by

$$x = A^+b. \tag{A2.23}$$

We will prove that equation A2.22 is the general solution to the problem of minimizing $\|Ax - b\|$. From equations A2.8–A2.11, we have

$$A^TAA^+ = A^T. \tag{A2.24}$$

Using this relation, we can show that

$$\begin{aligned} & b^T[I - (AA^+)^TAA^+]b \\ & + [x - A^+b - (I - A^+A)k]^T A^T A [x - A^+b - (I - A^+A)k] \\ & = x^T A^T A x - 2x^T A^T b + b^T b \\ & = \|Ax - b\|^2. \end{aligned} \tag{A2.25}$$

Since the first term on the left-hand side of equation A2.25 is independent of x , $\|Ax - b\|$ becomes minimum if and only if the second term is equal to zero, that is, if and only if

$$A[x - A^+b - (I - A^+A)k] = 0. \tag{A2.26}$$

Hence, equation A2.22 is a solution to the problem for any k . On the other hand, since equation A2.26 is equivalent to

$$A[x - A^+b] = 0 \tag{A2.27}$$

by equation A2.1, any solution x^* to the problem must satisfy

$$Ax^* = AA^+b. \tag{A2.28}$$

Thus,

$$\begin{aligned} x^* &= A^+b + x^* - A^+b \\ &= A^+b + x^* - A^+AA^+b \\ &= A^+b + (I - A^+A)x^*. \end{aligned} \tag{A2.29}$$

Therefore, \mathbf{x}^* can be expressed in the form of equation A2.22. This completes the proof.

The reader who wishes to learn more about pseudo-inverses is referred to the following:

A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications* (Wiley, 1974).

T. L. Boullion and P. L. Odell, *Generalized Inverse Matrices* (Wiley, 1971).

C. R. Rao and S. K. Mitra, *Generalized Inverse of Matrices and Its Applications* (Wiley, 1971).