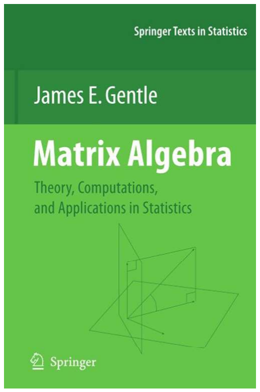
Theory of Matrices

Chapter 3. Basic Properties of Matrices

3.6. Generalized Inverses—Proofs of Theorems



Theorem 3.6.1

Theorem 3.6.1. Let A be an $n \times m$ matrix. Then a generalized inverse of A exists.

Proof. If $\text{rank}(A) = 0$ (and so $A = 0$) then every $m \times n$ matrix is a generalized inverse of A, as observed above.

If $\text{rank}(A) > 0$ then by Theorem 3.3.9, there are matrices P and Q, both products of elementary matrices, such that $A = P \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} Q$. All elementary matrices are invertible so P^{-1} and Q^{-1} exist. Consider $Q^{-1} \begin{bmatrix} I_r & U \\ V & W \end{bmatrix} P^{-1}$ where U, V, and W are *any* matrices where U is $r \times (m-r)$, V is $(n-r) \times r$, and W is $(n-r) \times (m-r)$.

Theorem 3.6.1 (continued)

Theorem 3.6.1. Let A be an $n \times m$ matrix. Then a generalized inverse of A exists.

Proof (continued). Then

$$A\left(Q^{-1}\begin{bmatrix} I_r & U \\ V & W \end{bmatrix}P^{-1}\right)A$$

$$= P\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}QQ^{-1}\begin{bmatrix} I_r & U \\ V & W \end{bmatrix}P^{-1}P\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}Q$$

$$= P\begin{bmatrix} I_r & U \\ 0 & 0 \end{bmatrix}\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}Q$$

$$= P\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}Q = A$$

and so $A^- = Q^{-1}\begin{bmatrix} I_r & U \\ V & W \end{bmatrix}P^{-1}$ is a generalized inverse of A. □

Theorem 3.6.2

Theorem 3.6.2. Every matrix A with $\text{rank}(A) = r > 0$ has a pseudoinverse given be $A^+ = R^T(L^TAR^T)^{-1}L^T$ where $A = LR$ is a full rank factorization of A (such a factorization exists as shown in equations $(**)$ and $(***)$ of Section 3.4).

Proof. Let A be $n \times m$ with $\text{rank}(A) = r > 0$. Then there is a full rank factorization of A, $A = LR$ where L is a $n \times r$ full column rank matrix and R is a $r \times m$ full row rank matrix. So by Theorem 3.3.14(5), L^TL is full rank r. Since R^T is full column rank, similarly (by Theorem 3.3.14(5)) $(R^T)^TR^T = RR^T$ is full rank r. So L^TL and RR^T are both invertible. Hence $(L^TL)(RR^T) = L^T(LR)R^T = L^TAR^T$ is invertible. Consider $B = R^T(L^TAR^T)^{-1}L^T$. The fact that B satisfies the four parts of the definition of pseudoinverse is to be given in Exercise 3.15. So $A^+ = B$ is a pseudoinverse of A. □

Theorem 3.6.3

Theorem 3.6.3. For any matrix A, the pseudoinverse A^+ is unique.

Proof. The case $A = 0$ is addressed in the note above. For A with rank$(A) > 0$, suppose both B and C are pseudoinverses. We use the four properties of a pseudoinverse to prove that $B = C$. We have

$$\begin{aligned} B &= BAB \text{ by Property (2) for } B \\ &= (BA)^T B = A^T B^T B \text{ by Property (3) for } B \\ &= (ACA)^T B^T B \text{ by Property (1) for } C \\ &= A^T C^T A^T B^T B = (CA)^T A^T B^T B \\ &= CAA^T B^T B \text{ by Property (3) for } C \\ &= CA(BA)^T B = CA(BA)B \text{ by Property (3) for } B \\ &= C(ABA)B = CAB \text{ by Property (1) for } B \\ &\quad \ldots \end{aligned}$$

Theorem 3.6.3 (continued)

Theorem 3.6.3. For any matrix A, the pseudoinverse A^+ is unique.

Proof (continued). . . .

$$\begin{aligned} CAB &= C(ACA)B = C(AC)AB \text{ by Property (1) for } C \\ &= C(AC)^T AB = CC^T A^T AB \text{ by Property (3) for } C \\ &= CC^T A^T (AB)^T = CC^T (ABA)^T \text{ by Property (4) for } B \\ &= CC^T A^T = C(AC)^T \text{ by Property (1) for } B \\ &= CAC \text{ by Property (4) for } C \\ &= C \text{ by Property (2) for } C. \end{aligned}$$

So $B = C$ and the pseudoinverse of A is unique. □