

**UNCLASSIFIED**

**AD**

**405 938**

**DEFENSE DOCUMENTATION CENTER**

**FOR**

**SCIENTIFIC AND TECHNICAL INFORMATION**

**CAMERON STATION, ALEXANDRIA, VIRGINIA**



**UNCLASSIFIED**

NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.

(18) (19) (4) 260 (5) 17505  
SSD-TDR 63/92 (7) + (8) NA (9) NA

(14)  
TDR 169/3550/01/TN-2  
# # # # #

(12) 14p. (13) NA

(16) (17) NA

(20) u  
(21) NA

fd

(6) THE MATRIX PSEUDOINVERSE AND  
MINIMAL VARIANCE ESTIMATES,

(10) by C.M. Price

(11) 25 Mar 1963

(15) Contract # AF04/695/169  
# # #

AEROSPACE CORPORATION  
2400 East El Segundo Boulevard  
El Segundo, California

Prepared for  
COMMANDER SPACE SYSTEMS DIVISION  
UNITED STATES AIR FORCE  
Inglewood, California

THE MATRIX PSEUDOINVERSE AND  
MINIMAL VARIANCE ESTIMATES

25 March 1963

Prepared by Charles M Price  
C.M. Price, Associate Head  
Astrodynamics Department  
Systems Research and Planning Division

Approved by E. Levin  
E. Levin, Head  
Astrodynamics Department  
Systems Research and Planning Division

AEROSPACE CORPORATION  
2400 East El Segundo Boulevard  
El Segundo, California

## ABSTRACT

↓ This paper reviews and applies certain results concerning the matrix pseudoinverse to the general theory of estimable functions and minimal variance estimates. The paper is divided into two sections. The first section reviews and extends certain known results concerning the matrix pseudoinverse. This section is essentially nonstatistical. The second section uses results in the first section to state and prove a generalized version of the Gauss-Markoff Theorem (~~Reference 4, page 14~~) concerning unbiased linear estimates having minimal variance. In the third section, an additional theorem is proven, which together with the preceding material provides a theoretical foundation for parameter estimation in orbit determination work. This foundation is then exploited to provide formulae for such parameters. ↑

CONTENTS -

	<u>Page</u>
I. PROPERTIES OF THE MATRIX PSEUDOINVERSE . . . . .	1
II. THE GAUSS-MARKOFF THEOREM. . . . .	8
III. PARAMETER ESTIMATION FOR ORBIT DETERMINATION. . . . .	11
REFERENCES . . . . .	14

## I. PROPERTIES OF THE MATRIX PSEUDOINVERSE

All matrices will be designated by capital English letters; the notation  $A(m \times n) = (a_{ij})$  means that  $A$  is an  $m \times n$  matrix having  $a_{ij}$  as the element in the  $i^{\text{th}}$  row and  $j^{\text{th}}$  column ( $i=1, \dots, m = \text{number of rows}; j=1, \dots, n = \text{number of columns}$ ). All matrices considered will be presumed to have elements  $a_{ij}$  which are real numbers.  $A^*$  designates the transpose of  $A$ ; thus if  $A(m \times n) = (a_{ij})$  then  $A^*(n \times m) = (b_{ij})$  where  $b_{ij} = a_{ji}$ .  $I_n$  is the  $(n \times n)$  identity matrix with 1's down the diagonal and zeros elsewhere; usually the subscript  $n$  will be dropped if the dimension of  $I$  is clear from the context.  $E_{rs}^n$  is the  $(n \times n)$  matrix  $(e_{ij})$  such that  $e_{rs} = 1$  and all other  $e_{ij} = 0$ ; usually the superscript  $n$  will be dropped if the dimension of  $E_{rs}$  is clear from the context.  $A(m \times n)$  is called square if  $m=n$ . If  $A$  is square then  $|A|$  designates the determinant of  $A$ . If  $A$  is square and  $|A| \neq 0$  then  $A^{-1}$  designates the inverse of  $A$ , and  $A$  is said to be nonsingular.  $\mathcal{E}_n$  designates the set of all real  $(n \times 1)$  matrices;  $\mathcal{E}_n$  will also be called Euclidean  $n$ -space, and elements of  $\mathcal{E}_n$  called  $n$  dimensional vectors, or just vectors if the dimension is clear from the context. The symbol  $0$  will be used to designate either a matrix which is identically zero or the scalar real number zero, depending upon the context. Given any  $x$  in  $\mathcal{E}_n$ ,  $\|x\| \equiv \sqrt{x^* x}$  and is called the "norm" of  $x$ . The symbol " $\epsilon$ " will sometimes be used for "in" in the set theoretic sense; e. g., "given any  $x \in \mathcal{E}_n$ " means given any  $x$  which is an element (a member) of the set  $\mathcal{E}_n$ . Let  $W(m \times m)$  be positive definite so that  $W = R^* R$  for some square  $R$ ,  $|R| \neq 0$ ,  $R = R^*$ . For any  $a \in \mathcal{E}_n$ , define  $\|a\|_w^2 \equiv a^* R^* R a = \|Ra\|^2$ .

Theorem 1:<sup>†</sup> For every  $(m \times n)$  matrix  $A$ , there exists a unique  $(n \times m)$  matrix, which we shall designate as  $A^+$ , that satisfies the following four identities:

<sup>†</sup> The bulk of the material in Section I constitutes a review and restatement of results to be found in (1) - (3) convenient for present purposes. In general, the absence of a proof following the statement of a result indicates that the result and proof may be found in (1) - (3), and vice versa.

- (1)  $AA^+A = A$
- (2)  $A^+AA^+ = A^+$
- (3)  $(AA^+)^* = AA^+ (m \times m)$
- (4)  $(A^+A)^* = A^+A (n \times n)$

Furthermore,

(R1) if  $D = (d_{ij})$  is square ( $m = n$ ) and diagonal ( $d_{ij} = 0$  for  $i \neq j$ ) then  $D^+ = (d_{ij}^+)$  is defined by  $d_{ij}^+ = 0$  for  $i \neq j$ ,  $d_{ii}^+ = 0$  if  $d_{ii} = 0$ ,  $d_{ii}^+ = d_{ii}^{-1}$  if  $d_{ii} \neq 0$ .

(R2) if  $A^*A = PDP^*$ , where  $PP^* = P^*P = I$ , and  $D$  is diagonal, then  $A^+ = PD^+P^*A^*$ .

(R3) if  $A = BC$ , where the columns of  $B$  are linearly independent and the rows of  $C$  are linearly independent, then  $A^+ = C^*(CC^*)^{-1}(B^*B)^{-1}B^*$ .  
Thus

(R3.1)  $A^+ = (A^*A)^{-1}A^*$  if the columns of  $A$  are linearly independent.

(R3.2)  $A^+ = A^*(AA^*)^{-1}$  if the rows of  $A$  are linearly independent.

(R3.3)  $A^+ = A^{-1}$  if  $A$  is square and nonsingular.

Theorem 2: The matrix correspondence  $A \rightarrow A^+$  satisfies the following

- (R1)  $(A^+)^+ = A$
- (R2)  $(A^*)^+ = (A^+)^* \equiv A^{+*} \equiv A^{**+}$
- (R3)  $A^+AA^* = A^*$
- (R4)  $A^*AA^+ = A^*$
- (R5)  $AA^+A^{+*} = A^{+*}$
- (R6)  $A^{+*}A^+A = A^{+*}$
- (R7)  $A^{**+}A^*A = A$



$$(R8) \quad A A^* A^{*+} = A$$

$$(R9) \quad A^* A^{+*} A^+ = A^+$$

$$(R10) \quad A^+ A^{**} A^* = A^+$$

(R11) The row spaces of  $A^+$  and  $A^*$  are identical, i. e., the rows of  $A^+$  are in the row space of  $A^*$  and the rows of  $A^*$  are in the row space of  $A^+$ .

(R12) The column spaces of  $A^+$  and  $A^*$  are identical.

(R13)  $A$ ,  $A^+$  and  $A^*$  all have the same rank.

$$(R14) \quad (A A^*)^+ = A^{+*} A^+$$

$$(R15) \quad (A A^*)^+ (A A^*) = A A^+$$

(R16) If  $A^+$  commutes with some power of  $A$  and  $\lambda$  is any non-zero eigenvalue of  $A$  corresponding to the eigenvector  $x$ , then  $\lambda^{-1}$  is an eigenvalue of  $A^+$  corresponding to the eigenvector  $x$ .

$$(R17) \quad \text{If } \alpha \neq 0 \text{ then } (\alpha A)^+ = \alpha^{-1} A^+$$

$$(R18) \quad 0^+ = 0$$

Proof: All of (R1) - (R18) follow directly with the possible exception of (R16). To prove the latter, let  $A^+$  commute with  $A^n$  for some positive interger  $n$ , and let  $\lambda \neq 0$  be an eigenvalue of  $A$  corresponding to the eigenvector  $x$  so that  $Ax = \lambda x$ ,  $x = \lambda^{-1} Ax$ ,  $A^+ A^n = A^n A^+$ . Then  $A^+ x = \lambda^{-1} A^+ Ax = \lambda^{-2} A^+ A^2 x = \lambda^{-n} A^+ A^n x = \lambda^{-n} A^n A^+ x = \lambda^{-n-1} A^n A^+ Ax = \lambda^{-n-1} A^{n-1} (A A^+ A) x = \lambda^{-n-1} A^{n-1} Ax = (\lambda^{-1}) (\lambda^{-n}) A^n x = \lambda^{-1} x$ , q. e. d.

Remark: It is not true in general that  $(AB)^+$  is  $B^+ A^+$ . For a counter-example, let

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

Then

$$(AB)^+ = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \quad C = B^+ A^+ = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Note that  $C$  satisfies properties (1), (2), (3) but not (4) for  $(AB)^+$  listed in Theorem 1.

Remark: Let

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$$

Then

$$A^+ = \begin{pmatrix} 0.1 & 0.1 \\ 0.2 & 0.2 \end{pmatrix} = (0.1) A^*$$

The eigenvalues of  $A$  are 3 and 0. The eigenvalues of  $A^+$  are clearly 0.3 and 0. The lack of any "inverse" relation between 3 and 0.3 shows that (R16) of Theorem 2 needs some hypothesis. More generally, let  $A$  be the singular  $2 \times 2$  matrix ( $\alpha\beta \neq -1$ ).

$$A = \begin{pmatrix} 1 & \beta \\ \alpha & \alpha\beta \end{pmatrix}; \text{ then } (1 + \alpha\beta) \leftrightarrow \begin{pmatrix} 1 \\ \alpha \end{pmatrix} \text{ and } 0 \leftrightarrow \begin{pmatrix} \beta \\ -1 \end{pmatrix},$$

the notation meaning  $A$  has the eigenvalue  $1 + \alpha\beta$  corresponding to the eigenvector

$$\begin{pmatrix} 1 \\ \alpha \end{pmatrix}$$

etc. Now  $A^+$  is readily computed to be  $A^+ = (1 + \beta^2)^{-1} (1 + \alpha^2)^{-1} A^*$ . The non-zero eigenvalue  $(1 + \alpha\beta)$  of  $A$  thus corresponds to the non-zero eigenvalue  $(1 + \beta^2)^{-1} (1 + \alpha^2)^{-1} (1 + \alpha\beta)$  of  $A^+$ . The latter is  $(1 + \alpha\beta)^{-1}$  if and only if  $\alpha = \beta$ , which is true if and only if  $AA^+ = A^+A$ .

Examples:

(E1)  $E_{ij}^+ = E_{ji}$

(E2) If  $A = A^* = A^2$ , then  $A^+ = A$

(E3)  $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}^+ = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$

(E4)  $\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}^+ = (0.1) \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$

Definition: For  $A (m \times n)$  and  $b (m \times 1)$ ,  $\mathcal{H}(A, b) \equiv A^+b + (I_n - A^+A) \mathcal{E}_n$ .Lemma 1: The element of least norm in  $\mathcal{H}(A, b)$  is  $A^+b$ .Proof: Let  $x \in \mathcal{H}(A, b)$ . Then  $x = A^+b + (I - A^+A)v$  for some  $v \in \mathcal{E}_n$ , and  $\|x\|^2 = \|A^+b\|^2 + \|(I - A^+A)v\|^2$  since  $(I - A^+A)v$  and  $A^+b$  are orthogonal;  $[(I - A^+A)v]^*(A^+b) = v^*(I - A^+A)^*A^+b = v^*(I - A^+A)A^+b = 0$ . Thus  $\|x\| > \|A^+b\|$  unless  $\|(I - A^+A)v\| = 0$ , i. e.,  $\|x\| > \|A^+b\|$  unless  $x = A^+b$ , q. e. d.Lemma 2: Let  $x \in \mathcal{E}_n$ . Then  $x \in \mathcal{H}(A, b)$  if and only if  $A(x - A^+b) = 0$ .Proof: Suppose  $A(x - A^+b) = 0$ . Let  $x = A^+b + y$ . Then  $Ay = 0$  so that  $y = A^+Ay + (I - A^+A)y = (I - A^+A)y$ ,  $x = A^+b + (I - A^+A)y \in \mathcal{H}(A, b)$ . The converse is trivial, q. e. d.Lemma 3: Let  $x \in \mathcal{E}_n$ . Then  $x \in \mathcal{H}(A, b)$  if and only if  $A^*(Ax - b) = 0$ .Proof: We shall use Lemma 2 in the proof. We must show that  $A(x - A^+b) = 0$  if and only if  $A^*(Ax - b) = 0$ . Suppose that  $A(x - A^+b) = 0$ . Then  $0 = A^*A(x - A^+b) = A^*Ax - A^*b$  using (R4) of Theorem 2. Conversely, let  $A^*Ax = A^*b$ . Then  $A^{**}A^*Ax = A^{**}A^*b = Ax = AA^+b$  using (R7) of Theorem 2 and the fact that  $(AA^+)^* = AA^+ = A^{**}A^*$ , q. e. d.Lemma 4: Let  $A (m \times n)$ ,  $N (n \times n)$  nonsingular,  $M (m \times m)$  nonsingular. Then  $(AN)(AN)^+ = AA^+$  and  $(MA)^+(MA) = A^+A$ .

Proof: Let  $L \equiv (AN)(AN)^+$ ,  $R \equiv AA^+$ . Then  $L^2 = L = L^*$  and  $R^2 = R = R^*$ ,  $LR = (AN)(AN)^+AA^+ = (AN)(AN)^+(AN)(N^{-1}A^+) = ANN^{-1}A^+ = AA^+ = R$ ,  $(LR)^* = R^* = R = R^*L^* = RL$ . Let  $B \equiv AN$ . Then  $RL = (BN^{-1})(BN^{-1})^+BB^+ = (BN^{-1})(BN^{-1})^+(BN^{-1})NB^+ = BN^{-1}NB^+ = BB^+ = L$ . Thus  $R = RL = L$ ,  $(AN)(AN)^+ = AA^+$ . Using this identity we see that  $(A^*M^*)^+ = A^*A^{**+}$ . Transposing we get  $(MA)^+(MA) = A^+A$ , q. e. d.

Lemma 5: Let  $A(m \times n)$ , and  $N(n \times n)$  nonsingular. Then  $\mathcal{H}(A, b) = N\mathcal{H}(AN, b)$ .

Proof: Let  $x = A^+b + (I - A^+A)v \in \mathcal{H}(A, b)$ . Then  $AN[N^{-1}x - (AN)^+b] = Ax - (AN)(AN)^+b = AA^+b - AA^+b = 0$ ,  $N^{-1}x \in \mathcal{H}(AN, b)$  using Lemma 2,  $N^{-1}\mathcal{H}(A, b) \subset \mathcal{H}(AN, b)$ ,  $\mathcal{H}(A, b) \subset N\mathcal{H}(AN, b)$ . Using this identity we see that  $\mathcal{H}(AN, b) \subset N^{-1}\mathcal{H}(ANN^{-1}, b) = N^{-1}\mathcal{H}(A, b)$ ,  $N\mathcal{H}(AN, b) \subset \mathcal{H}(A, b)$ , and so  $\mathcal{H}(A, b) = N\mathcal{H}(AN, b)$ .

Lemma 6: Let  $V = S^2$  be positive definite ( $n \times n$ ),  $A(m \times n)$ ,  $b(m \times 1)$ . Then the vector ( $n \times 1$ )  $x$  of least  $\|x\|_V$  in  $\mathcal{H}(A, b)$  is given by  $S^{-1}(AS^{-1})^+b$ .

Proof: Let  $u \equiv S^{-1}(AS^{-1})^+b$ . Then  $Au = AA^+b$  using Lemma 4,  $A(u - A^+b) = 0$ ,  $u \in \mathcal{H}(A, b)$  using Lemma 2. Now  $\mathcal{H}(A, b) = S^{-1}\mathcal{H}(A, S^{-1}b)$  using Lemma 5. Let  $x \in \mathcal{H}(A, b)$ . Then  $x = S^{-1}y$  where  $y \in \mathcal{H}(A, S^{-1}b)$ . Now  $\|y\| \geq \|(AS^{-1})^+b\|$  by Lemma 1. Thus  $\|x\|_V = \|Sx\| = \|y\| \geq \|(AS^{-1})^+b\| = \|S^{-1}(AS^{-1})^+b\|_V = \|u\|_V$  and so  $u$  is the element of least "V" norm in  $\mathcal{H}(A, b)$ .

Lemma 7: The following statements are equivalent:

- (1) The columns of  $A$  are linearly independent.
- (2)  $A^*A$  is nonsingular.
- (3)  $A^+A = I$ .

Proof: (1)  $\rightarrow$  (2): Let  $A^*Ax = 0$ . Then  $0 = x^*A^*Ax = \|Ax\|^2$ ,  $0 = Ax$ ,  $x = 0$ . We now show (2)  $\rightarrow$  (3). Let  $u = (A^+A - I)y$ . Then  $A^*Au = 0$ ,  $u = 0$ , and  $(A^+A - I)y = 0$  for all  $y$ ,  $A^+A = I$ . Lastly, we prove (3)  $\rightarrow$  (1). Let  $Ax = 0$ . Then  $A^+Ax = 0 = x$ , q. e. d.

Theorem 3: The equation  $Ax = b$  has a solution (vector)  $x$  if and only if  $AA^+b = b$ . If the latter equality holds then  $x$  is a solution if and only if  $x \in \mathcal{N}(A, b)$ .

Proof: If  $AA^+b = b$  then clearly  $A^+b$  is a solution. Conversely suppose  $Ax = b$  for some  $x \in \mathcal{E}_n$ . Then  $AA^+b = AA^+Ax = Ax = b$ . This proves the first statement. To prove the second statement, let  $AA^+b = b$ . Then  $x$  is a solution of  $Ax = b$  if and only if  $0 = Ax - b = Ax - AA^+b = A(x - A^+b)$ . By Lemma 2 this is true if and only if  $x \in \mathcal{N}(A, b)$ .

Theorem 4: (Least Squares) - For  $A(m \times n)$  and  $b(m \times 1)$ , the set of all  $(n \times 1)$  vectors  $x$  such that  $\|Ax - b\|$  is a minimum, is  $\mathcal{N}(A, b)$ . Also, the  $n \times 1$  matrix (vector) of least norm such that  $\|b - Ax\|$  is minimized, is  $A^+b$ .

Corollary 1: For  $A(m \times n)$  and  $b(m \times 1)$ , and  $W(m \times m) = R^*R$  which is positive definite, the set of all  $(n \times 1)$  vectors such that  $\|Ax - b\|_W$  is a minimum, is  $\mathcal{N}(RA, Rb)$ . The vector of least norm such that  $\|Ax - b\|_W$  is minimized, is  $(RA)^+Rb$ .

Corollary 2: Let  $V = S^2$  be positive definite  $(n \times n)$ ,  $W = R^2$  positive definite  $(m \times m)$ ,  $A(m \times n)$ ,  $b(m \times 1)$ . Then the set of all  $(n \times 1)$  vectors such that  $\|Ax - b\|_W$  is a minimum, is  $\mathcal{N}(RAS^{-1}, Rb)$ .

The vector of least  $\| \cdot \|_W$  norm such that  $\|Ax - b\|_W$  is minimized is  $S^{-1}(RAS^{-1})^+Rb$ .

Proof: Let  $x \in \mathcal{E}_n$ . Then

$$\|b - Ax\|^2 = \|AA^+(b - Ax)\|^2 + \|(I - AA^+)(b - Ax)\|^2 =$$

$$\|AA^+(b - Ax)\|^2 + \|(I - AA^+)b\|^2 =$$

$$\|A(A^+b - x)\|^2 + \|(I - AA^+)b\|^2$$

Thus  $x$  minimizes  $\|b - Ax\|$  if and only if  $A(A^+b - x) = 0$  which is true if and only if  $x \in \mathcal{N}(A, b)$  using Lemma 2. This proves the first statement of the theorem. The second follows immediately from Lemma 1. The proof of the

Corollary 1 follows directly from the identity  $\|Ax - b\|_w = \|R(Ax - b)\|$ . The first statement of Corollary 2 is a restatement of part of Corollary 1. The last statement of Corollary 2 is an immediate consequence of Lemma 6.

Theorem 5: Let  $A$  be  $m \times n$  and  $z$  be any  $m \times 1$  matrix (vector). Then there exists  $m \times 1$  matrices (vectors)  $x$  and  $y$  such that

- (1)  $z = x + y$
- (2)  $x$  is in the column space of  $A$
- (3)  $y$  is orthogonal to the column space of  $A$

Any vectors satisfying (1) - (3) above are unique, and

- (4)  $x = AA^+z$
- (5)  $y = z - AA^+z$
- (6)  $x^*y = 0$

Thus  $AA^+$  is the projection which takes any column vector ( $m \times 1$ ) into the column space of  $A$ ;  $I_m - AA^+$  is the projection which takes any ( $m$ ) vector into the orthogonal complement of the column space of  $A$ . The proof of this theorem follows directly from (4), (5), and (6) above.

## II. THE GAUSS-MARKOFF THEOREM

$\mathcal{V}$  will designate a vector space of real valued random variables over the real numbers  $\mathcal{R}$ . There is assumed to be a linear functional called "expected value" on  $\mathcal{V}$  to  $\mathcal{R}$ , and written  $\bar{v}$  (the "expected value" of  $v$ ) for any  $v \in \mathcal{V}$ . Thus for  $a, b \in \mathcal{R}$  and  $x, y \in \mathcal{V}$  we have  $\bar{a} = a$  and  $\overline{ax + by} = a\bar{x} + b\bar{y}$ . Capital English letters will designate matrices. Let  $A = (a_{ij})$ . Then the terminology "A constant" below will be used to indicate that the  $a_{ij}$  are in  $\mathcal{R}$ ; "A variable" will indicate that the  $a_{ij}$  are in  $\mathcal{V}$ . In general, capital English letters near the beginning of the alphabet will be used for constant matrices and letters near the end of the alphabet will be used for variable matrices. The expected value functional on  $\mathcal{V}$  to  $\mathcal{R}$  may be extended to a linear function on variable matrices to constant matrices by defining  $\bar{V} = (\bar{v}_{ij})$  for any variable  $V = (v_{ij})$ . The notation  $\mathcal{V}_n$  will designate the set of all variable ( $n \times 1$ ) matrices.

Elements of  $\mathcal{V}_n$  will also be called  $n$  dimensional variable vectors, or just variable vectors, if the dimension is clear from the context.

Lemma 8: Let  $A, B$  be constant and  $V$  variable. Then

$$(1) \quad (\overline{V})^* = \overline{(V^*)}$$

$$(2) \quad \overline{AVB} = A\overline{VB} \text{ if the multiplication is defined.}$$

Proof: Follows directly from the definitions and linearity of the "expected value" operator.

Definition: Let  $v$  be a variable vector. Then  $\ddagger(v) \equiv \overline{(v - \bar{v})(v - \bar{v})^*}$  and  $\ddagger(v)$  is called the variance-covariance matrix of  $v$ .

Lemma 9: Let  $A$  be constant and  $v$  be a variable vector. Then<sup>†</sup>  
 $\ddagger(Av) = A \ddagger(v) A^*$ .

Let (H) designate the hypothesis  $y$  is an  $(m \times 1)$  variable vector,  $x$  is an  $(n \times 1)$  constant vector,  $A$   $(m \times n)$  is constant,  $\bar{y} = Ax$ , and  $\ddagger(y) \equiv W = R^2$  is positive definite,  $R = R^*$ . Let  $\psi \equiv c^*x$ , for some constant  $c$   $(n \times 1)$ . We shall call an estimate  $\tilde{\psi}$  of  $\psi$  linear if  $\tilde{\psi} = g^*y$  for some constant  $g$   $(m \times 1)$ , and unbiased if  $\tilde{\psi} = \psi$ .

Theorem 6 (Gauss-Markoff): Let  $\psi$  be as above and assume (H). Then  $\psi$  has an unbiased linear estimate if and only if  $A^+Ac = c$ . In the latter case  $c^*(R^{-1}A)^+R^{-1}y \equiv \tilde{\psi}_{MV}$  is an unbiased linear estimate of  $\psi$  having minimal variance  $c^*(A^*W^{-1}A)^+c$  in the class of all unbiased linear estimates of  $\psi$ . If  $A^+A = I$ , then  $\tilde{\psi}_{MV} = c^*(A^*W^{-1}A)^{-1}A^*W^{-1}y$  and the variance of  $\tilde{\psi}_{MV}$  is  $c^*(A^*W^{-1}A)^{-1}c$ .

Proof: Suppose  $\psi$  has the unbiased linear estimate  $\tilde{\psi} = g^*y$ . Then  $c^*x = \psi = \tilde{\psi} = g^*\bar{y} = g^*Ax$  for all  $x$ ,  $c^* = g^*A$ ,  $c = A^*g$ ,  $A^+Ac = A^+AA^*g = A^*g = c$ .<sup>††</sup> Conversely suppose  $A^+Ac = c$ . Then  $c^*A^+A = c^*$  and so using Lemma 4, we have  $\tilde{\psi}_{MV} = c^*(R^{-1}A)^+R^{-1}\bar{y} = c^*(R^{-1}A)^+(R^{-1}A)x = c^*A^+Ax = c^*\psi$  so  $\tilde{\psi}_{MV}$  is clearly an unbiased linear estimate. This proves the first statement

<sup>†</sup>This may be proved by direct computation

<sup>††</sup>Consult Theorem 2 for the appropriate identity used.

of the theorem. By the first part of the proof, any unbiased linear estimate  $\tilde{\psi} \equiv g^* y$  must satisfy the property  $A^* g = c$ . Also  $\ddagger(\tilde{\psi}) = g^* \ddagger(y) g = g^* R^2 g$ . Thus the unbiased linear estimate of minimal variance is given by the solution  $A^* g = c$  such that  $\|g\|_w$  is minimized. The latter solution<sup>†</sup> is given by  $g_{MV} = R^{-1} (A^* R^{-1})^+ c$ , which yields  $\tilde{\psi}_{MV}$  as the estimate sought. Also  $\ddagger(\tilde{\psi}_{MV}) = g_{MV}^* R^2 g_{MV} = c^* (R^{-1} A)^+ R^{-1} R^2 R^{-1} (A^* R^{-1})^+ c = c^* (R^{-1} A)^+ (A^* R^{-1})^+ c^{\dagger\dagger} = c^* (A^* R^{-2} A)^+ c = c^* (A^* W^{-1} A)^+ c$ . If  $A^+ A = I$  then the columns of  $A$  are linearly independent. Then the columns of  $R^{-1} A$  are linearly independent and  $\tilde{\psi}_{MV} = c^* (R^{-1} A)^+ R^{-1} y = c^* [(R^{-1} A)^* (R^{-1} A)]^{-1} (R^{-1} A)^* R^{-1} y = c^* (A^* W^{-1} A)^{-1} A^* W^{-1} y$ , q. e. d.

Suppose we assume (H) and define  $x_{MV} \equiv (R^{-1} A)^+ R^{-1} y$ . Then  $\ddagger(x_{MV}) = (A^* W^{-1} A)^+$  using the same argument as in the proof of the above theorem. Note that  $x_{MV}$  now enjoys the following properties:

(1) If  $\psi \equiv c^* x$  has an unbiased linear estimate, then  $c^* x_{MV}$  is an unbiased linear estimate of  $\psi$  having minimal variance  $c^* \ddagger(x_{MV}) c$  in the class of all unbiased linear estimates. This follows directly from the theorem above.

(2) The vector  $x_{MV}$  is the vector of least norm such that  $S(W) \equiv (y - Ax)^* W^{-1} (y - Ax) = \|y - Ax\|_{W^{-1}}$  is minimized. This follows directly from the (least Squares) Theorem.<sup>†††</sup> If for an arbitrary positive definite matrix  $W$ ,  $x_{LS}(W)$  is defined to be the vector of least norm which minimizes  $S(W)$ , then  $x_{LS}(W)$  is called the generalized weighted least squares estimate corresponding to the weighting matrix  $W$ . Thus  $x_{MV}$  is  $x_{LS}(\ddagger(y))$ .

<sup>†</sup> See Theorem 4, Corollary 2

<sup>††</sup> Consult Theorem 2 for the appropriate identity used

<sup>†††</sup> See Theorem 4, Corollary 1



### III. PARAMETER ESTIMATION FOR ORBIT DETERMINATION

Theorem 7: Let  $v_1$  and  $v_2$  be two independent variable vectors in  $\mathcal{V}_n$  such that  $\bar{v}_1 = \bar{v}_2$  and  $V_i \equiv \mathbb{E}(v_i)$  is positive definite ( $i = 1, 2$ ). Let  $\tilde{v} \equiv (V_1^{-1} + V_2^{-1})^{-1} (V_1^{-1} v_1 + V_2^{-1} v_2)$ . Then  $\tilde{v}$  is unbiased (i. e.,  $\tilde{v} = \bar{v}_1 = \bar{v}_2$ ) and  $\tilde{v}$  has minimal variance covariance matrix  $(V_1^{-1} + V_2^{-1})^{-1}$  in the unbiased class of vectors  $Rv_1 + (I-R)v_2$  for  $R$  constant ( $n \times n$ ).

Proof: The fact that  $\mathbb{E}(\tilde{v}) = (V_1^{-1} + V_2^{-1})^{-1}$  follows by direct computation. Let  $y = Rv_1 + Sv_2$  where  $R$  is constant ( $n \times n$ ) and  $S \equiv I - R$ . Then clearly  $y$  is unbiased. We wish to show that  $\mathbb{E}(y) \geq \mathbb{E}(\tilde{v})$ , which is defined to mean  $A \equiv \mathbb{E}(y) - \mathbb{E}(\tilde{v})$  is positive semi-definite. By direct calculation  $A = RV_1 R^* + SV_2 S^* - (V_1^{-1} + V_2^{-1})^{-1}$ . Then  $A = (V_1^{-1} + V_2^{-1})^{-1} B (V_1^{-1} + V_2^{-1})^{-1}$ , where  $B = [(V_1^{-1} + V_2^{-1}) RV_1] V_1^{-1} [V_1 R^* (V_1^{-1} + V_2^{-1})] + [(V_1^{-1} + V_2^{-1}) SV_2] V_2^{-1} [V_2 S^* (V_1^{-1} + V_2^{-1})] - (V_1^{-1} + V_2^{-1})$ ,  $B = CV_1^{-1} C^* + DV_2^{-1} D^* - (V_1^{-1} + V_2^{-1})$  where  $C \equiv (V_1^{-1} + V_2^{-1}) RV_1$ ,  $D \equiv (V_1^{-1} + V_2^{-1}) SV_2$ . Thus  $B = CP^2 C^* + DQ^2 D^* - (P^2 + Q^2)$  where  $V_1^{-1} = P^2$ ,  $P = P^*$ ,  $V_2^{-1} = Q^2$ ,  $Q = Q^*$ . Let  $F \equiv (CP - P)(CP - P)^* + (DQ - Q)(DQ - Q)^*$ . Then  $F = (CP - P)(PC^* - P) + (DQ - Q)(QD^* - Q) = CP^2 C^* - CP^2 - P^2 C^* + P^2 + DQ^2 D^* - DQ^2 - Q^2 D^* + Q^2 = CP^2 C^* + DQ^2 D^* + [(P^2 + Q^2) - (CP^2 + DQ^2) - (P^2 C^* + Q^2 D^*)]$ . But  $CP^2 + DQ^2 = CV_1^{-1} + DV_2^{-1} = (V_1^{-1} + V_2^{-1})(R + S) = V_1^{-1} + V_2^{-1} = P^2 + Q^2$  and by transposing both sides,  $P^2 C^* + Q^2 D^* = P^2 + Q^2$ . Thus  $F = CP^2 C^* + DQ^2 D^* - (P^2 + Q^2) = B$  and  $B$  is positive semi-definite, since  $B = F$ , which is the sum of two terms of the form  $MM^*$  and hence positive semi-definite. Then  $A$  is also positive semi-definite since  $A$  is of the form  $MBM^*$ , q. e. d.

We shall now discuss applications of the preceding material to parameter estimation in orbit determination work.<sup>†</sup> Let  $f(a, p)$  be a vector ( $n \times 1$ ) valued function of the vector variables  $a$  ( $r \times 1$ ) and  $p$  ( $s \times 1$ );  $f(a, p)$  is the vector of observations without noise given the (vector) parameters  $a$  and  $p$ . Here  $a$  is some parameter vector we wish to estimate and determine the variance covariance matrix of the estimate;  $p$  is some parameter vector which we do not

<sup>†</sup>The general notation, necessary background and essential formulae parallel the development in Reference 5.

wish to estimate, but wish to account for the effect of an uncertainty in  $p$  on the variance-covariance of the estimate of  $a$ . We shall assume that  $a_t$  and  $p_t$  are the "true" estimates of  $a$  and  $p$ , respectively. Assume that  $f(a, p)$  is linear in the neighborhood of  $(a_t, p_t)$  so that  $f(a, p) = f(a_t, p_t) + A(a - a_t) + P(p - p_t)$  provided  $a$  and  $p$  are sufficiently close to  $a_t$  and  $p_t$ , respectively, and  $A, P$  are constant  $n \times r$  and  $n \times s$  matrices. Let  $R$  and  $\epsilon$  be  $(n \times 1)$  vector variables satisfying  $R = f(a_t, p_t) + \epsilon$ , with  $\bar{\epsilon} = 0$ . Note that  $R$  is the vector of observations with noise; the noise is assumed to have mean zero. Let  $c$  be a constant  $(r \times 1)$  vector. Then  $R - f(c, p_t) = f(a_t, p_t) - f(c, p_t) + \epsilon = A(a_t - c) + \epsilon$ ,  $\overline{R - f(c, p_t)} = A(a_t - c)$  and the generalized weighted least squares estimate for  $a_t - c$  would be given by  $(a_t - c)_{LS} = M(R - f(c, p_t)) = M(A(a_t - c) + \epsilon)$  where  $M = (A^* W^{-1} A)^{-1} A^* W^{-1}$  for the weighting matrix  $W$ .<sup>†</sup> Since  $MA = I$  we have  $(a_t)_{LS} - c = a_t - c + M\epsilon$ , and so  $(a_t)_{LS} - a_t = M\epsilon$ . Now suppose  $a_o$  is a  $(r \times 1)$  vector variable such that  $\bar{a}_o = a_t$  and  $a_o - a_t$  is independent of  $\epsilon$ . Then  $a_o - a_t$  is independent of  $(a_t)_{LS} - a_t$  and the minimal variance unbiased linear combination of the two is given by:

$$\tilde{a}_t = K(V(a_t)_{LS} + U a_o) \text{ where } V = \mathbb{E}((a_t)_{LS})^{-1}, U = \mathbb{E}(a_o)^{-1}$$

and

$$K \equiv (U + V)^{-1}. \text{ Then } \tilde{a}_t = K \left\{ V [c + M(R - f(c, p_t))] + U a_o \right\}$$

Now  $V = (M \mathbb{E}(\epsilon) M^*)^{-1}$  and if we assume  $c = a_o$  then  $\tilde{a}_t = a_o + KVM(R - f(a_o, p_t))$ . Define  $\hat{a}_t$  by

$$\begin{aligned} \hat{a}_t &\equiv a_o + KVM(R - f(a_o, p_o)) \\ &= a_o + KVM(R - f(a_t, p_t) + f(a_t, p_t) - f(a_o, p_o)) \\ &= a_o + KVM(\epsilon + A(a_t - a_o) + P(p_t - p_o)) \end{aligned}$$

<sup>†</sup>We have assumed  $A^* A$  is nonsingular, which is equivalent to  $A^+ A = I$ . See Lemma 7.

$$\begin{aligned}\hat{a}_t - a_t &= K(V+U)(a_o - a_t) + K(VM\epsilon + V(a_t - a_o) + VMP(p_t - p_o)) \\ &= K[VM\epsilon - U(a_t - a_o) + VMP(p_t - p_o)]\end{aligned}$$

$$\begin{aligned}\hat{\Phi}(KVM\epsilon) &= KVM\hat{\Phi}(\epsilon)M^*VK = KVK, \quad V = (M\hat{\Phi}(\epsilon)M^*)^{-1} = \\ &= [(A^*W^{-1}A)^{-1}A^*W^{-1}\hat{\Phi}(\epsilon)W^{-1}A(A^*W^{-1}A)^{-1}]^{-1},\end{aligned}$$

$$\hat{\Phi}(\hat{a}_t - a_t) = K[V + VMP\hat{\Phi}(p_t - p_o)P^*M^*V + \hat{\Phi}(a_o)^{-1}]K,$$

$$\hat{\Phi}(\hat{a}_t - a_t) = K + KVM\hat{\Phi}(p_t - p_o)P^*M^*VK.$$

Note that if  $W = \hat{\Phi}(\epsilon)$  then  $V = A^*W^{-1}A$ ,  $VM = A^*W^{-1}$ , and  $\hat{\Phi}(\hat{a}_t - a_t) = K + KA^*W^{-1}P\hat{\Phi}(p_t - p_o)P^*W^{-1}AK$ .

REFERENCES

1. R. Penrose, A Generalized Inverse for Matrices, Proc. Cambridge Philos. Soc., 51, pp 406-413, 1955.
2. R. Penrose, On Best Approximate Solutions of Linear Matrix Equations, Proc. Cambridge Philos. Soc., 52, pp 17-19, 1956.
3. T. N. E. Greville, The Pseudoinverse of a Rectangular or Singular Matrix and its Application to the Solution of Systems of Linear Equations, SIAM Review, pp 38-43, January 1959.
4. Henry Scheffe, The Analysis of Variance, John Wiley, New York, 1959.
5. L. Wong, A. S. Liu, M. C. Fujisaki, O. Senda, Computer Program Guide, Tracking Accuracy and Prediction Program, Space Technology Laboratories, 8976, 6005-NU-000.

UNCLASSIFIED

Aerospace Corporation, El Segundo, California. THE MATRIX PSEUDOINVERSE AND MINIMAL VARIANCE ESTIMATES by C. M. Price, 25 March 1963. 18 pages. (Report TDR-169(3550-01)TN-2, SSD-TDR-63-92) Contract No. AF04(695)-169.

Unclassified Report

This paper reviews and applies certain results concerning the matrix pseudoinverse to the general theory of estimable functions and minimal variance estimates. The paper is divided into two sections. The first section reviews and extends certain known results concerning the matrix pseudoinverse. This section is essentially nonstatistical. The second section uses results in the first section to state and prove a generalized version of the Gauss-Markoff Theorem concerning unbiased linear estimates having minimal variance. In the third section, an additional theorem is proven, which together with

(over)

UNCLASSIFIED

Aerospace Corporation, El Segundo, California. THE MATRIX PSEUDOINVERSE AND MINIMAL VARIANCE ESTIMATES by C. M. Price, 25 March 1963. 18 pages. (Report TDR-169(3550-01)TN-2, SSD-TDR-63-92) Contract No. AF04(695)-169.

Unclassified Report

This paper reviews and applies certain results concerning the matrix pseudoinverse to the general theory of estimable functions and minimal variance estimates. The paper is divided into two sections. The first section reviews and extends certain known results concerning the matrix pseudoinverse. This section is essentially nonstatistical. The second section uses results in the first section to state and prove a generalized version of the Gauss-Markoff Theorem concerning unbiased linear estimates having minimal variance. In the third section, an additional theorem is proven, which together with

(over)

UNCLASSIFIED

Aerospace Corporation, El Segundo, California. THE MATRIX PSEUDOINVERSE AND MINIMAL VARIANCE ESTIMATES by C. M. Price, 25 March 1963. 18 pages. (Report TDR-169(3550-01)TN-2, SSD-TDR-63-92) Contract No. AF04(695)-169.

Unclassified Report

This paper reviews and applies certain results concerning the matrix pseudoinverse to the general theory of estimable functions and minimal variance estimates. The paper is divided into two sections. The first section reviews and extends certain known results concerning the matrix pseudoinverse. This section is essentially nonstatistical. The second section uses results in the first section to state and prove a generalized version of the Gauss-Markoff Theorem concerning unbiased linear estimates having minimal variance. In the third section, an additional theorem is proven, which together with

(over)

UNCLASSIFIED

Aerospace Corporation, El Segundo, California. THE MATRIX PSEUDOINVERSE AND MINIMAL VARIANCE ESTIMATES by C. M. Price, 25 March 1963. 18 pages. (Report TDR-169(3550-01)TN-2, SSD-TDR-63-92) Contract No. AF04(695)-169.

Unclassified Report

This paper reviews and applies certain results concerning the matrix pseudoinverse to the general theory of estimable functions and minimal variance estimates. The paper is divided into two sections. The first section reviews and extends certain known results concerning the matrix pseudoinverse. This section is essentially nonstatistical. The second section uses results in the first section to state and prove a generalized version of the Gauss-Markoff Theorem concerning unbiased linear estimates having minimal variance. In the third section, an additional theorem is proven, which together with

(over)

UNCLASSIFIED

the preceding material provides a theoretical foundation for parameter estimation in orbit determination work. This foundation is then exploited to provide formulae for such parameters.

UNCLASSIFIED

UNCLASSIFIED

the preceding material provides a theoretical foundation for parameter estimation in orbit determination work. This foundation is then exploited to provide formulae for such parameters.

UNCLASSIFIED

UNCLASSIFIED

the preceding material provides a theoretical foundation for parameter estimation in orbit determination work. This foundation is then exploited to provide formulae for such parameters.

UNCLASSIFIED

UNCLASSIFIED

the preceding material provides a theoretical foundation for parameter estimation in orbit determination work. This foundation is then exploited to provide formulae for such parameters.

UNCLASSIFIED

UNCLASSIFIED

Aerospace Corporation, El Segundo, California.  
THE MATRIX PSEUDOINVERSE AND MINIMAL  
VARIANCE ESTIMATES by C. M. Price, 25 March  
1963. 18 pages. (Report TDR-169(3550-01)TN-2,  
SSD-TDR-63-92) Contract No. AF04(695)-169.

Unclassified Report

This paper reviews and applies certain results concerning the matrix pseudoinverse to the general theory of estimable functions and minimal variance estimates. The paper is divided into two sections. The first section reviews and extends certain known results concerning the matrix pseudoinverse. This section is essentially nonstatistical. The second section uses results in the first section to state and prove a generalized version of the Gauss-Markoff Theorem concerning unbiased linear estimates having minimal variance. In the third section, an additional theorem is proven, which together with

(over)

UNCLASSIFIED

UNCLASSIFIED

Aerospace Corporation, El Segundo, California.  
THE MATRIX PSEUDOINVERSE AND MINIMAL  
VARIANCE ESTIMATES by C. M. Price, 25 March  
1963. 18 pages. (Report TDR-169(3550-01)TN-2,  
SSD-TDR-63-92) Contract No. AF04(695)-169.

Unclassified Report

This paper reviews and applies certain results concerning the matrix pseudoinverse to the general theory of estimable functions and minimal variance estimates. The paper is divided into two sections. The first section reviews and extends certain known results concerning the matrix pseudoinverse. This section is essentially nonstatistical. The second section uses results in the first section to state and prove a generalized version of the Gauss-Markoff Theorem concerning unbiased linear estimates having minimal variance. In the third section, an additional theorem is proven, which together with

(over)

UNCLASSIFIED

UNCLASSIFIED

Aerospace Corporation, El Segundo, California.  
THE MATRIX PSEUDOINVERSE AND MINIMAL  
VARIANCE ESTIMATES by C. M. Price, 25 March  
1963. 18 pages. (Report TDR-169(3550-01)TN-2,  
SSD-TDR-63-92) Contract No. AF04(695)-169.

Unclassified Report

This paper reviews and applies certain results concerning the matrix pseudoinverse to the general theory of estimable functions and minimal variance estimates. The paper is divided into two sections. The first section reviews and extends certain known results concerning the matrix pseudoinverse. This section is essentially nonstatistical. The second section uses results in the first section to state and prove a generalized version of the Gauss-Markoff Theorem concerning unbiased linear estimates having minimal variance. In the third section, an additional theorem is proven, which together with

(over)

UNCLASSIFIED

UNCLASSIFIED

Aerospace Corporation, El Segundo, California.  
THE MATRIX PSEUDOINVERSE AND MINIMAL  
VARIANCE ESTIMATES by C. M. Price, 25 March  
1963. 18 pages. (Report TDR-169(3550-01)TN-2,  
SSD-TDR-63-92) Contract No. AF04(695)-169.

Unclassified Report

This paper reviews and applies certain results concerning the matrix pseudoinverse to the general theory of estimable functions and minimal variance estimates. The paper is divided into two sections. The first section reviews and extends certain known results concerning the matrix pseudoinverse. This section is essentially nonstatistical. The second section uses results in the first section to state and prove a generalized version of the Gauss-Markoff Theorem concerning unbiased linear estimates having minimal variance. In the third section, an additional theorem is proven, which together with

(over)

UNCLASSIFIED

UNCLASSIFIED

UNCLASSIFIED

the preceding material provides a theoretical foundation for parameter estimation in orbit determination work. This foundation is then exploited to provide formulae for such parameters.

UNCLASSIFIED

UNCLASSIFIED

the preceding material provides a theoretical foundation for parameter estimation in orbit determination work. This foundation is then exploited to provide formulae for such parameters.

UNCLASSIFIED

UNCLASSIFIED

the preceding material provides a theoretical foundation for parameter estimation in orbit determination work. This foundation is then exploited to provide formulae for such parameters.

UNCLASSIFIED

UNCLASSIFIED

the preceding material provides a theoretical foundation for parameter estimation in orbit determination work. This foundation is then exploited to provide formulae for such parameters.