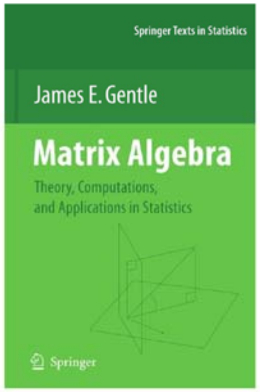
Theory of Matrices

Chapter 3. Basic Properties of Matrices

3.5. Linear Systems of Equations—Proofs of Theorems



Theorem 3.5.1

Theorem 3.5.1. If $Ax = b$ is an underdetermined system then there are an infinite number of solutions to the system.

Proof. If $Ax = b$ is an underdetermined system with A $n \times m$ then, since it is consistent by definition, there is a solution x_1 such that $Ax_1 = b$. Since $\text{rank}(A) < m$ and A has m columns, then by Exercise 2.1 the columns of A are not linearly independent. So with $a_1, a_2, \ldots, a_m$ as the columns of A, there are scalars $s_1, s_2, \ldots, s_m$ not all 0 for which $s_1a_1 + s_2a_2 + \cdots + s_ma_m = 0$. Let $s \in \mathbb{R}^m$ have components s_i and define $x_2 = s + x_1$. Then $Ax_2 = A(s + x_1) = As + Ax_1 = 0 + b = b$ and so x_2 is also a solution to the system $Ax = b$. Now let $w \in \mathbb{R}$ and consider $x_w = wx_1 + (1 - w)x_2$. We have

$$Ax_w = A(wx_1 + (1-w)x_2) = wAx_1 + (1-w)Ax_2 = wb + (1-w)b = b$$

and each x_w is a solution to $Ax = b$. Therefore, $Ax = b$ has an infinite number of solutions. □

Theorem 3.5.2

Theorem 3.5.2. Properties of the Generalized Inverse.

(1) If A^- is a generalized inverse of A then $(A^-)^T$ is a generalized inverse of A^T.

(2) $(A^-A)(A^-A) = A^-A$; that is, A^-A is idempotent.

(3) $\text{rank}(A^-A) = \text{rank}(A)$.

(4) $(I - A^-A)(A^-A) = 0$ and $(I - A^-A)(I - A^-A) = (I - A^-A)$.

(5) $\text{rank}(I - A^-A) = m - \text{rank}(A)$ where A is $n \times m$.

Proof. (1) We have $A = AA^-A$ so, by Theorem 3.2.1(1), $A^T = (AA^-A)^T = A^T(A^-)^TA^T$ and so $(A^-)^T$ is a generalized inverse of A^T.
(2) Since $A = AA^-A$ then $A^-A = A^-AA^-A = (A^-A)(A^-A)$.
(3) By Theorem 3.3.5, $\text{rank}(A^-A) \le \min\{\text{rank}(A^-), \text{rank}(A)\} \le \text{rank}(A)$. Since $A = AA^-A$ then again by Theorem 3.3.5, $\text{rank}(A) \le \min\{\text{rank}(A), \text{rank}(A^-A)\} \le \text{rank}(A^-A)$, and so $\text{rank}(A) = \text{rank}(A^-A)$.

Theorem 3.5.2 (continued)

Proof (continued). (4) We have $(I - A^-A)(A^-A) = IA^-A - A^-AA^-A = A^-A - A^-(AA^-A) = A^-A - A^-A = 0$. So $(I - A^-A)(I - A^-A) = I - A^-A - (I - A^-A)A^-A = I - A^-A - 0 = I - A^-A$.

(5) Notice that A^-A is $m \times m$ and by Part (4) $(I - A^-A)A^-A = 0$, so

$$\begin{aligned} 0 &= \text{rank}(0) = \text{rank}((I - A^-A)A^-A) \\ &\ge \text{rank}(I - A^-A) + \text{rank}(A^-A) - m \text{ by Theorem 3.3.15} \\ &= \text{rank}(I - A^-A) + \text{rank}(A) - m \text{ by Part (3)} \qquad (*) \end{aligned}$$

Next, $I = I - A^-A + A^-A$ and by Theorem 3.3.6,

$$\begin{aligned} m &= \text{rank}(I) = \text{rank}(I - A^-A + A^-A) \le \text{rank}(I - A^-A) + \text{rank}(A^-A) \\ &= \text{rank}(I - A^-A) + \text{rank}(A) \text{ by Part (3).} \qquad (**) \end{aligned}$$

Combining $(*)$ and $(**)$ gives $m = \text{rank}(I - A^-A) + \text{rank}(A)$ and the claim follows. □

Theorem 3.5.3

Theorem 3.5.3. Let $Ax = b$ be a consistent system of equations and let A^- be a generalized inverse of A.

(1) $x = A^-b$ is a solution.

(2) If $x = Gb$ is a solution of system $Ax = b$ for *all* b such that a solution exists, then $AGA = A$; that is, G is a generalized inverse of A.

(3) For any $z \in \mathbb{R}^m$, $A^-b + (I - A^-A)z$ is a solution.

(4) Every solution is of the form $x = A^-b + (I - A^-A)z$ for some $z \in \mathbb{R}^m$.

Proof. (1) We have $(AA^-A)x = Ax$ and with $Ax = b$ as the given system, we get $AA^-(Ax) = Ax$ or $AA^-b = b$ or $A(A^-b) = b$; that is, A^-b is a solution to $Ax = b$.

Theorem 3.5.3 (continued)

Proof (continued). (2) Let the columns of A be $a_1, a_2, \ldots, a_m$. The m systems $Ax = a_j$ (where $1 \le j \le n$) each have a solution (namely, the jth unit vector in $\mathbb{R}^m$). So by hypothesis, Ga_j is a solution of the system $Ax = a_j$ for each j (where $1 \le j \le n$). That is, $AGa_j = a_j$ for $1 \le j \le n$, or $AGA = A$.

(3) We have

$$\begin{aligned} A(A^-b + (I - A^-A)z) &= AA^-b + (A - AA^-A)z \\ &= b + (A - A)z \text{ by Part (1)} \\ &= b + 0 = b. \end{aligned}$$

(4) Let y be a solution of $Ax = b$. Then

$$\begin{aligned} y &= A^-b - A^-b + y = A^-b - A^-(Ay) + y \text{ since } Ay = b \\ &= A^-b - (A^-A - I)y = A^-b + (I - A^-A)z \text{ with } z = y. \end{aligned}$$

□

Theorem 3.5.4

Theorem 3.5.4. The nullity of $n \times m$ matrix A satisfies $\dim(\mathcal{N}(A)) = m - \text{rank}(A)$.

Proof. If $x \in \mathcal{N}(A)$ then $Ax = 0$ and by Theorem 3.5.3 (3 and 4) $x = 0 + (I - A^-A)z = (I - A^-A)z$ for any $z \in \mathbb{R}^m$ (and conversely every solution to $Ax = 0$ is of this form). Now $(I - A^-A)z$ is in the column space of $I - A^-A$ for every $z \in \mathbb{R}^m$, so by Theorem 3.5.2(5),

$$\dim(\mathcal{N}(A)) = \text{rank}(I - A^-A) = m - \text{rank}(A).$$

□

Theorem 3.5.5

Theorem 3.5.5.

(1) If system $Ax = b$ is consistent, then any solution is of the form $x = A^-b + z$ for some $z \in \mathcal{N}(A)$.

(2) For matrix A, the null space of A is orthogonal to the row space of A: $\mathcal{N}(A) \perp \mathcal{V}(A^T)$.

(3) For matrix A, $\mathcal{N}(A) \oplus \mathcal{V}(A^T) = \mathbb{R}^m$.

Proof. (1) Let y be a solution of $Ax = b$. Then $Ay = b = AA^-b$ by Theorem 3.5.3(1) and so $Ay - AA^-b = A(y - A^-b) = 0$. Therefore $z = y - A^-b \in \mathcal{N}(A)$. So $y = A^-b + z$ where $z \in \mathcal{N}(A)$.

(2) Let $a \in \mathcal{V}(A^T)$ and $b \in \mathcal{N}(A)$. Then

$$\begin{aligned} \langle b, a \rangle &= b^T a \\ &= b^T A^T s \text{ since } a \in \mathcal{V}(A^T) \text{ then } a = A^T s \text{ for some } s \in \mathbb{R}^n \\ &= (b^T A^T)s = (Ab)^T s \text{ by Theorem 3.2.1(1)} \\ &= 0s = 0 \text{ since } b \in \mathcal{N}(A). \end{aligned}$$

Theorem 3.5.5 (continued)

Theorem 3.5.5.

(1) If system $Ax = b$ is consistent, then any solution is of the form $x = A^{-}b + z$ for some $z \in \mathcal{N}(A)$.

(2) For matrix A, the null space of A is orthogonal to the row space of A: $\mathcal{N}(A) \perp \mathcal{V}(A^T)$.

(3) For matrix A, $\mathcal{N}(A) \oplus \mathcal{V}(A^T) = \mathbb{R}^m$.

Proof (continued). So $a \perp b$. Since a is an arbitrary element of $\mathcal{V}(A^T)$ and b is an arbitrary element of $\mathcal{N}(A)$ then $\mathcal{N}(A) \perp \mathcal{V}(A^T)$.

(3) From Theorem 3.5.4 (the rank-nullity equation), $\dim(\mathcal{N}(A)) + \dim(\mathcal{V}(A^T)) = m$. Now both $\mathcal{N}(A)$ and $\mathcal{V}(A^T)$ are subspaces of $\mathbb{R}^m$, so $\mathcal{N}(A) \oplus \mathcal{V}(A^T)$ is a m dimensional subspace of $\mathbb{R}^m$. That is, $\mathcal{N}(A) \oplus \mathcal{V}(A^T) = \mathbb{R}^m$ (technically, we need the Fundamental Theorem of Finite Dimensional Vector Spaces here). □