Section 3.5. Linear Systems of Equations

Note. Recall that we represent a system of n linear equations in m unknowns as $Ax = b$ where A is an $n \times m$ matrix, x is an m-vector of unknowns and b is an n-vector.

Definition. The system of equations $Ax = b$ where $b = 0$ is *homogeneous*.

Definition. If there is a solution x to the system $Ax = b$ then the system is *consistent*. If there is no solution then the system is *inconsistent*.

Note 3.5.A. In the system $Ax = b$, if A is invertible (and hence square) then there is a unique solution to the system, namely $x = A^{-1}b$.

Definition. A consistent system $Ax = b$ of n equations in m unknowns (so A is $n \times m$) is *underdetermined* if $\text{rank}(A) < m$.

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

Note. Since an underdetermined system of equations has multiple solutions, we might want to choose a "best" solution in some way that optimizes additional desirable properties.

Definition. A system $Ax = b$ of n equations in m unknowns where $n > m$ and $\text{rank}([A \mid b]) > \text{rank}(A)$ is *overdetermined*.

Note. In an overdetermined system, the condition $\text{rank}([A \mid b]) > \text{rank}(A)$ implies that b is not in the column space of A and so there is no solution of the system $Ax = b$. In an overdetermined system, we might seek an approximate solution; one approach to this is the method of least squares to be explored in Section 9.2.

Definition. A matrix G such that $AGA = A$ is a *generalized inverse* of A, denoted $G = A^-$.

Note. If A is $n \times m$ then a generalized inverse of A must be $m \times n$. If A is nonsingular (and hence is square and of full rank) then $A^- = A^{-1}$. In Section 3.6 we'll see that an A^- exists for any matrix A.

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) $(\mathcal{I} - A^-A)(A^-A) = 0$ and $(\mathcal{I} - A^-A)(\mathcal{I} - A^-A) = (\mathcal{I} - A^-A)$.

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

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

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

Note. Gentle's statement of (2) in Theorem 3.5.3 is not correct (see page 99). The correct statement given here is Exercise 7 from Section 3 of the Introduction (on page 3) of A. Ben-Isreal and T. Greville's *Generalized Inverses: Theory and Applications*, 2nd Edition, Springer (2003).

Note. From Theorem 3.5.3 (3 and 4) we see that the number of linearly independent solutions to $Ax = b$ (for a given b) is the rank of $\mathcal{I} - A^-A$ which by Theorem 3.5.2(5) is $m - \text{rank}(A)$.

Definition. For $n \times m$ matrix A, the set of all vectors generated by all solutions x of the homogeneous system $Ax = 0$ is the *null space* of A, denoted $\mathcal{N}(A)$. The dimension of the null space is the *nullity* of A.

Note. By the definition of vector space (see Section 2.1) we need only show that for $x, y \in \mathcal{N}(A)$ and for $a, b \in \mathbb{R}$ we have $ax + by \in \mathcal{N}(A)$ (which is "clear"). See also Theorem 2.2.2.

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

Note. The result of Theorem 3.5.4 can be rearranged to give the rank-nullity equation:

$$\text{rank}(A) + \text{nullity}(A) = \#\text{ columns of } A.$$

Note. By Theorem 3.3.5, for square matrix A,

$$\text{rank}(A) \geq \text{rank}(A^2) \geq \text{rank}(A^3) \geq \cdots.$$

So by Theorem 3.5.4,

$$\dim(\mathcal{N}(A)) \leq \dim(\mathcal{N}(A^2)) \leq \dim(\mathcal{N}(A^3)) \leq \cdots.$$

In fact, $\mathcal{N}(A) \subset \mathcal{N}(A^2) \subset \mathcal{N}(A^3) \subset \cdots$ since for any x with $A^i x = 0$ we also have $A^{i+1}x = A(A^i x) = A0 = 0$.

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

Note. Since $\mathcal{N}(A) \subset \mathcal{N}(A^2) \subset \mathcal{N}(A^3) \subset \cdots$ and $\mathcal{N}(A) \oplus \mathcal{V}(A^T) = \mathbb{R}^m$, then row spaces satisfy $\mathcal{V}(A^T) \supset \mathcal{V}((A^T)^2) \supset \mathcal{V}((A^T)^3) \subset \cdots$.

Note. By Theorem 3.5.5(1), the general solution to consistent system $Ax = b$ is $x = A^- b + z$ where A^- is some general inverse of A and z is any element of $\mathcal{N}(A)$. Here, $A^- b$ is a "particular solution" to $Ax = b$ and the "general solution" to $Ax = 0$ is $\mathcal{N}(A)$. We treat $A^- b$ as a "translation vector" and then the general solution of $Ax = b$ can be expressed as a *translation of the null space*: $A^- b + \mathcal{N}(A)$. Of course, $A^- b + \mathcal{N}(A)$ is not itself a vector space unless $A^- b \in \mathcal{N}(A)$ in which case it must be that $b = 0$. In the language of Section 2.2, $A^- b + \mathcal{N}(A)$ is a "flat" or "affine space" in $\mathbb{R}^m$.

Revised: 6/28/2018