Math 31AH Fall 2011
Solving Equations

Theorem (2.2.1)**.** *Given a system $A\mathbf{x} = \mathbf{b}$ of m linear equations in n unknowns represented by the $m \times (n+1)$ augmented matrix $[A \mid \mathbf{b}]$ with echelon form $[\widetilde{A} \mid \widetilde{\mathbf{b}}]$.*

1. *If $\widetilde{\mathbf{b}}$ contains a pivotal 1, the system has no solutions.*
2. *If $\widetilde{\mathbf{b}}$ does not contain a pivotal 1, then*
   (a) *If each column of $\widetilde{A}$ has a pivotal 1, the system has a unique solution.*
   (b) *If at least one column of $\widetilde{A}$ does not have a pivotal 1, the system has infinitely many solutions which may be found by assigning arbitrary values to the the unknowns corresponding to the nonpivotal columns of $\widetilde{A}$, which then uniquely determine the values of the unknowns corresponding to the pivotal columns of $\widetilde{A}$.*

Theorem 2.2.1 was proven in class (as well as in the book). The following corollary follows immediately.

Corollary (2.2.2)**.** *A system $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b}$ if and only if A row reduces to the identity.*

Proof.

$\Longleftarrow$) If A row reduces to the identity, then $\widetilde{A}$ has a pivotal 1 in each row and each column. In particular, A is an $n \times n$ square matrix, and $\widetilde{\mathbf{b}}$ does not have a pivotal 1 (since it is the $(n+1)^{\text{st}}$ column of $[A \mid \mathbf{b}]$). By Theorem 2.2.2:2.(a), the system $A\mathbf{x} = \mathbf{b}$ has a unique solution.

$\Longrightarrow$) If A does not row reduce to the identity, then either i) $m > n$, every column of $\widetilde{A}$ has a pivotal 1, and $A\mathbf{x} = \mathbf{e}_m$ will have no solution or ii) at least one column of A does not have a pivotal 1, in which case either $\widetilde{\mathbf{b}}$ has a pivotal 1 and the system has no solutions (by Theorem 2.2.1:1), or $\widetilde{\mathbf{b}}$ does not have a pivotal 1 and the system has infinitely many solutions (by Theorem 2.2.1:2.(b)).

□

Theorem (2.1.7.2)**.** *Given an $m \times n$ matrix A, its echelon form $\widetilde{A}$ is unique.*

Note: This is clear if A can be row reduced to the identity. In that case, every column of A is a pivotal column and every system $A\mathbf{x} = \mathbf{b}$ has a unique solution, by Theorem 2.2.1:2(a). If A had another echelon form, at least one of its columns would not have a pivotal 1 (otherwise it would be the identity), and by Theorem 2.2.1, a system of the form $A\mathbf{x} = \mathbf{b}$ would have either no solution or infinitely many solutions, which would contradict Theorem 2.1.2 (which asserts that the solutions to $A\mathbf{x} = \mathbf{b}$ and $\widetilde{A}\mathbf{x} = \widetilde{\mathbf{b}}$ are the same whenever the augmented matrix $[\widetilde{A} \mid \widetilde{\mathbf{b}}]$ can be obtained from $[A \mid \mathbf{b}]$ by row operations).

Proof. Write $A = [\mathbf{a}_1, \ldots, \mathbf{a}_n]$, where $\mathbf{a}_i$ is the i^{th} column of A. For each $k = 1, \ldots, n$, write $A[k] = [\mathbf{a}_1, \ldots, \mathbf{a}_k]$, the submatrix of A consisting of the first k columns of A. Observe that any sequence of row operations that reduces A to an echelon form $\widetilde{A}$ will also reduce $A[k]$ to an echelon form $\widetilde{A}[k]$, for every $k = 1, \ldots, n$.

Interpret $A[k]$ as the augmented matrix corresponding to the system of equations

$$x_1\mathbf{a}_1 + \cdots + x_{k-1}\mathbf{a}_{k-1} = \mathbf{a}_k. \tag{1}$$

$\square$

Note that equation (1) depends only on A, and not on $\widetilde{A}$. By Theorem 2.2.1, the system of equations (1) has no solutions if and only if the last column of $\widetilde{A}[k]$, that is, the k^{th} column of $\widetilde{A}$, is pivotal. (Note: When $k = 1$, we interpret the empty sum on the left as being the zero vector in $\mathbb{R}^m$, so $\mathbf{a}_1$ is pivotal if and only if it is nonzero.) Thus, A determines the pivotal columns of $\widetilde{A}$.

By Theorem 2.1.2, the solutions to equation (1) are the same as the solutions to

$$x_1\widetilde{\mathbf{a}}_1 + \cdots + x_{k-1}\widetilde{\mathbf{a}}_{k-1} = \widetilde{\mathbf{a}}_k, \tag{2}$$

where $\widetilde{\mathbf{a}}_i$ is the i^{th} column of $\widetilde{A}$. For each $k = 1, \ldots, n$, if the k^{th} column of $\widetilde{A}$ is pivotal, it is just $\mathbf{e}_k$, the k^{th} standard basis vector. If the k^{th} column of $\widetilde{A}$ is nonpivotal, then equation (1) (and equation (2)) has a unique solution for every choice of values for the variables corresponding to the nonpivotal columns among the first $k-1$ columns of $\widetilde{A}$, by Theorem 2.2.1. In particular, equation (1) (and equation (2)) has a unique solution with the nonpivotal variables set to zero. Thus, the k^{th} column of $\widetilde{A}$ is given by $x_1\mathbf{e}_1 + \cdots + x_j\mathbf{e}_j = \widetilde{\mathbf{a}}_k$, where $x_1, \ldots, x_j$ are the unique values of the pivotal variables that solve equation (1) (and equation (2)) with the nonpivotal variables set to zero.

Thus, each column of $\widetilde{A}$ is uniquely specified by the system of equations (1) (one for every k), hence by A.