

Consider the equation

$$A\mathbf{x} = \mathbf{b} \tag{1.1}$$

where A is an $n \times n$ matrix, $\mathbf{x}, \mathbf{b} \in \mathbb{R}^n$. In particular, consider also the following theorem regarding (1.1):

$A\mathbf{x} = \mathbf{0}$ has only the trivial solution $\mathbf{x} = \mathbf{0}$ if and only if (1.1) has a solution $\mathbf{x} \in \mathbb{R}^n$ for every $\mathbf{b} \in \mathbb{R}^n$.

Proof and other Pedagogy: Recall that the proof of this equivalence goes something like this: If (1.1) is consistent for every choice of $\mathbf{b} \in \mathbb{R}^n$, then we can solve systems $A\mathbf{x}_i = \mathbf{b}_i$, $i = 1, \dots, n$, with the \mathbf{b}_i 's being the relevant columns of the identity matrix I . Then the matrix $C = [\mathbf{x}_1 | \dots | \mathbf{x}_i | \dots | \mathbf{x}_n]$ certainly turns out to be a "right inverse" of A , which, by theorem 1.6.3, will also be "the inverse" of A , so that $A\mathbf{x} = \mathbf{0}$ has only the solution $\mathbf{x} = I\mathbf{x} = (A^{-1}A)\mathbf{x} = A^{-1}(A\mathbf{x}) = A^{-1}\mathbf{0} = \mathbf{0}$. [Note that theorem 1.6.3 could itself be proved there by showing that this right inverse is itself invertible—consider the system $C\mathbf{x} = \mathbf{0}$, which then has only the solution $\mathbf{x} = I\mathbf{x} = (AC)\mathbf{x} = A(C\mathbf{x}) = A\mathbf{0} = \mathbf{0}$, which, according to equivalent statements, gives C invertible—and then using that, under this circumstance, $AC = I \Rightarrow A = C^{-1} \Rightarrow CA = CC^{-1} = I$, which says (among other things) that C is also A 's inverse.]

To get the other part of this equivalence, and to illuminate how some of the other equivalent statements just used are indeed equivalent, recall the following: if $A\mathbf{x} = \mathbf{0}$ has only the trivial solution $\mathbf{x} = \mathbf{0}$, then row reduction of $[A | \mathbf{0}]$ must give $[I | \mathbf{0}]$ (which presumes that row reduction does not change the solution space, which we have never really proved but seems clear), i.e. row reduction of A must give I (ignoring the last columns in the above augmented matrices), which shows (with theorem 1.5.1) that the product of A with a finite number of elementary matrices is I , which then shows that A can be expressed as a product of (the inverses of the) elementary matrices, and, so, is itself invertible, giving (finally!) that $A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = I\mathbf{b} = \mathbf{b}$ for every $\mathbf{b} \in \mathbb{R}^n$, so that (1.1) has at least the solution $\mathbf{x} = A^{-1}\mathbf{b}$ for any specified \mathbf{b} .

Aside from row reduction preserving the solution space of a system of equations, the other “big idea” that may be buried in here is that fact that elementary matrices, or, more to the point, elementary row operations, are “truly” invertible, i.e. that the left or right inverses of such are in fact also right and left inverses. This last statement formed in terms of elementary row operations is the following: not only is it the case that for every elementary row operation there is another one that will “undo” it “afterwards”, but that same “afterwards inverse” done “before” the given elementary row operation will itself be undone by the given elementary row operation. Of course this distinction of “before” and “after” is at the heart of what we mean by “right” and “left” inverses. Whew!

Note that at one very basic level, the truth of all of our equivalent statements comes down to row operations, specifically that they don't alter the solution space of a system of equations, and (very much related) that they are “before/after”= “left/right” invertible. Perhaps these last two (but certainly related) claims should be thoroughly investigated by the serious student.