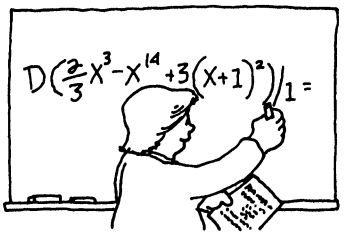
CLASSROOM CAPSULES

EDITOR

Thomas A. Farmer
Department of Mathematics and Statistics
Miami University
Oxford, OH 45056-1641



A Classroom Capsule is a short article that contains a new insight on a topic taught in the earlier years of undergraduate mathematics. Please submit manuscripts prepared according to the guidelines on the inside front cover to Tom Farmer.

A Picture Is Worth a Thousand Words

J. B. Thoo (jb2@ms.yuba.cc.ca.us), Yuba College, Marysville, CA 95901-7699

Four fundamental subspaces are associated with every $m \times n$ real matrix A: the null space of A and the column space of A^T are subspaces of $\mathbf{R}^n$; the null space of A^T and the column space of A are subspaces of $\mathbf{R}^m$. Strang [**3–5**] depicts these subspaces as pairs of orthogonal planes that he uses to illustrate the "true action of A times $\mathbf{x}$: row space to column space, null space to zero."

Here we will be exploring a slightly different diagram that has the same goal as Strang's and which I first encountered in my sophomore linear algebra course at UC–Berkeley, taught by Professor Beresford Parlett. I offer several examples of how the diagram helps explain some rudimentary concepts in linear algebra.

And here—drum roll, please—is our basic diagram: Figure 1. Note that $\mathcal{R}(\cdot)$ and $\mathcal{N}(\cdot)$ denote the column space and the null space, respectively.

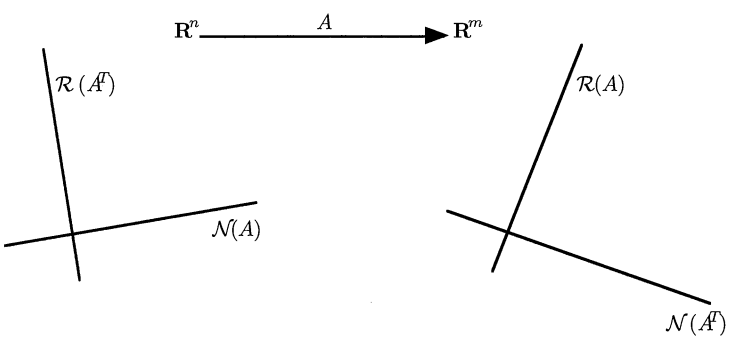


Figure 1. Decomposition of $\mathbf{R}^n$ and $\mathbf{R}^m$ into direct sums of null spaces and column spaces.

Example 1. By Figure 1 we want to suggest that $\mathcal{R}(A^T) \oplus \mathcal{N}(A) = \mathbf{R}^n$, $\mathcal{R}(A) \oplus \mathcal{N}(A^T) = \mathbf{R}^m$, and that the subspaces are orthogonal complements. Once it has

been established that rank A^T = rank A (by counting pivots, for example), the rank-nullity theorem is then apparent: rank A + null $A = n$.

If orthogonality has not made its appearance in the course at this time, it could be briefly and intuitively presented here. In fact, many students take a course in vector calculus before their first course in linear algebra, so would have seen orthogonality vis-à-vis the dot product. Those students, certainly, should be convinced from $A\mathbf{x} = \mathbf{0}$ that the rows of A are orthogonal to each of its null vectors, and from $A^T\mathbf{y} = \mathbf{0}$ that the columns of A are orthogonal to each of its left null vectors.

Example 2. Given the matrix A, if we introduce a vector $\mathbf{x}$ in $\mathbf{R}^n$, we argue by Figure 2 that $\mathbf{x}$ has a unique row space component, $\mathbf{x}_{\text{row}}$, and a unique null space component, $\mathbf{x}_{\text{null}}$. That is, $\mathbf{x} = \mathbf{x}_{\text{row}} + \mathbf{x}_{\text{null}}$. These components are obtained by finding the orthogonal projections of $\mathbf{x}$ in $\mathcal{R}(A^T)$ and $\mathcal{N}(A)$. Details will come later in the course.

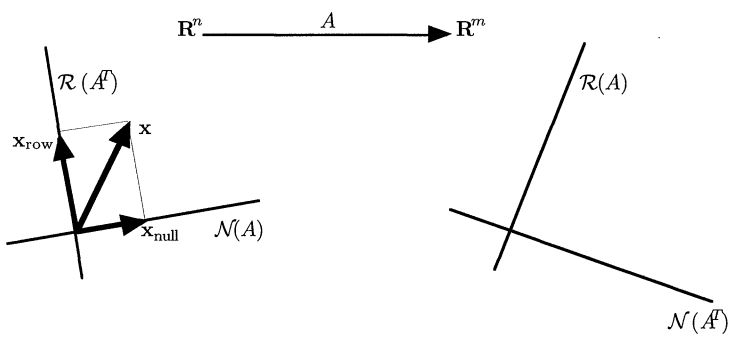


Figure 2. Every vector in $\mathbf{R}^n$ has a row space component and a null space component.

Example 3. Since $A\mathbf{u} = 0$ for every $\mathbf{u}$ in $\mathcal{N}(A)$, it follows that

$$A\mathbf{x} = A(\mathbf{x}_{\text{row}} + \mathbf{x}_{\text{null}}) = A\mathbf{x}_{\text{row}}.$$

That is, $\mathbf{x}_{\text{row}}$ is mapped into $\mathcal{R}(A)$ and $\mathbf{x}_{\text{null}}$ is mapped into $\mathbf{0}$ in $\mathbf{R}^m$. Hence, as we see in Figure 3, for $\mathbf{b}$ in $\mathcal{R}(A)$, $A\mathbf{x} = \mathbf{b}$ has a unique solution if and only if $\mathcal{N}(A)$

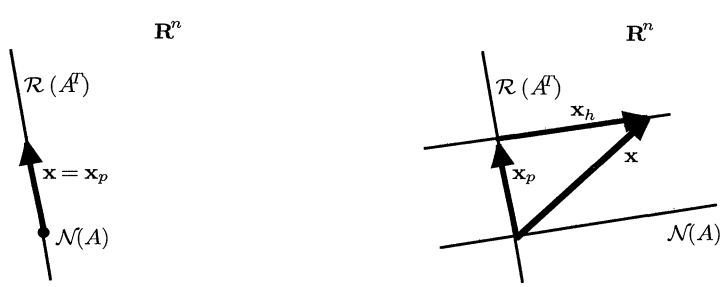


Figure 3. *Left:* $\mathbf{x} = \mathbf{x}_p$ is the unique solution to $A\mathbf{x} = \mathbf{b}$. *Right:* varying $\mathbf{x}_h$ in $\mathcal{N}(A)$ obtains different solutions $\mathbf{x}$.

contains only the zero vector. And if $\mathcal{N}(A)$ is not trivial, then the solution set to $A\mathbf{x} = \mathbf{b}$ is the translation of $\mathcal{N}(A)$ by a particular solution (there is only one in the row space). The general solution to $A\mathbf{x} = \mathbf{b}$ is $\mathbf{x} = \mathbf{x}_h + \mathbf{x}_p$, where $\mathbf{x}_h$ is the general solution to $A\mathbf{x} = \mathbf{0}$ and $\mathbf{x}_p$ is the particular row space solution to $A\mathbf{x} = \mathbf{b}$.

Example 4. Since we can solve $A\mathbf{x} = \mathbf{b}$ only when $\mathbf{b}$ lies in $\mathcal{R}(A)$, another idea can be suggested by using Figure 4. It is the *Fredholm alternative* [**1, 2**]: For $\mathbf{b}$ in $\mathbf{R}^m$, either $\mathbf{b} \cdot \mathbf{y} = \mathbf{0}$ for every $\mathbf{y}$ in $\mathcal{N}(A^T)$ or $A\mathbf{x} = \mathbf{b}$ has no solution, exclusively. In other words, a *solvability condition* for $A\mathbf{x} = \mathbf{b}$ is that $\mathbf{b}$ be orthogonal to $\mathcal{N}(A^T)$. Further, we see that $A\mathbf{x} = \mathbf{b}$ has a solution for *every* $\mathbf{b}$ in $\mathbf{R}^m$ if and only if $\mathcal{N}(A^T)$ contains only the zero vector, that is, if and only if rank $A = m$.

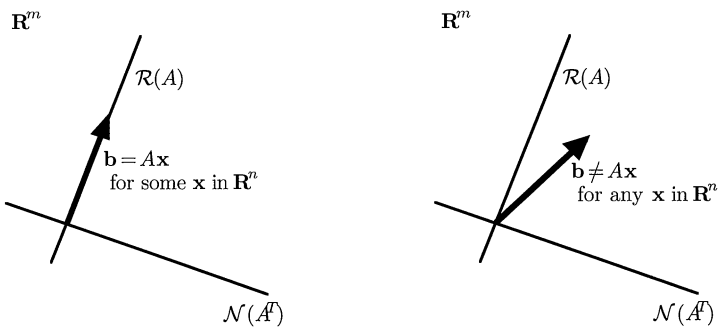


Figure 4. For $\mathbf{b} \in \mathbf{R}^m$, either $\mathbf{b} \perp \mathcal{N}(A^T)$ *(left)* or $A\mathbf{x} = \mathbf{b}$ has no solution *(right)*, exclusively.

Example 5. Figure 5 shows that if $\mathbf{b}$ is not in $\mathcal{R}(A)$, then we may solve, instead, $A\mathbf{x} = \mathbf{b}_{\text{proj}}$, where $\mathbf{b}_{\text{proj}}$ is the projection of $\mathbf{b}$ onto $\mathcal{R}(A)$. It is evident from the diagram that $\mathbf{b}_{\text{proj}}$ is the vector in $\mathcal{R}(A)$ that is closest to $\mathbf{b}$. This, of course, is the least-squares approximation, and the error is clearly seen to be $\mathbf{e}$, the orthogonal complement of $\mathbf{b}_{\text{proj}}$.

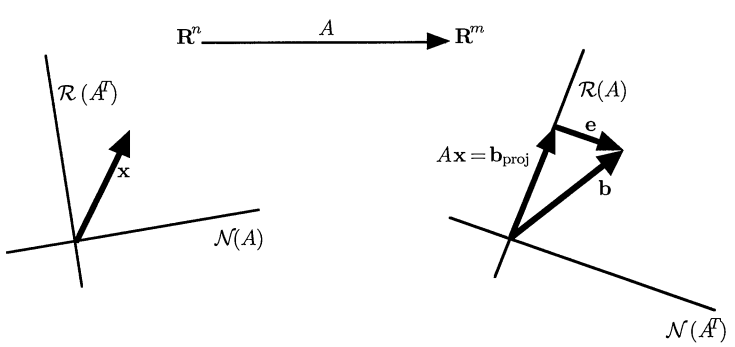


Figure 5. Least-squares approximation: $A\mathbf{x} = \mathbf{b}_{\text{proj}}$, the projection of $\mathbf{b}$ onto $\mathcal{R}(A)$; the error is $\mathbf{e}$, the orthogonal complement of $\mathbf{b}_{\text{proj}}$.

Example 6. Lastly, suppose that $\mathbf{b}$ in $\mathbf{R}^m$ is not in $\mathcal{R}(A)$ and that $\mathbf{x}^+$ is the row space component of a least-squares solution to $A\mathbf{x} = \mathbf{b}$ (see example 5). If A^+ is an $n \times m$ matrix such that $A^+\mathbf{b}_{\text{proj}} = \mathbf{x}^+$ and $A^+\mathbf{e} = \mathbf{0}$, then $A^+\mathbf{b} = A^+(\mathbf{b}_{\text{proj}} + \mathbf{e}) = \mathbf{x}^+$. In fact, as Figure 6 shows, we would have

$$A^+A\mathbf{x}^+ = A^+\mathbf{b}_{\text{proj}} = \mathbf{x}^+.$$

That is, A^+ would be a left-inverse of A on the latter's row space.

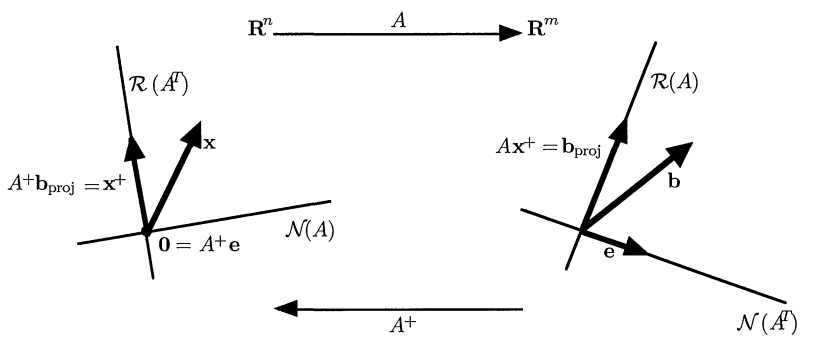


Figure 6. The matrix A^+, the pseudoinverse of A, is a left-inverse of A on the latter's row space: $A^+A\mathbf{x}^+ = \mathbf{x}^+$ for $\mathbf{x}^+ \in \mathcal{R}(A^T)$.

Actually, the matrix A^+ does exist: it is called the *pseudoinverse* of A [**4, 5**], and it is precisely A^{-1} when A is invertible. The pseudoinverse of A can be found using the singular value decomposition (SVD) of A; however, even without discussing the SVD of a matrix, one can still introduce the notion of a pseudoinverse by using Figure 6.

Acknowledgment. I thank an anonymous referee for some very useful suggestions that helped me convey my ideas more clearly.

References

1. A. W. Naylor and G. R. Sell, *Linear Operator Theory in Engineering and Science* (Appl. Math. Sci. series, vol. 40), Springer-Verlag, New York, 1982, 462–463.
2. J. Rauch, *Partial Differential Equations* (Grad. Texts in Mathematics, vol. 128), Springer-Verlag, New York, 1991, 200–207.
3. G. Strang, Graphs, matrices, and subspaces, *College Mathematics Journal* 24:1 (1993) 20–28.
4. ———, The fundamental theorem of linear algebra, *American Mathematical Monthly*, 100:9 (1993) 848–855.
5. ———, *Introduction to Linear Algebra*, Wellesley-Cambridge Press, Wellesley, MA, 1993, 150–188.