THE LANGUAGE OF LINEAR ALGEBRA

GILBERT STRANG

To express the ideas of linear algebra we need words. To use those words — to speak this subject — is to connect those ideas. Our goal in teaching is mastery of the language, which requires mastery of the words and ideas.

Certainly, the words don't all come at once. We build from one level to the next, as safely as possible. At level zero are the real numbers. At level one, m numbers are the components of a vector. At level two, we take linear combinations of n vectors; these vectors $a_1, \ldots, a_n$ go into the columns of a matrix A. At level three we take *all* linear combinations of those vectors; this produces the column space $\mathbf{C}(A)$. At level four, the step that started with coefficients $x_1, \ldots, x_n$ and ended with combinations $x_1a_1 + \cdots + x_na_n$ becomes a transformation from $\mathbf{R}^n$ to $\mathbf{R}^m$.

All together, the levels are

(0) **number** (1) **vector** (2) **matrix** (3) **subspace** (4) **transformation**

These are the nouns of linear algebra. Next come the adjectives.

(0) Numbers can be real or complex
(1) Vectors $a_1, \ldots, a_n$ can be independent, orthogonal, orthonormal
(2) Matrices can be invertible, triangular, symmetric, diagonalizable
(3) Subspaces can be orthogonal, intersecting, complementary
(4) Transformations can be linear, injective, surjective, bijective.

Every linguist will look for the verbs. But allow me to jump ahead to the "Big Picture" of linear algebra and two of the ideas behind it.

The picture shows two column spaces and two nullspaces: $\mathbf{C}(A)$ and $\mathbf{N}(A)$, $\mathbf{C}(A^T)$ and $\mathbf{N}(A^T)$. The first and last are orthogonal complements in $\mathbf{R}^m$, the other two are orthogonal complements in $\mathbf{R}^n$. Every m by n matrix fits this picture, and its rank r determines the dimensions of all four subspaces [1].

For the best matrices, square and invertible, the picture is too simple: The nullspaces contain only the zero vector, and $\mathbf{C}(A) = \mathbf{C}(A^T) = \mathbf{R}^n$. The matrix has full rank. Just short of that perfect case $m = n = r$ are two possibilities that are all-important for *rectangular matrices*. I would like to focus this paper on those two possibilities, in the belief that they are just about perfect in using the language of linear algebra (at all levels).

The first possibility will eventually be described by eleven equivalent sentences. I will ask my class for the eleven parallel sentences that describe the second possibility. The starting point of this short paper was in preparation for that class (which begins next week). What am I hoping to teach and what are the steps?

The Two Possibilities

For a square matrix, the most important question is *invertibility.* Computationally, this is not simple to decide. Algebraically, it requires independent columns and also independent rows. Magically, either one of these conditions is enough; a right inverse will also be a left inverse. But to understand these conditions we must separate them – and then the matrix can be rectangular.

For an m by n matrix, the first possibility is *independent columns.* That is a description at level 1 (and not trivial to explain). As the course moves forward, new words enter the language. We see the same idea at higher levels. Here are all four levels at once:

(Vectors)	The columns of A are independent.
(Matrices)	The rank of A is n.
(Subspaces)	$\mathbf{N}(A)$ contains only the zero vector.
(Transformations)	The transformation $T(x) = Ax$ is **1 to 1:**
	If $Ax = Ay$ then $x = y$.

The idea of a linear transformation does seem to come last (level 4). I teach it last. But a rewording into "$Ax = b$ has at most one solution for every b" is a statement about the linear equations that start the course.

The second possibility is *independent rows.* Where does that lead? Again this has a new meaning at each level. And again level 4 tells us a crucial fact about the linear system $Ax = b$: There is at least one solution for every b.

(Vectors)	The rows of A are independent.
(Matrices)	The rank of A is m.
(Subspaces)	The column space $\mathbf{C}(A)$ is all of $\mathbf{R}^m$.
(Transformations)	$T(x) = Ax$ maps $\mathbf{R}^n$ **onto** $\mathbf{R}^m$.

Inverse Matrices and Eleven Equivalent Sentences

I suppose the task of teaching is never ending. Four levels seem like enough but they are not. We used the word *independent* but not the word *span.* Both are needed for the central construction of linear algebra, a *basis* for a subspace. And we entirely missed the idea of an inverse matrix.

Let me improve one of the lists. If we start with independent columns, we also learn properties of the rows. The relation of rows in $\mathbf{R}^n$ to columns in $\mathbf{R}^m$ is at the heart of linear algebra, leading to the first great theorem in our subject: **Row rank equals column rank**. That goes beyond our two possibilities — it applies to all matrices. (Four proofs of this theorem are collected in [2].) This extended list mentions the reduced row echelon form R and the square symmetric matrix A^TA :

(1a) The columns of A are independent
(1b) The rows of A span $\mathbf{R}^n$
(2a) The rank of A is n : "full column rank"
(2b) All the columns of A are pivot columns so $R = \begin{bmatrix} I \\ 0 \end{bmatrix}$
(3a) The nullspace $\mathbf{N}(A)$ contains only the zero vector
(3b) The row space $\mathbf{C}(A^T)$ is all of $\mathbf{R}^n$
(3c) The columns of A are a basis for its column space
(4a) $T(x) = Ax$ is a 1 to 1 linear transformation: If $Ax = Ay$ then $x = y$ (uniqueness of solutions to $Ax = b$)
(4b) The kernel of T contains only $x = 0$
(5a) The matrix A^TA is invertible (and symmetric positive definite)
(5b) A has a left-inverse $B = (A^TA)^{-1}A^T$, with $BA = I$

I will commit a full class hour to presenting this summary of the theory and the language of linear algebra. Every student will be asked for eleven parallel statements, starting with independent rows. The hour will end with this last diagram, starting from all matrices and ending with square invertible matrices. In between come the two possibilities that are the focus of this article and this lecture.

Independent columns
rank $r = n$

All matrices
$r \le m, r \le n$

Square and invertible
rank $r = m = n$

Independent rows
rank $r = m$

A student would learn a lot by collecting equivalent conditions for that most special and attractive and frequently met case — when A is square and invertible. This brings with it two more words and ideas that are a big part of the basic course:

(1) The **determinant** of A is not zero.
(2) All the **eigenvalues** of A are nonzero.

Linear algebra is a wonderful subject to teach. Real mathematics comes real early.

References

[1] Gilbert Strang, The fundamental theorem of linear algebra, *American Mathematical Monthly* 100 (1993) 848-855.

[2] Gilbert Strang, Row rank equals column rank: Four approaches, IMAGE 53 (2014) 17. (`ilasic.org/IMAGE/IMAGES/image53.pdf`)

[3] Gilbert Strang, *Introduction to Linear Algebra,* 4^{th} edition, Wellesley-Cambridge Press, 2009.

[4] Paul R. Halmos, *Finite-Dimensional Vector Spaces,* Princeton University Press, 1942. Springer, 1958.