The Four Fundamental Subspaces

CS6015 / LARP

ACK : Linear Algebra and Its Applications - Gilbert Strang

Subspaces can be described in two ways. First, we may be given a set of vectors that span the space. (*Example*: The columns span the column space.) Second, we may be told which conditions the vectors in the space must satisfy. (*Example*: The null space consists of all vectors that satisfy $Ax = 0$.)

The first description may include useless vectors (dependent columns). The second description may include repeated conditions (dependent rows). We can't write a basis by inspection, and a systematic procedure is necessary.

When elimination on A produces an echelon matrix U or a reduced R, we will find a basis for each of the subspaces associated with A. Then we have to look at the extreme case of full rank:

> *When the rank is as large as possible, $r = n$ or $r = m$ or $r = m = n$, the matrix has a* **left-inverse** *B or a* **right-inverse** *C or a* **two-sided** *A^{-1}.*

To organize the whole discussion, we take each of the four subspaces in turn. Two of them are familiar and two are new.

1. The **column space** of A is denoted by $C(A)$. Its dimension is the rank r.
2. The **nullspace** of A is denoted by $N(A)$. Its dimension is $n - r$.
3. The **row space** of A is the column space of A^T. It is $C(A^T)$, and it is spanned by the rows of A. Its dimension is also r.
4. The **left nullspace** of A is the nullspace of A^T. It contains all vectors y such that $A^T y = 0$, and it is written $N(A^T)$.
5. The point about the last two subspaces is that they come from A^T. If A is an $m\ by\ n$ matrix, you can see which "host" spaces contain the four subspaces by looking at the number of components:

> The nullspace $\boldsymbol{N}(A)$ and row space $\boldsymbol{C}(A^{\mathrm{T}})$ are subspaces of $\mathbf{R}^n$.
> The left nullspace $\boldsymbol{N}(A^{\mathrm{T}})$ and column space $\boldsymbol{C}(A)$ are subspaces of $\mathbf{R}^m$.

The rows have n components and the columns have m.

For a simple matrix like

$$A = U = R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

the column space is the line through $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. The row space is the line through $[1\ 0\ 0]^T$. It is in $\boldsymbol{R}^3$. The nullspace is a plane in $\boldsymbol{R}^3$ and the left nullspace is a line in $\boldsymbol{R}^2$:

$$\boldsymbol{N}(A) \quad \text{contains} \quad \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \qquad \boldsymbol{N}(A^{\mathrm{T}}) \quad \text{contains} \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Note that all vectors are column vectors. Even the rows are transposed, and the row space of A is the column space of A^T. Our problem will be to connect the four spaces for U (after elimination) to the four spaces for A:

Basic example

$$U = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{came from} \quad A = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 7 \\ -1 & -3 & 3 & 4 \end{bmatrix}.$$

Row Space

The **row space** of A For an echelon matrix like U, the row space is clear. It contains all combinations of the rows, as every row space does—but here the third row contributes nothing. The first two rows are a basis for the row space. A similar rule applies to every echelon matrix U or R, with r pivots and r nonzero rows: **The nonzero rows are a basis, and the row space has dimension** r. That makes it easy to deal with the original matrix A.

2M The row space of A has the same dimension r as the row space of U, and it has the same bases, because **the row spaces of** A **and** U **(and** R**) are the same.**

The reason is that each elementary operation leaves the row space unchanged. The rows in U are combinations of the original rows in A. Therefore the row space of U contains nothing new. At the same time, because every step can be reversed, nothing is lost; the rows of A can be recovered from U. It is true that A and U have different rows, but the combinations of the rows are identical: *same space*!

Note that we did not start with the m rows of A, which span the row space, and discard $m - r$ of them to end up with a basis. According to 2L, we could have done so. But it might be hard to decide which rows to keep and which to discard, so it was easier just to take the nonzero rows of U.

Nullspace

The **nullspace** of A Elimination simplifies a system of linear equations without changing the solutions. The system $Ax = 0$ is reduced to $Ux = 0$, and this process is reversible. **The nullspace of A is the same as the nullspace of U and R.** Only r of the equations $Ax = 0$ are independent. Choosing the $n - r$ "*special solutions*" to $Ax = 0$ provides a definite basis for the nullspace:

> **2N** The nullspace $N(A)$ has dimension $n - r$. The "*special solutions*" are a basis—each free variable is given the value 1, while the other free variables are 0. Then $Ax = 0$ or $Ux = 0$ or $Rx = 0$ gives the pivot variables by back-substitution.

This is exactly the way we have been solving $Ux = 0$. The basic example above has pivots in columns 1 and 3. Therefore its free variables are the second and fourth v and y.

The basis for the nullspace is

$$\textbf{Special solutions} \quad \begin{aligned} v &= 1 \\ y &= 0 \end{aligned} \quad x_1 = \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix}; \qquad \begin{aligned} v &= 0 \\ y &= 1 \end{aligned} \quad x_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}.$$

Any combination $c_1x_1 + c_2x_2$ has c_1 as its v component, and c_2 as its y component. The only way to have $c_1x_1 + c_2x_2 = 0$ is to have $c_1 = c_2 = 0$, so these vectors are independent. They also span the nullspace; the complete solution is $vx_1 + yx_2$. Thus the $n - r = 4 - 2$ vectors are a basis.

The nullspace is also called the kernel of A, and its dimension $n - r$ is the nullity.

Column space

The **column space** of A The column space is sometimes called the range. This is consistent with the usual idea of the range, as the set of all possible values $f(x)$; x is in the domain and $f(x)$ is in the range. In our case the function is $f(x) = Ax$. Its domain consists of all x in $\boldsymbol{R}^n$; its range is all possible vectors Ax, which is the column space. (In an earlier edition of this book we called it $R(A)$.)

Our problem is to find bases for the column spaces of U and A. **Those spaces are different** (just look at the matrices!) but their dimensions are the same.

The first and third columns of U are a basis for its column space. They are the **columns with pivots**. Every other column is a combination of those two. Furthermore, the same is true of the original A—even though its columns are different. **The pivot columns of A are a basis for its column space**. The second column is three times the first, just as in U. The fourth column equals (column 3) − (column 1). The same nullspace is telling us those dependencies.

The reason is this: $Ax = 0$ *exactly when* $Ux = 0$. The two systems are equivalent and have the same solutions. The fourth column of U was also (column 3) − (column 1). Every linear dependence $Ax = 0$ among the columns of A is matched by a dependence $Ux = 0$ among the columns of U, with exactly the same coefficients. *If a set of columns of A is independent, then so are the corresponding columns of U, and vice versa.*

To find a basis for the column space $C(A)$, we use what is already done for U. The r columns containing pivots are a basis for the column space of U. We will pick those same r columns in A:

2O The dimension of the column space $C(A)$ equals the rank r, which also equals the dimension of the row space: **The number of independent columns equals the number of independent rows.** A basis for $C(A)$ is formed by the r columns of A that correspond, in U, to the columns containing pivots.

The row space and the column space have the same dimension r! This is one of the most important theorems in linear algebra. It is often abbreviated as "**row rank = column rank**." It expresses a result that, for a random $10\ by\ 12$ matrix, is not at all obvious. It also says something about square matrices: *If the rows of a square matrix are linearly independent, then so are the columns (and vice versa).* Again, that does not seem self-evident (at least, not to the author).

To see once more that both the row and column spaces of U have dimension r, consider a typical situation with rank $r = 3$. The echelon matrix U certainly has three independent rows:

$$U = \begin{bmatrix} d_1 & * & * & * & * & * \\ 0 & 0 & 0 & d_2 & * & * \\ 0 & 0 & 0 & 0 & 0 & d_3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

We claim that U also has three independent columns, and no more, The columns have only three nonzero components. If we can show that the pivot columns—the first, fourth, and sixth—are linearly independent, they must be a basis (for the column space of U, not A!). Suppose a combination of these pivot columns produced zero:

$$c_1 \begin{bmatrix} d_1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} * \\ d_2 \\ 0 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} * \\ * \\ d_3 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Working upward in the usual way, c_3 must be zero because the pivot $d_3 \neq 0$, then c_2 must be zero because $d_2 \neq 0$, and finally $c_1 = 0$. This establishes independence and completes the proof. Since $Ax = 0$ if and only if $Ux = 0$, the first, fourth, and sixth columns of A— whatever the original matrix A was, which we do not even know in this example—are a basis for $C(A)$.

The row space and column space both became clear after elimination on A. Now comes the fourth fundamental subspace, which has been keeping quietly out of sight. Since the first three spaces were $C(A)$, $N(A)$, and $C(A^T)$, the fourth space must be $N(A^T)$, It is the nullspace of the transpose, or the left nullspace of A. $A^T y = 0$ means $y^T A = 0$, and the vector appears on the left-hand side of A.

Left nullspace

The left nullspace of A (= **the nullspace of** A^T) If A is an m by n matrix, then A^T is $n\ by\ m$. Its nullspace is a subspace of $\boldsymbol{R}^m$; the vector y has m components. Written as $y^TA = 0$, those components multiply the rows of A to produce the zero row:

$$y^TA = [y_1 \cdots y_m][A] = [0 \cdots 0].$$

The dimension of this nullspace $N(A^T)$ is easy to find, For *any* matrix, **the number of pivot variables plus the number of free variables must match the total number of columns**. For A, that was $r + (n - r) = n$. In other words, rank plus nullity equals n:

$$dimension\ of\ C(A) + dimension\ of\ N(A) = number\ of\ columns.$$

This law applies equally to A^T, which has m columns. A^T is just as good a matrix as A. But the dimension of its column space is also r, so

$$r + dimension\left(N(A^T)\right) = m$$

2P The left nullspace $N(A^T)$ has dimension $m - r$.

The $m - r$ solutions to $y^T A = 0$ are hiding somewhere in elimination. The rows of A combine to produce the $m - r$ zero rows of U. Start from $PA = LU$, or $L^{-1}PA = U$. The last $m - r$ rows of the invertible matrix $L^{-1}P$ must be a basis of y's in the left nullspace— because they multiply A to give the zero rows in U.

In our $3\ by\ 4$ example, the zero row was row $3 - 2(row\ 2) + 5(row\ 1)$. Therefore the components of y are 5, −2, 1. This is the same combination as in $b_3 - 2b_2 + 5b_1$ on the right-hand side, leading to 0 = 0 as the final equation. That vector y is a basis for the left nullspace, which has dimension $m - r\ =\ 3\ -\ 2\ =\ 1$. It is the last row of $L^{-1}P$, and produces the zero row in U—and we can often see it without computing L^{-1}. When desperate, it is always possible just to solve $A^T y = 0$.

Example 1. $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$ has $m = n = 2$, and rank $r = 1$.

1. The **column space** contains all multiples of $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$. The second column is in the same direction and contributes nothing new.
2. The **nullspace** contains all multiples of $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$. The vector satisfies $Ax = 0$.
3. The **row space** contains all multiples $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$. It is written as a column vector, since strictly speaking it is in the column space of A^T.
4. The left nullspace contains all multiples of y=$\begin{bmatrix} -3 \\ 1 \end{bmatrix}$. The rows of A with coefficients -3 and 1 add to zero, so $A^T y = 0$.

In the previous example, *all four subspaces are lines*. That is an accident, coming from $r = 1$ and $n - r = 1$ and $m - r = 1$. Figure 2.5 shows that two pairs of lines are perpendicular. That is no accident!

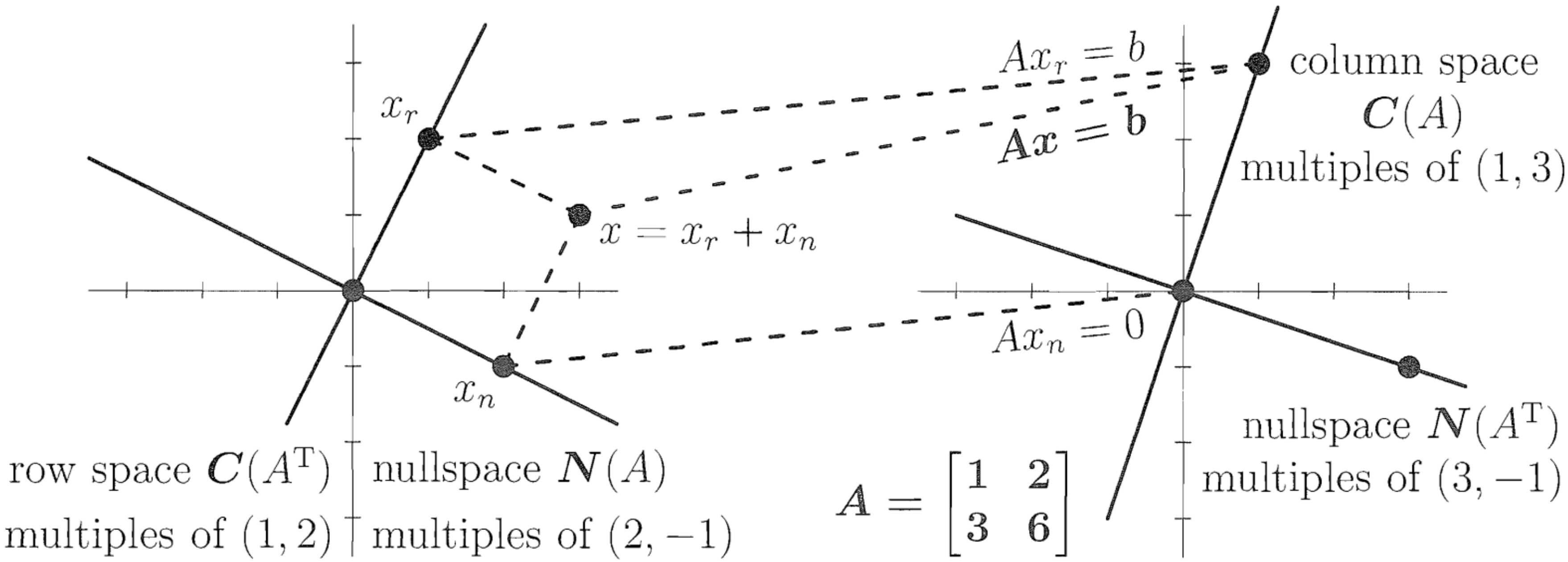

Figure 2.5: The four fundamental subspaces (lines) for the singular matrix A.

If you change the last entry of A from 6 to 7, all the dimensions are different. The column space and row space have dimension $r = 2$. The nullspace and left nullspace contain only the vectors $x = 0$ and $y = 0$. *The matrix is invertible.*

Existence of Inverses

We know that if A has a left-inverse $(BA = I)$ and a right-inverse $(AC = I)$, then the two inverses are equal: $B = B(AC)(BA)C = C$. Now, from the rank of a matrix, it is easy to decide which matrices actually have these inverses. Roughly speaking, **an inverse exists only when the rank is as large as possible.**

The rank always satisfies $r \leq m$ and also $r \leq n$. An $m\ by\ n$ matrix cannot have more than m independent rows or n independent columns. There is not space for more than m pivots, or more than n. We want to prove that when $r = m$ there is a right-inverse, and $Ax = b$ always has a solution. When $r = n$ there is a left-inverse, and the solution (if it exists) is unique.

Only a square matrix can have both $r = m$ and $r = n$, and therefore only a square matrix can achieve both existence and uniqueness. Only a square matrix has a two-sided inverse.

2Q EXISTENCE: Full row rank $r = m$. $Ax = b$ has at least one solution x for every b if and only if the columns span $\boldsymbol{R}^m$. Then A has a right-inverse C such that $AC = Im$ $(m\ by\ m)$. This is possible only if $m \leq n$.

UNIQUENESS: Full column rank $r = n$. $Ax = b$ has **at most** one solution x for every b if and only if the columns are linearly independent. Then A has an $n\ by\ m$ left-inverse B such that $BA = In$. This is possible only if $m \geq n$.

In the existence case, one possible solution is $x = Cb$, since then $Ax = ACb = b$. But there will be other solutions if there are other right-inverses. The number of solutions when the columns span $\boldsymbol{R}^m$ is 1 or ∞.

In the uniqueness case, if there is a solution to $Ax = b$, it has to be $x = BAx = Bb$. But there may be no solution. The number of solutions is 0 or 1.

There are simple formulas for the best left and right inverses, if they exist:

One-sided inverses $B = (A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$ and $C = A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}$.

Certainly $BA = I$ and $AC = I$. What is not so certain is that $A^T A$ and AA^T are actually invertible. We show in next chapter that $A^T A$ does have an inverse if the rank is n, and AA^T has an inverse when the rank is m. Thus the formulas make sense exactly when the rank is as large as possible, and the one-sided inverses are found.

Example 2. Consider a simple 2 by 3 matrix of rank 2:

$$A = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \end{bmatrix}.$$

Since $r = m = 2$, the theorem guarantees a right-inverse C:

$$AC = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{5} \\ c_{31} & c_{32} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

There are many right-inverses because the last row of C is completely arbitrary. This is a case of existence but not uniqueness. The matrix A has no left-inverse because the last column of BA is certain to be zero. The specific right-inverse $C = A^T(AA^T)^{-1}$ chooses c_{31} and c_{32} to be zero:

Best right-inverse $$A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1} = \begin{bmatrix} 4 & 0 \\ 0 & 5 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{16} & 0 \\ 0 & \frac{1}{25} \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{5} \\ 0 & 0 \end{bmatrix} = C.$$

This is the *pseudoinverse*—a way of choosing the best C. The transpose of A yields an example with infinitely many left-inverses:

$$BA^{\mathrm{T}} = \begin{bmatrix} \frac{1}{4} & 0 & b_{13} \\ 0 & \frac{1}{5} & b_{23} \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 5 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Now it is the last column of B that is completely arbitrary. The best left-inverse (also the pseudoinverse) has $b_{13} = b_{23} = 0$. This is a "*uniqueness case*," when the rank is $r = n$. There are no free variables, since $n - r = 0$. If there is a solution it will be the only one. You can see when this example has one solution or no solution:

$$\begin{bmatrix} 4 & 0 \\ 0 & 5 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad \text{is solvable exactly when} \quad b_3 = 0.$$

A rectangular matrix cannot have both existence and uniqueness. If m is different from n, we cannot have $r = m$ and $r = n$.

A square matrix is the opposite. If $m = n$, we cannot have one property without the other. A square matrix has a left-inverse if and only if it has a right-inverse. There is only one inverse, namely $B = C = A^{-1}$. *Existence implies uniqueness and uniqueness implies existence, when the matrix is square.* The condition for invertibility is **full rank**: $r = m = n$. Each of these conditions is a necessary and sufficient test:

1. The columns span $\boldsymbol{R}^n$, so $Ax = b$ has at least one solution for every b.
2. The columns are independent, so $Ax = 0$ has only the solution $x = 0$.

This list can be made much longer, especially if we look ahead to later chapters. Every condition is equivalent to every other, and ensures that A is invertible.

3. The rows of A span $\boldsymbol{R}^n$.
4. The rows are linearly independent.
5. Elimination can be completed: $PA = LDU$, with all n pivots.
6. The determinant of A is not zero.

7. Zero is not an eigenvalue of A.

8. $A^T A$ is positive definite.

Here is a typical application to polynomials $P(t)$ of degree $n-1$. The only such polynomial that vanishes at $t_1, \dots, t_n$ is $P(t) \equiv 0$. No other polynomial of degree $n-1$ can have n roots. This is uniqueness, and it implies existence: Given any values $b_1, \dots, b_n$, there exists a polynomial of degree $n-1$ interpolating these values: $P(t_i) = b_i$. The point is that we are dealing with a square matrix; the number n of coefficients in $P(t) = x_1 + x_2 t + \cdots + x_n t^{n-1}$ matches the number of equations:

Interpolation
$P(t_i) = b_i$

$$\begin{bmatrix} 1 & t_1 & t_1^2 & \cdots & t_1^{n-1} \\ 1 & t_2 & t_2^2 & \cdots & t_2^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & t_n & t_n^2 & \cdots & t_n^{n-1} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

That *Vandermonde* matrix is n *by* n and full rank. $Ax = b$ always has a solution—a polynomial can be passed through any b_i at distinct points t_i. Later we shall actually find the determinant of A; it is not zero.

Matrices of Rank 1

Finally comes the easiest case, when the rank is as small as possible (except for the zero matrix with rank 0), One basic theme of mathematics is, given something complicated, to show how it can be broken into simple pieces. For linear algebra, the simple pieces are matrices of **rank 1**:

$$\textbf{Rank 1} \qquad A = \begin{bmatrix} 2 & 1 & 1 \\ 4 & 2 & 2 \\ 8 & 4 & 4 \\ -2 & -1 & -1 \end{bmatrix} \quad \text{has} \quad r = 1.$$

Every row is a multiple of the first row, so the row space is one-dimensional. In fact, we can write the whole matrix *as the product of a column vector and a row vector*:

$$A = (\textbf{column})(\textbf{row}) \qquad \begin{bmatrix} 2 & 1 & 1 \\ 4 & 2 & 2 \\ 8 & 4 & 4 \\ -2 & -1 & -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \\ -1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \end{bmatrix}.$$

The product of a $4\ by\ 1$ matrix and a $1\ by\ 3$ matrix is a $4\ by\ 3$ matrix. *This product has rank* 1. At the same time, the columns are all multiples of the same column vector; the column space shares the dimension $r = 1$ and reduces to a line.

Every matrix of rank* 1 *has the simple form* $A = uv^{\mathrm{T}}$ = *column times row.

The rows are all multiples of the same vector v^T , and the columns are all multiples of u. The row space and column space are lines—the easiest case.