Fall 2019 Math 201-003 at Drexel: blackboard notes of 2019-10-07

Darij Grinberg

December 5, 2019

1. Gaussian elimination (continued)

1.1. Linear dependence (continued)

Proposition 1.1.1. Let $v_1, v_2, \ldots, v_k$ be k vectors in $\mathbb{R}^n$. Then, $v_1, v_2, \ldots, v_k$ are dependent **if and only if** one of them is a combination of the others.

(Remember: "dependent" = "linearly dependent", and "combination" = "linear combination".)

Proof. $\Longrightarrow$: Assume that $v_1, v_2, \ldots, v_k$ are dependent. We must show that one of them is a combination of the others.

We have assumed that $v_1, v_2, \ldots, v_k$ are dependent. Hence, there exists a nontrivial relation

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k = 0.$$

Since this relation is nontrivial, there is some i such that $\lambda_i \neq 0$. Now, we can rewrite this relation as follows:

$$\lambda_i v_i = -\lambda_1 v_1 - \lambda_2 v_2 - \cdots - \lambda_{i-1} v_{i-1} - \lambda_{i+1} v_{i+1} - \cdots - \lambda_k v_k.$$

But since $\lambda_i \neq 0$, we can multiply this equality by $\frac{1}{\lambda_i}$. We get

$$\begin{aligned} v_i &= \frac{1}{\lambda_i}\left(-\lambda_1 v_1 - \lambda_2 v_2 - \cdots - \lambda_{i-1} v_{i-1} - \lambda_{i+1} v_{i+1} - \cdots - \lambda_k v_k\right) \\ &= \frac{-\lambda_1}{\lambda_i} v_1 + \frac{-\lambda_2}{\lambda_i} v_2 + \cdots + \frac{-\lambda_{i-1}}{\lambda_i} v_{i-1} + \frac{-\lambda_{i+1}}{\lambda_i} v_{i+1} + \cdots + \frac{-\lambda_k}{\lambda_i} v_k. \end{aligned}$$

Thus, v_i is a linear combination of all the remaining v's (that is, of $v_1, v_2, \ldots, v_{i-1}, v_{i+1}, \ldots, v_k$).

$\Longleftarrow$: Assume that one of $v_1, v_2, \ldots, v_k$ is a combination of the others. We must show that $v_1, v_2, \ldots, v_k$ are dependent.

We have assumed that one of $v_1, v_2, \ldots, v_k$ – say, v_i – is a combination of the others. Thus,

$$v_i = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_{i-1} v_{i-1} + \alpha_{i+1} v_{i+1} + \cdots + \alpha_k v_k$$

for some numbers $\alpha_1, \alpha_2, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_k$. Subtracting v_i from both sides, we get

$$\begin{aligned} 0 &= (\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_{i-1} v_{i-1} + \alpha_{i+1} v_{i+1} + \cdots + \alpha_k v_k) - v_i \\ &= \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_{i-1} v_{i-1} + (-1)\, v_i + \alpha_{i+1} v_{i+1} + \cdots + \alpha_k v_k. \end{aligned}$$

This is a linear relation between $v_1, v_2, \ldots, v_k$, and it is nontrivial since $-1 \neq 0$. So $v_1, v_2, \ldots, v_k$ are dependent. □

Warning: "One of $v_1, v_2, \ldots, v_k$ is a combination of others" is not the same as saying "any of $v_1, v_2, \ldots, v_k$ is a combination of others".

Example 1.1.2. Let $k = 3$ and $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $v_2 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$ and $v_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Then, v_1, v_2, v_3 are dependent. Thus, the proposition shows that one of them is a combination of the others. And indeed,

$$\begin{aligned} &v_1 \text{ is a combination of } v_2 \text{ and } v_3\text{, since } v_1 = \frac{1}{2} v_2 + 0 v_3; \\ &v_2 \text{ is a combination of } v_1 \text{ and } v_3\text{, since } v_2 = 2 v_1 + 0 v_3. \end{aligned}$$

But v_3 is not a combination of v_1 and v_2.

Corollary 1.1.3. Let v and w be two vectors in $\mathbb{R}^n$.

(a) The one-element list (v) is linearly dependent if and only if $v = 0$.

(b) The two-element list (v, w) is linearly dependent if and only if one of v and w is a scalar multiple of the other (i.e., there is $\lambda \in \mathbb{R}$ such that either $v = \lambda w$ or $w = \lambda v$).

Proof. **(a)** Left to the reader.

(b) According to the proposition, the two-element list (v, w) is linearly dependent if and only if one of v and w is a combination of the other. But a linear combination of a single vector u is the same as a scalar multiple of u. □

Slogan:

- Linear dependence is a sort of redundancy – if a list of vectors is linearly dependent, then all its combinations can be obtained even if one of the vectors is removed (but we have to remove the right one, not just any one).
- Linear dependence is a generalization of proportionality between two lists of numbers.

Note that in the plane $\mathbb{R}^2$, linearly dependent vectors are said to be **parallel**.

Proposition 1.1.4. If a list of vectors contains the 0 vector, then it is automatically dependent.

Proof. If $v_i = 0$, then $0 \cdot v_1 + 0 \cdot v_2 + \cdots + 0 \cdot v_{i-1} + 1 \cdot v_i + 0 \cdot v_{i+1} + \cdots + 0 \cdot v_k = 0$ is a nontrivial relation. □

1.1.1. Linear combinations as matrix-by-vector products

Definition 1.1.5. Let $v_1, v_2, \ldots, v_m$ be m column vectors in $\mathbb{R}^n$. Then,

$$[v_1 \mid v_2 \mid \cdots \mid v_m]$$

shall denote the $n \times m$-matrix whose columns are $v_1, v_2, \ldots, v_m$.

This notation just lets us paste together some column vectors into a matrix. For example,

$$\left[\begin{pmatrix} a \\ b \end{pmatrix} \mid \begin{pmatrix} c \\ d \end{pmatrix} \mid \begin{pmatrix} e \\ f \end{pmatrix}\right] = \begin{pmatrix} a & c & e \\ b & d & f \end{pmatrix}.$$

In what follows, we will need a lemma:

Lemma 1.1.6. Let $v_1, v_2, \ldots, v_m$ be m column vectors in $\mathbb{R}^n$. Let $\lambda_1, \lambda_2, \ldots, \lambda_m$ be m real numbers. Then,

$$[v_1 \mid v_2 \mid \cdots \mid v_m] \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{pmatrix} = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_m v_m.$$

This lemma shows that the linear combination $\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_m v_m$ can be written as a "matrix · column vector" product.

Proof of Lemma 1.1.6. Just multiply things out! Write each vector v_i as $\begin{pmatrix} v_{i,1} \\ v_{i,2} \\ \vdots \\ v_{i,n} \end{pmatrix}$.

Thus,

$$[v_1 \mid v_2 \mid \cdots \mid v_m] = \begin{pmatrix} v_{1,1} & v_{2,1} & \cdots & v_{m,1} \\ v_{1,2} & v_{2,2} & \cdots & v_{m,2} \\ \vdots & \vdots & \ddots & \vdots \\ v_{1,n} & v_{2,n} & \cdots & v_{m,n} \end{pmatrix}.$$

Hence,

$$
\begin{aligned}
&[v_1 \mid v_2 \mid \cdots \mid v_m] \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{pmatrix} \\
&= \begin{pmatrix} v_{1,1} & v_{2,1} & \cdots & v_{m,1} \\ v_{1,2} & v_{2,2} & \cdots & v_{m,2} \\ \vdots & \vdots & \ddots & \vdots \\ v_{1,n} & v_{2,n} & \cdots & v_{m,n} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{pmatrix} \\
&= \begin{pmatrix} v_{1,1}\lambda_1 + v_{2,1}\lambda_2 + \cdots + v_{m,1}\lambda_m \\ v_{1,2}\lambda_1 + v_{2,2}\lambda_2 + \cdots + v_{m,2}\lambda_m \\ \vdots \\ v_{1,n}\lambda_1 + v_{2,n}\lambda_2 + \cdots + v_{m,n}\lambda_m \end{pmatrix} \\
&= \begin{pmatrix} v_{1,1}\lambda_1 \\ v_{1,2}\lambda_1 \\ \vdots \\ v_{1,n}\lambda_1 \end{pmatrix} + \begin{pmatrix} v_{2,1}\lambda_2 \\ v_{2,2}\lambda_2 \\ \vdots \\ v_{2,n}\lambda_2 \end{pmatrix} + \cdots + \begin{pmatrix} v_{m,1}\lambda_m \\ v_{m,2}\lambda_m \\ \vdots \\ v_{m,n}\lambda_m \end{pmatrix} \\
&= \lambda_1 \underbrace{\begin{pmatrix} v_{1,1} \\ v_{1,2} \\ \vdots \\ v_{1,n} \end{pmatrix}}_{=v_1} + \lambda_2 \underbrace{\begin{pmatrix} v_{2,1} \\ v_{2,2} \\ \vdots \\ v_{2,n} \end{pmatrix}}_{=v_2} + \cdots + \lambda_m \underbrace{\begin{pmatrix} v_{m,1} \\ v_{m,2} \\ \vdots \\ v_{m,n} \end{pmatrix}}_{=v_m} \\
&= \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_m v_m.
\end{aligned}
$$

□

1.1.2. Checking linear dependence via RREF

Question: How do we check whether m vectors $v_1, v_2, \ldots, v_m$ are dependent?

We already know a way: We want to know whether there is a nontrivial relation $\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_m v_m = 0$. We translate this relation into a system in the unknowns $\lambda_1, \lambda_2, \ldots, \lambda_m$. If this system has any nontrivial solution (i.e., if it has any free variables), then the vectors $v_1, v_2, \ldots, v_m$ are dependent. If this system has only the trivial solution, then the vectors $v_1, v_2, \ldots, v_m$ are independent.

But there is a better method ([Strickland, Method 8.8]):

Theorem 1.1.7. Let $v_1, v_2, \ldots, v_m$ be m vectors in $\mathbb{R}^n$. To see whether they are dependent, the following algorithm works:

(a) Form the $n \times m$-matrix A whose columns are $v_1, v_2, \ldots, v_m$. This is usually written as follows:

$$A = [v_1 \mid v_2 \mid \cdots \mid v_m].$$

(b) Row-reduce A to get an $n \times m$-matrix B in RREF.

(c) If every column of B has a pivot, then $v_1, v_2, \ldots, v_m$ are independent.

(d) If some column of B has no pivot, then $v_1, v_2, \ldots, v_m$ are dependent. Moreover, solutions to the system $B\lambda = 0$ (where $\lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{pmatrix}$) correspond precisely to the relations between $v_1, v_2, \ldots, v_m$.

Proof. We are trying to see whether the vectors $v_1, v_2, \ldots, v_m$ are dependent. In other words, we are trying to see whether there exist numbers $\lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{R}$ that are not all 0 but satisfy $\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_m v_m = 0$. In other words, we are trying to see whether there exists a column vector $\lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{pmatrix}$ satisfying $A\lambda = 0$ (because if $\lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{pmatrix}$ is any vector, then

$$A\lambda = [v_1 \mid v_2 \mid \cdots \mid v_m] \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{pmatrix} = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_m v_m$$

(by Lemma 1.1.6), and therefore the equation $\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_m v_m = 0$ can be rewritten as $A\lambda = 0$). In other words, we are trying to see whether the system $A\lambda = 0$ of linear equations has a nonzero solution λ. Since this system always has the zero solution $\lambda = 0$ (because $A0 = 0$), this is the same as checking whether this system has a **unique** solution. (Indeed, if it has a unique solution, then it must be the zero solution, and thus our vectors $v_1, v_2, \ldots, v_m$ are independent; but if its solution is not unique, then it must have a nonzero solution, and thus our vectors $v_1, v_2, \ldots, v_m$ are dependent.)

How do we tell whether the system $A\lambda = 0$ has a unique solution? We can solve the system using the (systematic) Gaussian elimination method that we learned last time. This method starts by constructing the augmented matrix of the system, which we shall denote by A' this time because the letter A is already taken. Clearly,

$$A' = [v_1 \mid v_2 \mid \cdots \mid v_m \mid 0].$$

That is, the augmented matrix A' is A with an extra column full of zeroes attached to it on the right end.

The next step in Gaussian elimination is to bring the augmented matrix A' into RREF. But note that the last column of A' is full of zeroes. This fact does not change as we perform the row operations of Gaussian elimination; indeed, adding, scaling or switching zeroes only produces zeroes.[1] Thus, the row operations that we apply to A' only change the first m columns of A' while the last column remains filled with zeroes. This means that these row operations are tantamount to the row operations that bring the matrix $A = [v_1 \mid v_2 \mid \cdots \mid v_m]$ into RREF; the only difference is that now we have a useless column full of zeroes dangling around at the right end, which never changes.

Thus, the RREF of A' will be the matrix B with a column full of zeroes attached at the right end (since the RREF of A was B). Let me call this matrix B'. So B' is obtained from B by adding an extra column full of zeroes at the right end. Here is a (symbolic) picture of our situation:

$$\begin{array}{ccc} A & \xrightarrow{\text{bring to RREF}} & B \\ {\scriptstyle\text{attach column full of zeroes}}\Big\downarrow & & \Big\downarrow{\scriptstyle\text{attach column full of zeroes}} \\ A' & \xrightarrow[\text{bring to RREF}]{} & B' \end{array}$$

But B' is the RREF of A'; thus the Gaussian elimination method tells us the following about solutions to $A\lambda = 0$:

- A solution to $A\lambda = 0$ exists if and only if the matrix B' has no pivot in the last column.
- If B' has no pivot in the last column but a pivot in each of the other columns, then there is only one solution to $A\lambda = 0$ (since there are no free variables).
- If B' has no pivot in at least one column other than the last one, then there are multiple solutions to $A\lambda = 0$ (since there are free variables).

We can restate this in terms of B, since B is simply B' without the last column (and since the last column of B' is full of zeroes and thus has no pivot). We thus obtain the following:

---

[1] Example:

$$\begin{pmatrix} a & b & 0 \\ c & d & 0 \\ e & f & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} a & b & 0 \\ a+c & b+d & 0 \\ e & f & 0 \end{pmatrix}.$$

Thus, the last column of A', being initially a zero vector, will remain a zero vector throughout the row-reduction process, and will not impact the life of the other columns. Hence, row-reducing the augmented matrix A' is the same as row-reducing the matrix A, except that the last column is just "hanging around".

- A solution to $A\lambda = 0$ always exists. (Not surprising – it's just $\lambda = 0$.)
- If B has a pivot in each of its columns, then there is only one solution to $A\lambda = 0$ (since there are no free variables).
- If B has no pivot in at least one column, then there are multiple solutions to $A\lambda = 0$ (since there are free variables).

Now, recall that "there is only one solution to $A\lambda = 0$" means "$v_1, v_2, \ldots, v_m$ are independent", whereas "there are multiple solutions to $A\lambda = 0$" means "$v_1, v_2, \ldots, v_m$ are dependent". Thus, what we just proved is precisely the theorem. □

Example 1.1.8. Consider the three vectors

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \qquad v_2 = \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix}, \qquad v_3 = \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix}.$$

Are they dependent?

The algorithm described in the previous theorem tells us to construct the matrix

$$A = [v_1 \mid v_2 \mid v_3] = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 3 \\ 3 & 3 & 3 \end{pmatrix}$$

and to row-reduce it:

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 3 \\ 3 & 3 & 3 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & -2 & -3 \\ 3 & 3 & 3 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & -2 & -3 \\ 0 & -3 & -6 \end{pmatrix}$$
$$\longrightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 3/2 \\ 0 & -3 & -6 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 3/2 \\ 0 & 0 & -3/2 \end{pmatrix}$$
$$\longrightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 3/2 \\ 0 & 0 & 1 \end{pmatrix}.$$

We are not done row-reducing the matrix yet, but we already see that it will have a pivot in each column. Thus, the vectors v_1, v_2, v_3 are independent.

Example 1.1.9. Are the vectors

$$v_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \qquad v_2 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \qquad v_3 = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

dependent?

We can proceed by the same method:

$$A = [v_1 \mid v_2 \mid v_3] = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \end{pmatrix}.$$

Again, we have not row-reduced the matrix yet, but we already see the pivots, and in particular we see that there won't be a pivot in the last column. Thus, the vectors v_1, v_2, v_3 are dependent.

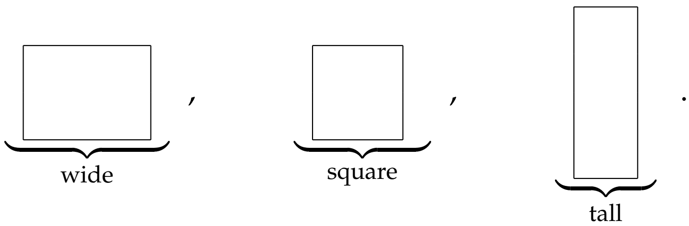
I claim that this was clear even earlier. We didn't have to do any row operations! Why not? Because A was a 2×3-matrix. It has more columns than rows. Of course, this will still hold after row-reduction. Hence, if each column had a pivot, then the matrix would have 3 pivots, which would thus lie in 3 different rows (since each pivot has to lie in a different row). But A does not have 3 different rows to begin with. So there is at least one column without pivot. Thus, v_1, v_2, v_3 are dependent.

This generalizes:

Definition 1.1.10. An $m \times n$-matrix is said to be

- **wide** if $n > m$;
- **square** if $n = m$;
- **tall** if $n < m$.

These words refer to how the matrix looks like:



Let us now generalize what we have observed in the last example:

Proposition 1.1.11. The RREF of a wide matrix cannot have a pivot in each column.

Proof. The RREF of a wide matrix is still a wide matrix. If it had a pivot in each column, then it would have as many pivots as it has columns, and thus it would have at least as many rows as it has columns (since each pivot must lie in a different row). But this is not true for a wide matrix. □

Corollary 1.1.12. Any m vectors in $\mathbb{R}^n$ are dependent if $m > n$.

Proof. Apply the method above to check whether these m vectors are dependent. The matrix B will be wide, so by the previous proposition it cannot have a pivot in each column. According to the method, this means that the m vectors are dependent. □

So we can answer certain dependence questions without even computing the RREF.

Example 1.1.13. Any 16 vectors in $\mathbb{R}^{15}$ are dependent.

Let us extend the last proposition a little bit ([Strickland, Lemma 8.7]):

Lemma 1.1.14. Let B be a $p \times q$-matrix in RREF.

(a) If B is wide, then B cannot have a pivot in each column.

(b) If B is square, then the only way for B to have a pivot in each column is when $B = I_p$.

(c) The only way for B to have a pivot in each column is when $p \geq q$ and

$$B = \begin{bmatrix} I_q \\ 0_{(p-q)\times q} \end{bmatrix}.$$

(The right hand side of this equality is to be understood as the $p \times q$-matrix obtained by piling the $q \times q$ identity matrix I_q on top of the $(p-q) \times q$ zero matrix $0_{(p-q)\times q}$. For example, if $q = 2$ and $p = 5$, then it means $\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$. If $p = q$, then it simply means I_q.)

Proof. Assume that B has a pivot in each column. Thus, B has q pivots in total (since B has q columns, and there can only be 1 pivot per column). Property RREF3 requires each pivot of an RREF matrix to "clear out" its column (in the sense that all other entries in its column are 0's). Thus, all non-pivot entries of B are 0's (since B has a pivot in each column). Due to property RREF2, we furthermore know that the pivots move right as we move down the matrix. Thus,

- the topmost pivot has to be in column 1 (since otherwise, column 1 would have no pivot);
- the second-topmost pivot has to be in column 2 (since otherwise, column 2 would have no pivot);

- the third-topmost pivot has to be in column 3 (since otherwise, column 3 would have no pivot);
- ....

Also, property RREF0 shows that these pivots must be in the first q rows of B, since all rows with no pivot are zero rows. Thus, the first q rows of B have pivots: row 1 has a pivot in column 1; row 2 has a pivot in column 2; row 3 has a pivot in column 3; and so on. The remaining rows are zero rows. This shows that $p \geq q$ (because B must have q rows to begin with) and $B = \begin{bmatrix} I_q \\ 0_{(p-q)\times q} \end{bmatrix}$ (since all non-pivot entries of B are 0's). This proves part **(c)** of the lemma.

Part **(b)** is just the particular case of part **(c)** for $p = q$.

Part **(a)** follows from the $p \geq q$ part of part **(c)** (since "B is wide" means the same as "$p < q$"). □

1.2. Spans and spanning

Proposition 1.2.1. Let $v_1, v_2, \ldots, v_k$ be some vectors in $\mathbb{R}^n$. Then, any combination of combinations of $v_1, v_2, \ldots, v_k$ is a combination of $v_1, v_2, \ldots, v_k$.

Proof. For example, any combination of **two** combinations of $v_1, v_2, \ldots, v_k$ is a combination of $v_1, v_2, \ldots, v_k$, because

$$\alpha \left(\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k\right) + \beta \left(\mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_k v_k\right) \\ = \left(\alpha\lambda_1 + \beta\mu_1\right) v_1 + \left(\alpha\lambda_2 + \beta\mu_2\right) v_2 + \cdots + \left(\alpha\lambda_k + \beta\mu_k\right) v_k.$$

A similar argument works for more than two combinations. □

Definition 1.2.2. Let $v_1, v_2, \ldots, v_k$ be some vectors in $\mathbb{R}^n$.

(a) The **span** of $v_1, v_2, \ldots, v_k$ is the set of all combinations of $v_1, v_2, \ldots, v_k$.

This is a subset of $\mathbb{R}^n$, and is closed under combination, as the previous proposition shows: i.e., any combination of elements in this span must also lie in this span.

The span of $v_1, v_2, \ldots, v_k$ is called $\operatorname{span}(v_1, v_2, \ldots, v_k)$ (or sometimes $\langle v_1, v_2, \ldots, v_k \rangle$).

(b) We say that the vectors $v_1, v_2, \ldots, v_k$ (or, more precisely, the list $(v_1, v_2, \ldots, v_k)$) **span** $\mathbb{R}^n$ if and only if $\operatorname{span}(v_1, v_2, \ldots, v_k) = \mathbb{R}^n$. In other words, they span $\mathbb{R}^n$ if and only if each vector in $\mathbb{R}^n$ is a combination of $v_1, v_2, \ldots, v_k$.

(In part **(a)**, the word "span" is a noun; in part **(b)**, it is a verb.)

Example 1.2.3. Consider the list $(v_1, v_2, v_3, v_4, v_5)$ where

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}, \quad v_4 = \begin{pmatrix} 1 \\ 4 \\ 1 \end{pmatrix}, \quad v_5 = \begin{pmatrix} 1 \\ 5 \\ 1 \end{pmatrix}.$$

What is $\operatorname{span}(v_1, v_2, \ldots, v_5)$?

An arbitrary combination of $v_1, v_2, \ldots, v_5$ is

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_5 v_5 = \begin{pmatrix} \lambda_1 + \lambda_2 + \cdots + \lambda_5 \\ \lambda_1 + 2\lambda_2 + 3\lambda_3 + 4\lambda_4 + 5\lambda_5 \\ \lambda_1 + \lambda_2 + \cdots + \lambda_5 \end{pmatrix}.$$

This is a vector whose first entry equals its third entry. Thus, not every vector in $\mathbb{R}^3$ can be a combination of $v_1, v_2, \ldots, v_5$. Hence, $v_1, v_2, \ldots, v_5$ do not span $\mathbb{R}^3$.

What is $\operatorname{span}(v_1, v_2, \ldots, v_5)$? We have just shown that

$$\operatorname{span}(v_1, v_2, \ldots, v_5) \subseteq \left\{ w \in \mathbb{R}^3 \mid \text{the first entry of } w \text{ equals its third entry} \right\}.$$

Do we have

$$\operatorname{span}(v_1, v_2, \ldots, v_5) = \left\{ w \in \mathbb{R}^3 \mid \text{the first entry of } w \text{ equals its third entry} \right\}$$

as well? In other words, can **each** vector in $\mathbb{R}^3$ whose first entry equals its third entry be written as a combination of $v_1, v_2, \ldots, v_5$?

Let's try: Given such a vector $\begin{pmatrix} a \\ b \\ a \end{pmatrix}$, we want to write it as $\begin{pmatrix} a \\ b \\ a \end{pmatrix} = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_5 v_5$. This means solving the system

$$\begin{cases} \lambda_1 + \lambda_2 + \cdots + \lambda_5 = a \\ \lambda_1 + 2\lambda_2 + 3\lambda_3 + 4\lambda_4 + 5\lambda_5 = b \\ \lambda_1 + \lambda_2 + \cdots + \lambda_5 = a \end{cases}.$$

This system always has a solution – for example $\lambda_1 = 2a - b$ and $\lambda_2 = b - a$ and $\lambda_3 = \lambda_4 = \lambda_5 = 0$. So we indeed have

$$\operatorname{span}(v_1, v_2, \ldots, v_5) = \left\{ w \in \mathbb{R}^3 \mid \text{the first entry of } w \text{ equals its third entry} \right\}.$$

Question: How do we find out whether a given list of vectors spans $\mathbb{R}^n$?

Simple but slow answer: We can try to write each of the standard basis vectors $e_1, e_2, \ldots, e_n$ (see HW2 Exercise 4) as a combination of this list.

(Remember: The standard basis vectors $e_1, e_2, \ldots, e_n$ are defined by $e_i = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ with the 1 being placed in position i.)

If any of $e_1, e_2, \ldots, e_n$ is not a combination of our list, then our list does not span $\mathbb{R}^n$.

If all of $e_1, e_2, \ldots, e_n$ are combinations of our list, then we conclude that each vector in $\mathbb{R}^n$ is a combination of our list as well (because HW2 Exercise 4 shows that each vector in $\mathbb{R}^n$ is a combination of $e_1, e_2, \ldots, e_n$), and therefore our list does span $\mathbb{R}^n$.

However, this method is slow, since it requires solving n systems of equations.

Here is a faster method ([Strickland, Method 9.7]):

Theorem 1.2.4. Let $\mathcal{V} = (v_1, v_2, \ldots, v_m)$ be a list of vectors in $\mathbb{R}^n$. We can check whether this list spans $\mathbb{R}^n$ as follows:

(a) Form the $m \times n$-matrix

$$C = \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_m^T \end{bmatrix}.$$

This is the matrix whose rows are $v_1^T, v_2^T, \ldots, v_m^T$ (the transposes of the vectors in $\mathcal{V}$).

(b) Row-reduce C to get a matrix D in RREF.

(c) If D has a pivot in each column, then $\mathcal{V}$ spans $\mathbb{R}^n$.

(d) If D has no pivot in some column, then $\mathcal{V}$ does not span $\mathbb{R}^n$.

The proof of this theorem relies on tracking down what happens to the span of the rows of a matrix[2] when we apply row operations to the matrix:

Lemma 1.2.5. Let $v_1^T, v_2^T, \ldots, v_m^T$ be the rows of a matrix C, and let $w_1^T, w_2^T, \ldots, w_m^T$ be the rows of a matrix C' that is obtained from C by row operations. Then,

$$\operatorname{span}(v_1, v_2, \ldots, v_m) = \operatorname{span}(w_1, w_2, \ldots, w_m).$$

In other words, row operations do not change the span of the rows.

[2]We are being slightly indirect here: We have defined the span for column vectors only, but the rows of a matrix are row vectors. So, to be precise, we should speak not of the span of the rows, but of the span of the **transposes** of the rows. (The transpose of a row vector is a column vector.) This explains why you see so many "^{T}" signs in the following lemma.

Proof. Clearly, it suffices to show that any **single** row operation does not change the span of the rows.

In other words, we have to show that any column vectors $v_1, v_2, \ldots, v_m$ satisfy:

1. $\operatorname{span}(\lambda v_1, v_2, v_3, \ldots, v_m) = \operatorname{span}(v_1, v_2, v_3, \ldots, v_m)$ for any $\lambda \neq 0$.
2. $\operatorname{span}(v_1 + \lambda v_2, v_2, v_3, \ldots, v_m) = \operatorname{span}(v_1, v_2, v_3, \ldots, v_m)$ for any $\lambda \in \mathbb{R}$.
3. $\operatorname{span}(v_2, v_1, v_3, \ldots, v_m) = \operatorname{span}(v_1, v_2, v_3, \ldots, v_m)$.

(At least, this corresponds to the row operations "scaling the first row by λ", "adding λ times the second row to the first row" and "swapping the first two rows". To be honest, we would have to show this not just for the first two rows, but also for any pair of rows; but the arguments are the same.)

Proof of 1: Any linear combination

$$\mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_m v_m \qquad \text{of } v_1, v_2, \ldots, v_m$$

will still be a linear combination of $\lambda v_1, v_2, v_3, \ldots, v_m$, since it can be rewritten as

$$\left(\mu_1 \lambda^{-1}\right)(\lambda v_1) + \mu_2 v_2 + \cdots + \mu_m v_m.$$

Conversely, any linear combination

$$\mu_1 (\lambda v_1) + \mu_2 v_2 + \cdots + \mu_m v_m \qquad \text{of } \lambda v_1, v_2, \ldots, v_m$$

will be a linear combination of $v_1, v_2, \ldots, v_m$, since it can be rewritten as

$$(\mu_1 \lambda) v_1 + \mu_2 v_2 + \cdots + \mu_m v_m.$$

Thus, the linear combinations of $v_1, v_2, \ldots, v_m$ are the same as the linear combinations of $\lambda v_1, v_2, \ldots, v_m$. In other words, $\operatorname{span}(v_1, v_2, \ldots, v_m) = \operatorname{span}(\lambda v_1, v_2, \ldots, v_m)$.

Proof of 2: Any linear combination

$$\mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_m v_m \qquad \text{of } v_1, v_2, \ldots, v_m$$

will still be a linear combination of $v_1 + \lambda v_2, v_2, v_3, \ldots, v_m$, since it can be rewritten as

$$\mu_1 (v_1 + \lambda v_2) + (\mu_2 - \lambda \mu_1) v_2 + \mu_3 v_3 + \cdots + \mu_m v_m.$$

Similarly, the converse direction can be shown. Thus, $\operatorname{span}(v_1 + \lambda v_2, v_2, v_3, \ldots, v_m) = \operatorname{span}(v_1, v_2, v_3, \ldots, v_m)$.

Proof of 3: Any linear combination

$$\mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_m v_m \qquad \text{of } v_1, v_2, \ldots, v_m$$

will still be a linear combination of $v_2, v_1, v_3, \ldots, v_m$, since it can be rewritten as

$$\mu_2 v_2 + \mu_1 v_1 + \mu_3 v_3 + \cdots + \mu_m v_m.$$

Similarly, the converse direction can be shown. Thus, $\operatorname{span}(v_2, v_1, v_3, \ldots, v_m) = \operatorname{span}(v_1, v_2, v_3, \ldots, v_m)$. □

We just showed that when we apply row operations to a matrix, the span of its rows does not change. This means that the span of its rows is the span of the rows of its RREF (since the RREF is obtained by row operations).

Next, we claim the following lemma ([Strickland, Lemma 9.17]):

Lemma 1.2.6. Let D be an $m \times n$-matrix in RREF.

(a) If every column of D contains a pivot, then the transposes of the rows of D span $\mathbb{R}^n$.

(b) If some column of D does not contain a pivot, then the transposes of the rows of D do not span $\mathbb{R}^n$.

Proof. **(a)** Assume that every column of D contains a pivot. Then, Lemma 1.1.14 **(c)** shows that $D = \begin{bmatrix} I_n \\ 0_{(m-n)\times n} \end{bmatrix}$ (and $m \geq n$). Hence, the transposes of the rows of D are $e_1, e_2, \ldots, e_n, \underbrace{0, 0, \ldots, 0}_{m-n \text{ zero vectors}}$. These m vectors do span $\mathbb{R}^n$, since the $e_1, e_2, \ldots, e_n$ span $\mathbb{R}^n$ already (by homework set #2 Exercise 4).

(b) Assume that column i of D does not contain a pivot. I claim that the transposes of the rows of D do not span $\mathbb{R}^n$. More specifically, I am claiming that the standard basis vector $e_i \in \mathbb{R}^n$ is not a combination of the transposes of the rows of D.

Indeed, assume that e_i is a combination of the transposes of the rows of D. That is,

$$e_i = \lambda_1 \left(\operatorname{row}_1 D\right)^T + \lambda_2 \left(\operatorname{row}_2 D\right)^T + \cdots + \lambda_m \left(\operatorname{row}_m D\right)^T$$

for some $\lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{R}$. Clearly, not all of $\lambda_1, \lambda_2, \ldots, \lambda_m$ are 0, since then this equality would say $e_i = 0$ (which is absurd). Thus, some λ_j is $\neq 0$. Pick the **smallest** $j \in \{1, 2, \ldots, m\}$ such that $\lambda_j \neq 0$. Thus, $\lambda_1 = \lambda_2 = \cdots = \lambda_{j-1} = 0$. We furthermore assume that D has no zero rows, since we could otherwise simply remove them and nothing important would change (since zero vectors do not contribute to a span: $\operatorname{span}(v_1, v_2, \ldots, v_k, 0, 0, \ldots, 0) = \operatorname{span}(v_1, v_2, \ldots, v_k)$). Hence, each row of D has a pivot somewhere, and each row's pivot is further right than the previous row's pivot (by property RREF2). Now,[3]

$$\begin{aligned} e_i &= \lambda_1 \left(\operatorname{row}_1 D\right)^T + \lambda_2 \left(\operatorname{row}_2 D\right)^T + \cdots + \lambda_m \left(\operatorname{row}_m D\right)^T \\ &= \lambda_j \left(\operatorname{row}_j D\right)^T + \lambda_{j+1} \left(\operatorname{row}_{j+1} D\right)^T + \cdots + \lambda_m \left(\operatorname{row}_m D\right)^T \\ &\qquad \left(\begin{array}{c} \text{here, we have thrown the first } j-1 \text{ addends away,} \\ \text{since they are all 0 (because } \lambda_1 = \lambda_2 = \cdots = \lambda_{j-1} = 0) \end{array}\right) \end{aligned}$$

[3] We put asterisks ("$*$") to signify entries whose values we don't care about.

$$= \lambda_j \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ * \\ * \\ * \\ \vdots \\ * \end{pmatrix} + \lambda_{j+1} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ * \\ * \\ \vdots \\ * \end{pmatrix} + \cdots + \lambda_m \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ * \\ \vdots \end{pmatrix}$$

$$\left(\begin{array}{c} \text{because } D \text{ is a RREF matrix with no zero rows,} \\ \text{so each row of } D \text{ has a pivot equal to 1,} \\ \text{and each row's pivot is further right than the previous row's pivot} \end{array} \right)$$

$$= \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \lambda_j \\ * \\ * \\ \vdots \\ * \end{pmatrix}.$$

Thus, the first nonzero entry of the vector e_i is in the same position as the first nonzero entry in $\left(\operatorname{row}_j D\right)^T$ (because the $\lambda_j \left(\operatorname{row}_j D\right)^T$ addend is a vector whose first nonzero entry is in this position, whereas all the other addends have their first nonzero entries further down). But this is of course the position in which the j-th row of D has its pivot. Thus, it cannot be in position i (since D has no pivot in the i-th column). But this contradicts the fact that the first nonzero entry in e_i is in position i. □

From the last two lemmas, our theorem follows: The second-to-last lemma shows that row operations don't change the span of the transposes of the rows; the last lemma explains how we can tell from the RREF whether this span is $\mathbb{R}^n$ or not.

Corollary 1.2.7. A list of m vectors can never span $\mathbb{R}^n$ if $m < n$.

Proof. Consider some m vectors $v_1, v_2, \ldots, v_m$ in $\mathbb{R}^n$, where $m < n$. Apply the method above to check whether these m vectors span $\mathbb{R}^n$. The matrix D will be wide, so it cannot have a pivot in each column. According to the method, this means that the m vectors cannot span $\mathbb{R}^n$. □

References

[Strickland] Neil Strickland, *MAS201 Linear Mathematics for Applications*, lecture notes, 28 September 2013.
http://neil-strickland.staff.shef.ac.uk/courses/MAS201/