LECTURES 14/15: LINEAR INDEPENDENCE AND BASES

MA1111: LINEAR ALGEBRA I, MICHAELMAS 2016

1. Linear Independence

We have seen in examples of span sets of vectors that sometimes adding additional vectors doesn't increase the span of a set of vectors. For example, if v is a vector in $\mathbb{R}^3$, then $\operatorname{span}(v) = \operatorname{span}(v, 2v)$. We want to develop language to exclude such cases and throw away redundant information in the description of subspaces built out of span sets. To this end, we set the following terminology.

Definition. We say that a non-empty set S of vectors in V is **linearly dependent** if there are vectors $v_1, \ldots, v_n \in V$ and scalars $c_1, \ldots, c_n \in F$ **not all equal to zero** for which

$$c_1 v_1 + \ldots + c_n v_n = 0.$$

If no such **non-trivial linear dependency** exists, we say that the set S is **linearly independent**.

Example. *A set with only one non-zero vector is linearly independent, as if $cv = 0$, then we saw before that $c = 0$ or $v = 0$, and $v \neq 0$ by assumption. Thus, any equation must have $c = 0$, and then it violates the condition of not all scalars in a linear dependency being* 0.

Example. *In $\mathbb{R}^3$, for any vector v the set $S = \{e_1, e_2, e_3, v\}$ is linearly dependent, as we can write any such vector as $v = c_1e_1 + c_2e_2 + c_3e_3$ for real numbers c_1, c_2, c_3, and thus*

$$-c_1e_1 - c_2e_2 - c_3e_3 + v = 0,$$

and the coefficient in front of v is $1 \neq 0$.

Example. *Any set which contains the zero vector is linearly dependent. For example, we have the linear dependency $1 \cdot 0 = 0$.*

Example. *By definition the empty set $\emptyset$ is always linearly independent as there are no possible linear combinations in the definition above to check!*

As we have seen, properties about linear combinations of vectors can be expressed in terms of solution sets to systems of linear equations. In the case of linear independence, suppose that we wish to determine whether $S = \{v_1, \ldots, v_n\}$ is linearly independent,

Date: November 2, 4, 2016.

where $v_i = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix} \in \mathbb{R}^m$. Then we are asking whether there are scalars $c_1, \ldots, c_n$, not all zero, for which

$$c_1 v_1 + \ldots + c_n v_n = c_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \ldots + c_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} = 0,$$

or

$$Ac = 0$$

with A defined by $A_{ij} = a_{ij}$ (i.e., A is the matrix whose columns are $v_1, \ldots, v_n$) and $c = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$, and we only require that $c \neq 0$. That is, a finite set of vectors in $\mathbb{R}^n$ is linearly independent if and only if the matrix A with those vectors as its columns has $\ker(A) = \{0\}$. As performing elementary row operations on an augmented matrix preserves the solution sets, and as any elementary row operations keep the vector 0 unchanged, this is equivalent to there being a non-trivial kernel of the RREF of A, which happens exactly when there is a column without a pivot in the RREF of A. That is, we have shown the following.

Theorem. *The following are equivalent for a set of vectors $v_1, \ldots, v_n \in \mathbb{R}^m$, where $A = \begin{pmatrix} v_1 & \ldots & v_n \end{pmatrix}$ is the matrix with columns $v_1, \ldots, v_n$.*

(1) *The vectors $v_1, \ldots, v_n$ are linearly independent.*
(2) *The kernel of A is the trivial subspace $\{0\}$.*
(3) *The kernel of the RREF of A is the trivial subspace $\{0\}$.*
(4) *The RREF of A has a pivot in every column.*

In particular, if $n = m$, that is, if the matrix A is square, this is equivalent to $\det A \neq 0$.

Example. *In $\mathbb{R}^3$, the set of vectors $v_1 = (1, 4, 7)$, $v_2 = (2, 5, 8)$, $v_3 = (3, 6, 9)$ is linearly dependent as*

$$\det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = 0.$$

However, the set of vectors $v_1 = (1, 4, 7)$, $v_2 = (2, 5, 0)$, $v_3 = (3, 6, 9)$ is linearly independent as

$$\det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 0 & 9 \end{pmatrix} = -48 \neq 0.$$

Note that this doesn't say exactly the same things as the example in the last lecture when we discussed the spans of these vectors. However, we will see that spans of n vectors in

$\mathbb{R}^n$ are all of $\mathbb{R}^n$ if and only if they are linearly independent. This is essentially a special case of the Fundamental Theorem of Linear Algebrta.

Example. *In $\mathbb{R}^4$, the vectors $v_1 = (1,4,7,1)$, $v_2 = (2,5,8,2)$, $v_3 = (3,6,9,3)$ are linearly dependent as the RREF of*

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 1 & 2 & 3 \end{pmatrix}$$

is

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

which has 2 pivots and 3 columns, and hence there are infinitely many solutions to $Ax = 0$. However, the vectors $v_1 = (1,4,7,1)$, $v_2 = (2,5,8,2)$, $v_3 = (3,6,9,4)$ are linearly independent as the RREF of

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 1 & 2 & 4 \end{pmatrix}$$

is

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

which has a pivot in each column, and hence trivial kernel.

Example. *If we have a set of n vectors $v_1 \dots, v_n$ in $\mathbb{R}^m$, if $n > m$, then they must be linearly dependent as the corresponding matrix A has n columns, but only m rows. In order to be linearly independent, there must be a pivot in each column, that is there must be n pivots. However, there can only be one pivot in each row, so there are at most $m < n$ pivots.*

Example. *By the last example, it is automatic that the set of vectors $v_1 = (1,4,7)$, $v_2 = (2,5,8)$, $v_3 = (3,6,9), v_4 = (-1,4,-5)$ is linearly dependent, and we don't have to do any work with row reduction to determine this.*

2. Bases of vector spaces

We now combine the notions of span and linear independence to describe set of vectors which build up vector spaces without "redundant information."

Definition. A **basis** for a vector space V is a set $S \subset V$ which is linearly independent and which spans V.

Example. *We defined the span of the empty set $\emptyset$ to be $\{0\}$, and clearly there can be no linear dependence relations in a set with no elements, and so $\emptyset$ is a basis for $\{0\}$.*

Example. *In $\mathbb{R}^n$, we have regularly referred to the vectors $e_1, \ldots, e_n$ as standard unit basis vectors. They do indeed form a basis of $\mathbb{R}^n$, as we have repeatedly seen that they span $\mathbb{R}^n$, and they are linearly independent as the matrix whose columns are the standard unit basis vectors is the identity matrix I_n, which of course has non-zero determinant. The same construction, setting e_i to be the vector with a 1 in the i-th position and zero elsewhere, provides a basis of F^n for any field F.*

Example. *The set $\{1, x, x^2, \ldots\}$ of all non-negative integral powers of x is a basis of $\mathcal{P}(F)$.*

Example. *The set $\{M^{ij} : 1 \leq i \leq m,\ 1 \leq j \leq n\}$ where M^{ij} is the $m \times n$ matrix with entry 1 in the (i, j)-th entry and 0 elsewhere is a basis of the space of matrices $M_{m\times n}(F)$.*

The idea of a basis providing the smallest possible set to describe all elements of a vector space (i.e., without redundant information) is made precise by the following result.

Theorem. *A set $S \subset V$ is a basis for V if and only if every element of V can be expressed* **uniquely** *as a linear combination of elements of S.*

Proof. Suppose that S is a basis for V. Then by definition $\text{span}(S) = V$, so every element of V can be written as a linear combination of elements of S. To show the uniqueness of this expression, suppose that $v \in V$ has two representations as linear combinations of elements $v_1, \ldots, v_n \in S$:

$$v = c_1v_1 + \ldots + c_nv_n = c'_1v_1 + \ldots + c'_nv_n.$$

Taking the difference of these two representations, we find that

$$(c_1 - c'_1)v_1 + \ldots + (c_n - c'_n)v_n = 0,$$

and since S is linearly independent we have $c_1 - c'_1 = 0, \ldots, c_n - c'_n = 0$. Hence, $c_1 = c'_1, \ldots, c_n = c'_n$, and so the two representations of v were the same all along.

Conversely, if every element of V is a unique linear combination of elements of S, then clearly S spans V. To show that S is linearly independent, suppose that

$$c_1v_1 + \ldots + c_nv_n = 0$$

with $v_1 \ldots, v_n \in S$ and with not all $c_1, \ldots, c_n$ equal to zero. Then we can find many non-unique representations of elements of V as linear combinations of elements from S. For example, $v_1 = 1 \cdot v_1 = (c_1 + 1)v_1 + \ldots + c_nv_n$ gives two different representations of v_1.

$\square$

We will later return to this fact, and the coefficients in these unique representations will be called the **coordinates** of a vector with respect to a basis. In fact, this idea will allow us to show that all vector spaces spanned by a finite number of vectors (this is most of them that we have seen, other than vector spaces built out of functions like $\mathcal{P}(\mathbb{R})$ and $\mathcal{C}^\infty$) are really the same thing as the "simple" vector space F^n for some n.

Our two most important examples of computing bases are for our two most important subspaces: kernels (or null spaces) and column spaces (or spans of vectors in $\mathbb{R}^n$). We now discuss each case.

Example. *The kernel of a matrix can be found using row reduction, as we have seen. In particular, we have seen above that the kernel of a matrix A is the same as the kernel of the RREF of A, and we can use our Gaussian elimination algorithm to easily find a basis for the kernel of any matrix. For example, we saw above that*

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 1 & 2 & 3 \end{pmatrix}$$

has a non-trivial kernel. This is the same as the kernel of the RREF of A, which we have seen is

$$R = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The solutions to the equation $R = 0$ are obtained by setting the free variable $x_3 = t$ for $t \in \mathbb{R}$, and solving to get the one-parameter infinite set of solutions

$$x_1 = t, \qquad x_2 = -2t, \qquad x_3 = t,$$

which is a line in $\mathbb{R}^3$ with basis $(1, -2, 1)$.

As another example, consider the matrix

$$A = \begin{pmatrix} 1 & 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 2 & 1 \end{pmatrix}.$$

This matrix has RREF

$$R = \begin{pmatrix} 1 & 0 & 0 & -2 & -2 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 2 & 3 \end{pmatrix}.$$

The solution set to $Rx = 0$ is found by setting $x_4 = t_4$, $x_5 = t_5$ for $t_4, t_5 \in \mathbb{R}$, and solving for the pivotal variables:

$$x_1 = 2t_4 + 2t_5, \quad x_2 = -t_4, \quad x_3 = -2t_4 - 3t_5, \quad x_4 = t_4, \quad x_5 = t_5.$$

That is, the the vectors in the kernel are those vectors $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix}$ *of the form*

$$x = t_4 \begin{pmatrix} 2 \\ -1 \\ -2 \\ 1 \\ 0 \end{pmatrix} + t_5 \begin{pmatrix} 2 \\ 0 \\ -3 \\ 0 \\ 1 \end{pmatrix},$$

i.e., the span of

$$\left\{ \begin{pmatrix} 2 \\ -1 \\ -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ -3 \\ 0 \\ 1 \end{pmatrix} \right\}$$

is $\ker(A)$. *It is also easy to check that these two vectors are linearly independent, so that they are indeed a basis for* $\ker(A)$. *In general, the basis for the kernel of a matrix will always have the same number of elements in it as the number of free variables in its RREF, i.e., the vectors that this method yields will always form a linearly independent set.*

The second main example of a vector space we would like to find bases for is a subspace spanned by a set of vectors in $\mathbb{R}^n$. We will give two methods for this. We can solve this problem by computing the row space or the column space of a matrix A; of course, we can switch between the two perspectives using transposes.

We begin with the case of row space, the span of the rows a a matrix.

Theorem. *The row space of A is the same as the row space of the RREF of A.*

Proof. It suffices to show that elementary row operations don't change row space. Indeed, suppose that the rows of a matrix A are $v_1, \dots, v_n$. If we multiply row j by a non-zero constant c, this then the row space is of the form

$$\operatorname{span}(v_1, \dots, cv_j, \dots, v_n) = \{c_1 v_1 + \ldots + (c_j c) v_j + \ldots + c_n v_n : c_1, \dots, c_n \in F\},$$

which is the same as the original span since as c_j ranges over all elements of F so does $c_j c$, as $c \neq 0$ (think of the case of real numbers; this is true since in a field we can always divide by non-zero elements). Clearly, if we swap two rows, this doesn't change the span, as addition in vector spaces is commutative. Finally, if we add a multiple of one row to another it doesn't change the span. For example, by swapping rows suppose that we add c times the second row to the first, with $c \neq 0$. Then the row space of the resulting matrix is

$$\operatorname{span}(v_1, v_2 + cv_1, \ldots, v_n) = \{c_1v_1 + c_2(v_2 + cv_1) + \ldots + c_nv_n : c_1, \ldots, c_n \in F\}.$$

Now $c_1v_1 + c_2(v_2 + cv_1) + \ldots + c_nv_n = (c_1 + c_2c)v_1 + c_2v_2 + \ldots + c_nv_n$, and clearly as c_1, c_2 range over all choices of n elements of F, so does $c_1 + c_2c$. $\square$

In general, the column space of a matrix isn't the same as the column space of its RREF. But, you can always find the column space by using elementary column operations or by taking the transpose of the matrix and then using row reduction to find the row space.

Moreover, it isn't hard to show the following.

Theorem. *The non-zero rows in a reduced row echelon form matrix are linearly independent. Thus, the non-zero rows of the RREF of a matrix A form a basis of the row space of A.*

Another method is to use pivots in row recution to find which column vectors in matrix are linearly independent. To see this, suppose that a number of columns $v_{i_1}, \ldots, v_{i_n}$ of a matrix have a non-trivial linear dependency $c_1v_{i_1} + \ldots + c_nv_{i_n} = 0$. Note that $v_{i_j} = Ae_{i_j}$, where $e_1, \ldots$ are the usual standard basis vectors. Then if $E_1, \ldots, E_k$ are elementary matrices with $R = E_1 \cdots E_kA$ where R is the RREF of A we clearly have

$$\begin{aligned} 0 &= (E_1, \cdots, E_k)(c_1v_{i_1} + \ldots + c_nv_{i_n}) = (E_1, \cdots, E_k)(c_1Ae_{i_1} + \ldots + c_nAe_{i_n}) \\ &= c_1(E_1, \cdots, E_k)Ae_{i_1} + \ldots + c_n(E_1, \cdots, E_k)Ae_{i_n} = c_1Re_{i_1} + \ldots + c_nRe_{i_n}, \end{aligned}$$

so that the same linear relation holds between the corresponding columns of R. The converse is also clearly true, as elementary matrices are always invertible. Thus, since the column space of a matrix in RREF is just the span of its pivotal columns, we have shown the following result.

Theorem. *The column space of a matrix A has as a basis the set of column vectors corresponding to columns with a pivot in the RREF of A.*

Example. *Let's compute bases for the kernel, column space, and row space of a big matrix. Take*

$$A = \begin{pmatrix} 3 & 4 & 0 & 7 \\ 1 & -5 & 2 & -2 \\ -1 & 4 & 0 & 3 \\ 1 & -1 & 2 & 2 \end{pmatrix}.$$

The first step is to compute the RREF of A, which is

$$R = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

There are 3 pivotal variables and one free variable, so the kernel will have a basis with 1 *element and a column space with a basis of* 3 *elements. The kernel is found by solving* $Rx = 0$*, which has solution set*

$$x_1 = -t, \quad x_2 = -t, \quad x_3 = -t, \quad x_4 = t,$$

which is clearly $\ker(A) = \operatorname{span}(-1, -1, -1, 1)$*, with basis* $\{(-1, -1, -1, 1)\}$.

We have already basically found the column space as well, as the RREF above has pivots in the first three columns and so $\operatorname{col}(A)$ *has as a basis* $\{(3, 1, -1, 1), (4, -5, 4, -1), (0, 2, 0, 2)\}$.

We can find another basis for the column space by noting that it is the same as the row space of A^T*, and we compute that the RREF of* A^T *is*

$$\begin{pmatrix} 1 & 0 & 0 & \frac{1}{4} \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & \frac{3}{4} \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Thus, the row space of this matrix, and hence the column space of our original matrix A *also has basis* $\{(1, 0, 0, 1/4), (0, 1, 0, 1), (0, 0, 1, 3/4)\}$.

Finally, we can find the row space of our original matrix A *by noting that it has as a basis the non-zero rows of* R *above, namely* $\{(1, 0, 0, 1), (0, 1, 0, 1), (0, 0, 1, 1)\}$.

Summarizing our results on kernels, column spaces, and row spaces, we have the following.

Theorem. *For an* $m \times n$ *matrix* A*, the number of free variables in the RREF of* A *is the number of elements in a basis for the kernel* $\ker(A)$*. Moreover, the columns of* A *are linearly independent if and only if the kernel is the trivial subspace* $\{0\}$*, which is true if and only if the RREF of* A *has a pivot in each column.*

Furthermore, the set of non-zero rows in the RREF of A *forms a basis of the row space, and the set of columns corresponding to pivotal variables of the RREF of* A *forms a basis of the column space of* A*. In particular, the row space and column space both have a basis which has as its number of elements the number of pivots in the RREF of* A*. Furthermore, the columns of* A *span all of* $\mathbb{R}^m$ *if and only if there is a pivot in each row.*

Finally, for a square $n \times n$ *matrix* A*, the following are equivalent.*

(1) *The columns of* A *span* $\mathbb{R}^n$
(2) *The columns of* A *are linearly independent.*
(3) *The columns of* A *form a basis of* $\mathbb{R}^n$.

(4) *The rows of A form a basis of* $\mathbb{R}^n$.
(5) *The matrix A is invertible.*
(6) $\det(A) \neq 0$.