LECTURE 4

Elementary Operations and Matrices

In the last lecture we developed a procedure for simplying the set of generators for a given subspace of the form

$$S = span_{\mathbb{F}}(v_1, \dots, v_k) := \{\alpha_1 v_1 + \cdots + \alpha_k v_k \mid \alpha_1, \dots, \alpha_k \in \mathbb{F}\}$$

It went like this

- Find a dependence relation amongst the vectors $v_1, \dots, v_k$; that is to say, a valid equation of the form
$$\alpha_1 v_1 + \cdots + \alpha_k v_k = \mathbf{0}_V$$
with at least on coefficient, say α_i, not equal to $0_{\mathbb{F}}$. We'll then have
$$S = span_{\mathbb{F}}(v_1, \dots, v_k) = span_{\mathbb{F}}(v_1, \dots, \widehat{v_i}, \dots, v_k)$$
(here a "$\widehat{\ }$" over a list entry is to signify that this element **does not** appear in the list) and so we can express S as being generated by $k-1$ vectors.
- Find a dependence relation amongst the vectors $v_1, \dots, \widehat{v_i}, \dots, v_k$, and use that to eliminate one more vector, say v_j, from the list of geneators
$$\Rightarrow \quad S = span_{\mathbb{F}}(v_1, \dots, \widehat{v_i}, \dots, \widehat{v_j}, \dots, v_k)$$
- Keep repeating this process until one can no longer find a dependence relation amongst the generators of S.

Of course, what's missing in this discussion is exactly how we are to find a dependence relation. Today we'll try to fill that gap.

1. Elementary Operations

LEMMA 4.1. *Suppose $\{v_1, \dots, v_n\}$ be a set of vectors and let $S = span_{\mathbb{F}}(v_1, \dots, v_n)$ be the subspace generated by $v_1, \dots, v_n$.*

(i) If λ is a non-zero element of $\mathbb{F}$, then

$$S = span_{\mathbb{F}}(v_1, \dots, v_{i-1}, \lambda v_i, v_{i+1}, \dots, v_n)$$

(ii) If $1 \leq i < j \leq n$, then

$$S = span_{\mathbb{F}}(v_1, \dots, v_{j-1}, v_j + v_i, v_{j+1}, \dots, v_n)$$

Proof.

(i) Let $S' = span_{\mathbb{F}}(v_1, \dots, v_{i-1}, \lambda v_i, v_{i+1}, \dots, v_n)$. If $v \in S'$ then

$$\begin{aligned} v &= \alpha_1 v_1 + \cdots + \alpha_i (\lambda v_i) + \cdots + \alpha_n v_n \\ &= \alpha_1 v_1 + \cdots + (\alpha_i \lambda) v_i + \cdots + \alpha_n v_n \\ &\in S. \end{aligned}$$

and so each element of S' is an element of S.

On the other hand, let v be an arbitrary element of S. Then

$$\begin{aligned} v &= \alpha_1 v_1 + \cdots + \alpha_i v_i + \cdots + \alpha_n v_n \\ &= \alpha_1 v_1 + \cdots + \frac{\alpha_i}{\lambda} \lambda v_i + \cdots + \alpha_n v_n \qquad \text{(which is valid since } \lambda \neq 0_{\mathbb{F}}\text{)} \\ &= \sigma_1 v_1 + \cdots + \frac{\alpha_i}{\lambda} (\lambda v_i) + \cdots + \alpha_n v_n \in S' \end{aligned}$$

So any element $v \in S$ is also an element of S'.

Since $S' \subseteq S$ and $S \subseteq S'$ implies $S = S'$, (i) is proved.

(ii) Let $S' = span_{\mathbb{F}}(v_1, \ldots, v_{j-1}, v_j + v_i, v_{j+1}, \ldots, v_n)$. if $v \in S'$, then v has the form

$$\begin{aligned} v &= \alpha_1 v_1 + \cdots + \alpha_i v_i + \cdots + \alpha_j (v_i + v_j) + \cdots + \alpha_n v_n \\ &= \alpha_1 v_1 + \cdots + (\alpha_i + \alpha_j) v_i + \cdots + \alpha_j v_j + \cdots + \alpha_n v_n \\ &\in S \end{aligned}$$

and so, since each element of S' is in S,

$$S' \subseteq S.$$

On the other hand, if $v \in S$, then v has the form

$$\begin{aligned} v &= \alpha_1 v_1 + \cdots + \alpha_i v_i + \cdots + \alpha_j v_j + \cdots \alpha_n v_n \\ &= \alpha_1 v_1 + \cdots + (\alpha_i - \alpha_j + \alpha_j) v_i + \cdots + \alpha_j v_j + \cdots + \alpha_n v_n \\ &= \alpha_1 v_1 + \cdots + (\alpha_i - \alpha_j) v_i + \cdots + \alpha_j (v_j + v_i) + \cdots + \alpha_n v_n \\ &\in S' \end{aligned}$$

And so,

$$S' \subseteq S \quad .$$

But $S \subseteq S'$ and $S' \subseteq S$ implies $S = S'$.

□

We have thus discovered two operations that we can perform on a set of generators of a subspace that, which modifying particular generators, do not change the subspace generated. Note that these two operations are essentially the two operations we have been stressing all along: scalar multiplication and vector addition. It is a bit more conventional, however, to utilize as basic operations the following two operations.

Corollary 4.2. *Let $L = \{v_1, \ldots, v_n\}$ be an ordered list of generating vectors for a subspace S of a vector space V over a field $\mathbb{F}$. The following three operations on the list L do not change the subspace generated by the vectors in L.*

(i) *replacing a generating vector v_i with a non-zero scalar multiple of itself: $v_i \to \lambda v_i$*
(ii) *replaciing a generating vector v_j with its sum with a scalar multiple of another generator: $v_j \to v_j + \lambda v_i$*
(iii) *interchanging two vectors: $v_i \longleftrightarrow v_j$*

Proof. That the first operation does not affect the subspace generated was already demonstrated in Lemma 4.1. (ii) is proved by

$$\begin{aligned} span_{\mathbb{F}}(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) &= span(v_1, \ldots, \lambda v_i, \ldots, v_j, \ldots, v_n) \qquad \text{by Lemma 4.1 (i)} \\ &= span(v_1, \ldots, \lambda v_i, \ldots, v_j + \lambda v_i, \ldots, v_n) \qquad \text{by Lemma 4.1 (ii)} \\ &= span(v_1, \ldots, v_i, \ldots, v_j + \lambda v_i, \ldots, v_n) \qquad \text{by Lemma 4.1 (i)} \end{aligned}$$

As for (iii), we can use the commutativity of vector addition to interchange terms in the sum

$$\begin{aligned} span\left(v_1,\ldots,v_i,\ldots,v_j,\ldots,v_n\right) &\ni \alpha_1 v_1+\cdots+\alpha_i v_i+\cdots+\alpha_j v_j+\cdots+\alpha_n v_n \\ &= \alpha_1 v_1+\cdots+\alpha_j v_j+\cdots+\alpha_i v_i+\cdots+\alpha_n v_n \\ &\in span\left(v_1,\ldots,v_j,\ldots,v_i,\ldots,v_n\right) \end{aligned}$$

□

Definition 4.3. *An* ***elementary operation*** *on a list of vectors is an operation of the type (i), (ii), or (iii) as in Corollary 4.2.*

Okay, so here's our plan for simplifying the set vectors used to generate a subspace. We'll look for dependence relations by trying to set up relations of the form

$$v_i+\lambda v_j=\mathbf{0}_V \quad .$$

Each time we can do this we can elminate v_j from the set of generators (or v_i if $\lambda=0$). What follows is a procedure for doing this in such a way that we are systematically and inevitably led to minimal set of generators.

2. Matrices

Recall that the (underlying set of the) vector space $\mathbb{F}^m$ consists of all possible ordered lists of m elements of $\mathbb{F}$. Below we develop some calculational tools for working with subspaces of $\mathbb{F}^n$. Later we will show how these calculational tools can be applied in the more general setting of a vector space V over a field $\mathbb{F}$.

Definition 4.4. *Let $\mathbb{F}$ be a field. An $n\times m$* ***matrix*** *over $\mathbb{F}$ is an ordered list of n elements of $\mathbb{F}^m$.*

Notation 4.5. We will use the notation $Mat_{n,m}(\mathbb{F})$ to denote the set of $n\times m$ matrices with entries in $\mathbb{F}$.

The way I have defined a matrix, an example of a 3×2 matrix over $\mathbb{R}$ might be

$$[[-1,2]],[0,2]],[4,-1]]$$

This is, of course, not the way one is accustomed to viewing matrices. What is much more common is to write each vector in the order list as the row of a rectangular array of real numbers; that is, to say as

$$\begin{bmatrix} -1 & 2 \\ 0 & 2 \\ 4 & -1 \end{bmatrix}$$

with the ordering of the rows from top to bottom following the ordering of the vectors from first to last. I have adopted the former definition (an $n\times m$ matrix is an ordered list of n m-dimensional vectors) as that is more mathematical than the usual *visual* definition (a $n\times m$ matrix is an arrangement of nm numbers into an array with n rows and m columns). But of course the two definitions are equivalent. Moreover, despite the good mathematical intentions of the first defintion, the second presentation (in terms of arrays of numbers) is more helpful in carrying out calculations.

Definition 4.6. *The* ***row space*** *$RowSp(\mathbf{A})$ of an $n\times m$ matrix $\mathbf{A}\in Mat_{n,m}(\mathbb{F})$ is the subspace of $\mathbb{F}^m$ that is generated by the n (row-) vectors of the matrix. The* ***column space*** *$ColSp(\mathbf{A})$ of an $n\times m$ matrix $\mathbf{A}$ is the subspace of $\mathbb{F}^n$ generated by the column vectors of $\mathbf{A}$.*

Proposition 4.7. *Let $[v_1,\ldots,v_n]$ be an $n\times m$ matrix. If an elementary operation (see Corollary 4.2 and Definition 4.3) is applied to this list of vectors, the new list of vectors is matrix that has the same row space. More generally, if $\mathbf{M}$ is a matrix and $\mathbf{M}'$ is a matrix obtained from $\mathbf{M}$ by applying a sequence of elementary row operations to the (row) vectors of $\mathbf{M}$ (and the intermediary matrices). Then*

$$RowSp(\mathbf{M}')=RowSp(\mathbf{M})$$

Proof. This is just Corollary 2 translated into the matrix setting and applied over and over again. □

DEFINITION 4.8. *We say that two matrices* $\mathbf{A}$ *and* $\mathbf{B}$ *are* ***row-equivalent*** *if one is obtainable from another via a sequence of elementary operations.*

Note that if $\mathbf{A}$ and $\mathbf{B}$ are row-equivalent then $RowSp(\mathbf{A}) = RowSp(\mathbf{B})$.

DEFINITION 4.9. $v = [\alpha_1, \dots, \alpha_n]$ *be an element of* $\mathbb{F}$. *The* ***pivot*** *of* v *is the first (following the natural ordering of the list entries)* α_i *that is not equal to* $0_{\mathbb{F}}$. *If* α_i *is the pivot of* $v = [\alpha_1, \dots, \alpha_i, \dots, \alpha_n]$. *then* i *is the* ***pivot position*** *of* v.

DEFINITION 4.10. *A matrix* $\mathbf{A} = [v_1, \dots, v_n]$ *is in* ***row echelon form***, *if the pivot position of* v_i *is less than the pivot position of* v_j *whenever* $i < j$.

PROPOSITION 4.11. *Let* $\mathbf{A}$ *be an* $n \times m$ *matrix. Then there exists a sequence of elementary operations that converts* $\mathbf{A}$ *to a matrix in row echelon form.*

Proof. Let us say that a matrix $\mathbf{A} = [v_1, \dots, v_n]$ is in *semi-row-echelon form* if we have

$$PivPos(v_i) \le PivPos(v_j)$$

whenever $i \le j$. It is always possible to convert a matrix to semi-row-echelon for simply by interchanging setting up a correspondence between row vectors and their pivot positoins

$$\begin{array}{cccc} v_1 & v_2 & \cdots & v_n \\ \updownarrow & \updownarrow & \cdots & \updownarrow \\ PivPos(v_1) & PivPos(v_2) & \cdots & PivPos(v_n) \end{array}$$

and then by interchanging the columns of this correspondence, arranging matters so that the pivot positions listed at the bottom form a non-decreasing sequence of numbers.

We emphasize again that we can alway transform matrix into semi-row-echelon form simply by interchanging rows until we have achieved this.

Let us now say that a matrix $\mathbf{A} = [v_1, \dots, v_n]$ is in row echelon form *up to row* k if it is in semi-row echelon form and if for every $i < k$ we have $PivPos(v_j) < PivPos(v_i)$ for all $1 \le j < i \le k$. Clearly, when $\mathbf{A}$ is an $n \times m$ matrix that is in row echelong form up to row n, it is also in row echelon form. We will show that if $\mathbf{A}$ is a matrix that is in row echelon form up row k that it will always be possible to apply elementary operations to $\mathbf{A}$ and thereby obtain a matrix $\mathbf{A}'$ that is in row echelon form up to row $k+1$. Iterating this process successively, we can eventually convert $\mathbf{A}$ to a matrix that is in row echelon form up to row n; that is a matrix that is in row echelon form.

So let k be the largest integer (between 1 and n) such that $\mathbf{A} = [v_1, \dots, v_n]$ is in row echelon form up to row k If $k = n$ we are done, as $\mathbf{A}$ will already be in row echelon form. Otherwise, there has to be row vector $v_k, v_{k+1}, \dots, v_{k+j}$ for which $PivPos(v_{k+1}) = \cdots = PivPos(v_{k+j})$. Say $i = PivPos(v_k)$. Then these vectors have the form

$$\begin{aligned} v_k &= [0, \dots, 0, \alpha_i, \dots, \alpha_m] \\ v_{k+1} &= [0, \dots, 0, \beta_i, \dots, \beta_m] \\ &\vdots \\ v_{k+j} &= [0, \dots, 0, \gamma_i, \dots, \gamma_m] \end{aligned}$$

with none of the pivots $\alpha_i, \beta_i, \ldots, \gamma_i$ equal to $0_{\mathbb{F}}$. We can then use the elementary operation (ii) to replace the row vectors, $v_{k+1}, \ldots, v_{k+j}$ with

$$\begin{aligned} v_{k+1} \quad &\rightarrow \quad v'_{k+1} = v_{k+1} - \frac{\beta_i}{\alpha_i} v_k = \left[0, \ldots, 0, 0, \beta_{i+1} - \frac{\beta_i}{\alpha_i}\alpha_{i+1}, \ldots, \beta_m - \frac{\beta_i}{\alpha_i}\alpha_m \right] \\ &\quad\ \ \vdots \\ v_{k+j} \quad &\rightarrow \quad v'_{k+j} = v_{k+j} - \frac{\gamma_i}{\alpha_i} v_k = \left[0, \ldots, 0, 0, \gamma_{i+1} - \frac{\gamma_i}{\alpha_i}\alpha_{i+1}, \ldots, \gamma_m - \frac{\gamma_i}{\alpha_i}\alpha_m \right] \end{aligned}$$

Each of these vectors $v'_{k+1}, \ldots, v'_{k+j}$ will have a pivot position that's at least $k + 1. > k$ And so now, if we reorder the vectors $v'_{k+1}, \ldots, v'_{k+j}, v_{k+j+1}, \ldots, v_n$ so that they are in semi-row echelon form,

$$v'_{k+1}, \ldots, v'_{k+j}, v_{k+j+1}, \ldots, v_n \quad \xrightarrow{\text{reordering}} \quad v''_{k+1}, \ldots, v''_n$$

and adjoin this list to tail of the list $v_1, \ldots, v_k$, we will end up with a list of vectors

$$v_1, \ldots, v_k, v''_{k+1}, \ldots, v''_n$$

that is in row echelon form (at least) up to row $k + 1$.

We have now successfully demonstrated that we can use elementary row operations to convert a matrix that is in row echelon from up to row k into a matrix that is in row echelon form up to row $k + 1$. Since we are dealing with only a finite matrices (finite lists of vectors), it is clear the we can apply this the process again and again until we have arrived at a matrix that is in row echelon form up to row n - that is, until we have arrived at a matrix in row echelon form. □

3. Coordinate Vectors and Coefficient Matrices

In the previous section we specialized to vector spaces of the form $\mathbb{F}^m$ and introduced an $n \times m$ matrix over $\mathbb{F}$ as an ordered list of n elements of $\mathbb{F}^m$. We'll now return ot the general setting of a general vector space V over a field $\mathbb{F}$. The first thing to point out is how the calculational tools developed in the last section can be brought to bear on this more general situation.

LEMMA 4.12. *Let $B = [v_1, \ldots, v_m]$ be a basis for a vector space V over a field $\mathbb{F}$, then for each vector $v \in V$ there is a unique $a = [a_1, \ldots, a_m] \in \mathbb{F}^m$ such that*

$$v = a_1 v_1 + \cdots + a_m v_m \quad .$$

Proof. Suppose we had two such expressions for a vector v

$$\begin{aligned} v &= a_1 v_1 + \cdots + a_m v_m & (1) \\ v &= b_1 v_1 + \cdots + b_m v_m & (2) \end{aligned}$$

Subtracting the second equation from the first yields

$$(*) \qquad \mathbf{0}_V = (a_1 - b_1) v_1 + \cdots + (a_m - b_m) v_m \quad .$$

Now if any of the coefficients $(a_i - b_i)$ on the right hand side different from $0_{\mathbb{F}}$ then (*) would furnish us with a dependence relation amongst the basis vectors $v_1, \ldots, v_n$. But, by definition, basis vectors must be linearly independent and so cannot have any such dependence relation. Thus, we conclude that $a_i - b_i = 0_{\mathbb{F}}$ for each i between 1 and m. Thus, the two expansions (1) and (2) must coincide. □

DEFINITION 4.13. *Let $B = [v_1, \ldots, v_m]$ be a basis for a vector space V and let $v \in V$. The **coordinate vector** $\mathbf{v}_B$ **of** v **with respect to** B is the ordered list of coefficients $[a_1, \ldots, a_m]$ corresponding to the expansion of v with respect to the basis B :*

$$v = a_1 v_1 + \cdots + a_m v_m \qquad \Longleftrightarrow \qquad \mathbf{v}_B = [a_1, \ldots, a_m] \in \mathbb{F}^n \quad .$$

Note that if we are given the coordinate vector $\mathbf{v}_B = [a_1, \dots, a_m] \in \mathbb{F}^m$ with respect to a basis $B = [v_1, \dots, v_m]$ of V we can immediately write down the element of V that corresponds to v. (Just read the above relation right to left).

DEFINITION 4.14. *Let $B = [v_1, \dots, v_m]$ be a basis for a vector space V over a field $\mathbb{F}$ and let $[u_1, u_2, \dots, u_n]$ be an ordered list of n vectors in V. To this data we can attach an $n \times m$ matrix $\mathbf{A}$ with entries in $\mathbb{F}$. The entries of the i^{th} row of this matrix ($1 \le i \le n$) are taken to coincide with the entries of coordinate vector of the the i^{th} vector u_i with respect to the basis B.*

THEOREM 4.15. *Let V be an m-dimensional vector space with basis $B = [v_1, \dots, v_m]$ and $\mathbf{A}$ be the coefficient matrix of a set of n non-zero vectors $[u_1, \dots, u_n]$ with respect to B. Suppose that the row vectors $\mathbf{r}_1, \dots, \mathbf{r}_n \in \mathbb{F}^m$ of $\mathbf{A}$ are in row echelon form. Then the vectors $u_1, \dots, u_n$ are linearly independent.*

Proof. We will do a proof by induction, inducing on the number n of vectors in the list $[u_1, \dots, u_n]$. If $n = 1$, that $\{u_1\}$ is linearly independent is obvious, for the only what to hve

$$\alpha_1 u_i = 0$$

is to take $\alpha_i = 0_{\mathbb{F}}$. Assume now that the statement is true when we have a coefficient matrix with $n-1$ rows. We shall show if the coefficient matrix for $[u_1, \dots, u_n]$ is in row echelon form then the only way we can have

$$(***) \qquad \beta_1 u_1 + \beta_2 u_2 + \cdots + \beta_n u_n = \mathbf{0}_V$$

is by taking $\beta_1 = 0_{\mathbb{F}}$, $\beta_2 = 0_{\mathbb{F}}, \dots, \beta_n = 0_{\mathbb{F}}$.

We'll first show that $\beta_1 = 0_{\mathbb{F}}$. Let λ_{1i} be the first non-zero entry of the first row of $\mathbf{A}$. This element corresponds to the coefficient of u_1 with respect to the i^{th} basis vector v_i. Since the matrix $\mathbf{A}$ is in row echelon form, the leading and subsequent entries of the subsequent rows correlate to components of the vectors $u_2, \dots, u_n$ along basis vectors v_j with $j > i$. But then if we expand each u_k , $k = 1, \dots, n$ on the left hand side with respect to the basis B, there is nothing on the left hand side of (***) to cancel the term $\beta_1 \lambda_{1i} v_i$ that arises from $\beta_1 u_1$. So β_1 must vanish.

Therefore, the first term in the expansion (***) must vanish and so we have

$$\beta_2 u_2 + \cdots + \beta_n u_n = \mathbf{0}_V$$

But now, by our induction hypothesis, it must be that $u_2, \dots, u_n$ are linearly independent, so all $\beta_2, \dots, \beta_n = 0_{\mathbb{F}}$ as well. And so each β_i in (***) must separately equal $0_{\mathbb{F}}$ and so the vectors $u_1, \dots, u_n$ must be linearly independent. □

THEOREM 4.16. *Let $\mathbf{A}$ be the coefficient matrix expressing a list $[u_1, \dots, u_n]$ of vectors in V in terms of their coefficients with respect to a basis $B = [v_1, \dots, v_m]$ of V. Then the following statements hold.*

(i) *There exists a matrix $\mathbf{A}'$ row equivalent to $\mathbf{A}$, such that either $\mathbf{A}' = \mathbf{0}$ or there is a uniquely determined positive integer k (between 1 and n) such that the first k rows of $\mathbf{A}'$ are in row echelon form and the remaining rows are all zero.*
(ii) *The vectors $[w_1, \dots, w_k]$ corresponding to the first k rows of $\mathbf{A}'$ form a basis for span $(u_1, \dots, u_n)$.*
(iii) *The original set of vectors are linearly independent if and only if $n = k$.*

Proof. By Proposition 4.10 we can certainly convert $\mathbf{A}$ to a matrix in row echelon form by applying elementary row operations. Because the elementary operations applied to a list of vectors do not affect the subspace generated by the vectors (Corollary 4.2), the span of the vectors $[w_1, \dots, w_n]$ corresponding to the rows of $\mathbf{A}'$ is the same as span of the vectors $[u_1, \dots, u_n]$ corresponding to the rows of $\mathbf{A}$. But in fact, since the 0 rows of $\mathbf{A}'$ correspond to the $\mathbf{0}$-vector in V and so do not contribute to the span of vectors $[w_1, \dots, w_n]$ we have

$$span\,(u_1, \dots, u_n) = span\,(w_1, \dots, w_k)$$

We also know that by the preceding theorem that since the first k rows of $\mathbf{A}'$ are in row echelon form, the corresponding vectors $w_1, \ldots, w_k$ must be linear independent. And so the vectors $w_1, \ldots, w_k$ actually form a basis for $span\,(u_1, \ldots, u_n)$. The uniqueness of k follows from Corollary 3.4 (each basis of a subspace has the same number of vectors). $\square$