APPRENTICE PROGRAM: LINEAR ALGEBRA

INSTRUCTOR: MIKLÓS ABÉRT

Contents

1. Vector Spaces 1
2. Applications to Matrix Theory 5

1. Vector Spaces

Definition 1.1. Let K be a field. $(V, +, \cdot)$ is a *vector space* over K, if $(V, +)$ is an Abelian group, and $\cdot : K \times V \to V$ (called *scalar multiplication*) is distributive, associative and multiplication by 1 is the identity on V. This compact definition unwinds to give us the following ten(!) axioms:

(1) $+ : V \times V \to V$
(2) $u + (v + w) = (u + v) + w$.
(3) $u + v = v + u$.
(4) $\exists 0 \in V$ such that $v + 0 = 0 + v = v$, for all $v \in V$.
(5) $\forall v \in V$, $\exists - v \in V$ such that $v + (-v) = 0$.
(6) $\cdot : K \times V \to V$.
(7) $(\alpha\beta) \cdot v = \alpha \cdot (\beta \cdot v)$.
(8) $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$.
(9) $\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v$.
(10) $1 \cdot v = v$.

Remark 1.2. In Axiom (7) above the product $\alpha\beta$ arises from multiplication in the *field* K. It has nothing to do with the scalar product $\cdot$.

Exercise 1.1. *Show that $0_K \cdot v = 0$, for all $v \in V$. Use this to show that $(-1) \cdot v = -v$, for all v.*

Definition 1.3. A *linear combination* of elements $v_1, \ldots, v_k \in V$ is a sum of the form $\alpha_1 \cdot v_1 + \ldots + \alpha_k \cdot v_k$, where $\alpha_1, \alpha_2, \ldots, \alpha_k \in K$.

A subset $X \subset V$ is *linearly dependent* if there is a *non-trivial* linear combination from X equals 0. By a non-trivial linear combination from X we mean a linear combination $\alpha_1 v_1 + \ldots + \alpha_k v_k$ with $v_i \in X$ and $\alpha_i \in K$, with $k \geq 1$ and $\alpha_i \neq 0$ for at least one i.

A *list* of vectors $v_1, \ldots, v_k$ is linearly dependent if there is a non-trivial linear combination of the v_i, which equals 0. In a list, we're allowed to repeat an element more than once. For example v, v is an admissible list, and this will always be linearly dependent, but the set $\{v, v\} = \{v\}$ is not linearly dependent unless $v = 0$.

Date: 29 July, 2006.

An element $x \in V$ *depends* on $X \subset V$ if x can be expressed as a linear combination of elements of X.

A subset $X \subset V$ is *linearly independent* if it's not linearly dependent.

Exercise 1.2. *Show that any subset consisting of three vectors in a plane is linearly dependent.*

Note on Notation 1. We'll use the word 'iff' to mean 'if and only if'.

Lemma 1.4. *$X \subset V$ is linearly dependent iff there exists $x \in X$ that depends on $X \setminus \{x\}$.*

Proof. ($\Rightarrow$)Suppose X is linearly dependent; then we have

$$\sum_{i=1}^{k} \alpha_i v_i = 0 \ , \exists \alpha_i \in K \ , v_i \in X,$$

for some non-trivial linear combination of elements in X. Since at least one of the α_i is non-zero, we might as well assume $\alpha_1 \neq 0$. Take $y = v_1$; this will work.

($\Leftarrow$)Just reverse the steps above. $\square$

Lemma 1.5. *Suppose $X \subset V$ is a linearly independent subset, and let $y \in V$; then $X \cup \{y\}$ is linearly dependent iff y depends on X.*

Proof. ($\Rightarrow$)If $X \cup \{y\}$ is linearly dependent, then we have

$$\sum_{i=1}^{k} \alpha_i v_i = 0 \ , \exists \alpha_i \in K \ , v_i \in X \cup \{y\},$$

for some non-trivial linear combination of elements in $X \cup \{y\}$. Since X is linearly independent, there is some i such that $v_i = y$ and $\alpha_i \neq 0$ (Why?). This gives us an expression of y as a linear combination of elements in X, which means precisely that y depends on X.

($\Leftarrow$)This is immediate from the definition. $\square$

Definition 1.6. A subset $U \subset V$ is a *vector subspace*, if U is a vector space over K for the operations $+$ and $\cdot$. What we mean is that if $u, v \in U$, and $\alpha \in K$, then $u + v \in U$ and $\alpha \cdot u \in U$: that is, U is closed under addition and multiplication by scalars. To make it absolutely clear that U is a sub*space* and not just a subset we'll say $U \leq V$.

Exercise 1.3. *Give some examples of subspaces of vector spaces.*

Solution. For any vector space V, $\{0\}$ is a subspace, and so is V itself. Any line passing through the origin in $\mathbb{R}^2$ is a subspace of $\mathbb{R}$. The subset $\{(x, y, y) \in \mathbb{R}^3 \mid x, y \in \mathbb{R}\}$ is a subspace of $\mathbb{R}^3$. What does this look like? What are the kinds of subspaces $\mathbb{R}^3$ can have? $\square$

Lemma 1.7. *Suppose $U_i \leq V$, for $i \in I$; then the intersection $\bigcap_i U_i$ is also a subspace of V.*

Definition 1.8. Let $X \subset V$ be any subset. Define $\operatorname{Span} X = \cap_{X \subset U \leq V} U$ to be the *span* of X or the *subspace generated by* X. This is the smallest subspace of V containing X. If $U = \operatorname{Span} X$, we'll say that X *spans* the subspace U.

Remark 1.9. Note that an equivalent way of saying that a vector $y \in V$ depends on a subset $X \subset V$ is to say that $y \in \operatorname{Span} X$.

Exercise 1.4. *Show that the span of the subset $X \subset V$ is the collection of all linear combinations $\alpha_1 v_1 + \ldots + \alpha_k v_k$, for $v_1, \ldots, v_k \in V$. That is*

$$\operatorname{Span} X = \{\sum_{i=1}^{k} \alpha_i v_i \mid v_i \in X,\ \alpha_i \in K\}.$$

Solution. Call the subset on the right hand side U. U is a vector subspace of V (Why?). Now show that every vector subspace of V that contains X must also contain U. ☐

Exercise 1.5. *Consider the vector space of all polynomials over $\mathbb{R}$:*

$$\mathbb{R}[x] = \{\sum_{i=1}^{k} a_i x^i \mid\ a_i \in \mathbb{R}\}.$$

Show that this cannot be generated by finitely many elements.

Solution. Suppose we have finitely many polynomials $f_1(x), \ldots, f_r(x)$, with $\deg f_i = d_i$. Now, consider x^n, where $n > d_i$, for all $1 \leq i \leq r$. Can this be in $\operatorname{Span}\{f_1, \ldots, f_r\}$? ☐

Definition 1.10. A subset $B \subset V$ is a *basis* if $\operatorname{Span} B = V$ and B is linearly independent.

Lemma 1.11. *A subset $B \subset V$ is a basis iff every element of V can be* uniquely *expressed as a linear combination of elements in B.*

Proof. ($\Rightarrow$)First assume that B is a basis. Recalling the constructive description of $\operatorname{Span} B$ from Exercise (1.4), we see that saying $\operatorname{Span} B = V$ is the same as saying that every element can be expressed as a linear combination of B. If some element $v \in V$ can be expressed as a linear combination of elements in B in two different ways, say

$$v = \sum_{i=1}^{k} \alpha_i v_i = \sum_{j=1}^{l} \beta_j w_j,$$

for $v_1, \ldots, v_k, w_1, \ldots, w_l \in B$, then

$$\sum_{i=1}^{k} \alpha_i v_i - \sum_{j=1}^{l} \beta_j w_j = 0$$

gives us a linear combination of elements in B which equals 0. But B was linearly independent to begin with, so all the coefficients in this linear combination must be 0. This shows that the two different expressions that we had for v were in fact the same, and so we indeed have a unique way of expressing every element of V as a linear combination of elements in B.

($\Leftarrow$)Now assume that every element in V can be uniquely expressed as a linear combination of elements in B. Trivially, we have $\operatorname{Span} B = V$. We must show that B is linearly independent. Suppose $\sum_{i=1} \alpha_i v_i = 0$, for some non-trivial linear combination of elements $v_1, \ldots, v_n \in B$. Then, we have two different ways of expressing 0 as a linear combination of elements in B. Contradiction! ☐

Lemma 1.12. *Let $B \subset V$ be a subset. Then the following are equivalent.*

1. *B is a basis.*
2. *B is a maximal linearly independent subset; that is, B is linearly independent, and if $B' \supsetneq B$ is a bigger subset, then B' must be linearly dependent.*
3. *B is a minimal generating set; that is, $\operatorname{Span} B = V$, and if $\widetilde{B} \subsetneq B$ is a smaller subset, then $\operatorname{Span} \widetilde{B} \neq V$.*

Proof. We will show (1) $\Leftrightarrow$ (2) and (1) $\Leftrightarrow$ (3).

[(1) $\Leftrightarrow$ (2)]If B is a basis and $b \in B$; then $B \setminus \{b\}$ cannot span V (Why? Is $b \in \operatorname{Span} B \setminus \{b\}$?). This shows that B is a minimal generating set. Conversely, if B is a minimal generating set, then B must be linearly independent. Otherwise, there is some element $b \in B$, which depends on $B \setminus \{b\}$, but then $\operatorname{Span} B \setminus \{b\} = V$ (Why?), which contradicts the fact that B is a minimal generating set.

[(1) $\Leftrightarrow$ (3)]Suppose B is a basis and $y \in V \setminus B$. Then, since B spans V, y depends on B. But then, by Lemma 1.5, $B \cup \{y\}$ is linearly dependent. This shows that B is a maximal linearly independent subset. Conversely, suppose B is a maximal linearly independent subset. Then, we claim that $\operatorname{Span} B = V$; for otherwise, if we can find $y \in V \setminus \operatorname{Span} B$, then $B \cup \{y\}$ will still be linearly independent, contradicting the maximality of B. $\square$

Theorem 1.13 (Exchange Principle)**.** *If $v_1, v_2, \ldots, v_n$ is a linearly independent list, and $\operatorname{Span}\{w_1, \ldots, w_k\}$ contains all the v_i. Then, $\forall 1 \leq i \leq n$, $\exists 1 \leq j \leq k$ such that $v_1, \ldots, v_{i-1}, w_j, v_{i+1}, \ldots, v_n$ is also a linearly independent list.*

Proof. Consider the list $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n$: this is still linearly independent. Let $U = \operatorname{Span}\{v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n\}$ be the subspace of V generated by this list. If $w_j \in U$ for every $1 \leq j \leq k$, then we have a problem, because v_i will then not be in the span of $w_1, \ldots, w_k$, contradicting our assumption. So there is at least one j such that $w_j \notin U$. This j will work. $\square$

Corollary 1.14. *If $\{v_1, \ldots, v_n\} \subset V$ is a linearly independent subset and $\{w_1, \ldots, w_k\}$ is a generating subset, then $n \leq k$.*

Proof. By the Exchange Principle, we can exchange each of the v_i one by one with some w_j, while still keeping our list linearly independent. Eventually, we can replace each of the v_i with some w_j. If $n > k$, then some w_j has to repeat, but then the list won't be linearly independent. This is a contradiction, and so $n \leq k$. $\square$

Corollary 1.15. *Suppose a vector space V has a finite generating subset. Then V has a basis. Moreover, every basis has the same size, and we denote this common size by* $\dim V$, *the* dimension *of V.*

Proof. Let $X \subset V$ be a finite subset that spans V. If X is linearly independent, then it's already a basis according to Lemma (1.12). If X is linearly dependent, then we can find a $y \in X$ that depends on $X \setminus \{y\}$. So $X \setminus \{y\}$ must also span V. Is $X \setminus \{y\}$ linearly independent? If so, then we're done. If not, rinse and repeat till we hit a linearly independent subset. Note that this *won't* work if X is infinite. Why?

Suppose $\{v_1, \ldots, v_n\}$ and $\{w_1, \ldots, w_k\}$ are two bases for V. Use the previous Corollary to show that $n \geq k$ and $k \geq n$. $\square$

2. Applications to Matrix Theory

Definition 2.1. Suppose we have an $n \times k$ matrix

$$M = \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nk} \end{pmatrix}$$

For $1 \leq i \leq n$, let v_i be the vector $(a_{i1}, \dots, a_{ik}) \in K^k$. The v_i are the *row vectors* of the matrix M. For $1 \leq j \leq n$, let w_j be the vector $(a_{1j}, \dots, a_{nj}) \in K^n$. The w_j are the *column vectors* of the matrix M.

Definition 2.2. The *subspace spanned by the rows* of M is

$$R(M) = \mathrm{Span}\{v_1, \dots, v_n\} \subset K^k,$$

and the *subspace spanned by the columns* of M is

$$C(M) = \mathrm{Span}\{w_1, \dots, w_k\} \subset K^n.$$

Suppose now that M is an $n \times n$ square matrix, and suppose that the list $v_1, \dots, v_n$ is linearly dependent. In this case, we'll say that the rows of the matrix are linearly dependent. Then there is a non-trivial linear combination of the v_i that gives us 0. Since the determinant of M doesn't change under row operations, using this linear dependence relation, we can change M to another matrix M' using just row operations so that M' now has a row with all zeros. But then $\det M' = 0$. Since row operations preserve the determinant, $\det M$ must also be 0.

In fact the converse is also true: if $\det M = 0$, then the rows of M will be linearly dependent. We'll prove a more general theorem soon, but, for now, given any $n \times k$ matrix M, consider the subspace $R(M) \subset K^k$ generated by its rows. It is clear that row operations on M do not change $R(M)$. But what about column operations? For example, consider

$$M = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

It's easy to see that

$$R(M) = \{(x, 0) : x \in K\} \subset K^2.$$

But now suppose we add the first column to the second; then we get the following matrix.

$$M' = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$$

Now,

$$R(M') = \{(x, x) : x \in K\} \subset K^2.$$

So a column operation might change the subspace spanned by the rows, but it won't change the *dimension* of $R(M)$ (for example, we had above $\dim R(M) = \dim R(M') = 1$). This will follows from the next Lemma.

Lemma 2.3. *Any column operation on M cannot decrease the dimension of $R(M)$. Symmetrically, any row operation on M cannot decrease the dimension of $C(M)$.*

Proof. There are three basic kinds of column operations we can perform on M:

(1) We can scale one of the rows by a non-zero scalar. So suppose we've scaled the first column of M by $\alpha \neq 0$ (it doesn't matter which column we pick; the proof's the same). So now M has changed to
$$M' = \begin{pmatrix} \alpha \cdot a_{11} & \dots & a_{1k} \\ \vdots & \ddots & \vdots \\ \alpha \cdot a_{n1} & \dots & a_{nk} \end{pmatrix}.$$
(2) We can exchange one column with another.
(3) We can add one column to another

□

Why does this tell us that the dimension of $R(M)$ is *invariant* under column operations? (Hint: Note that a column operation is invertible).

So we can run any row or column operation we want on M without changing the dimension of $R(M)$ or $C(M)$. What's the simplest form into which we can transform M via row and column operations?

Lemma 2.4. *Let M be an $n \times k$ matrix. Then using row and column operations we can transform M into a matrix that has 0's everywhere off the diagonal, and has a sequence of 1's followed by a sequence of 0's on the diagonal. That is, we can transform M into a matrix of the form*
$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}.$$

Proof. We'll do this by induction on n. The base case is when $n = 1$. In this case M is just a row vector
$$M = (a_1\ a_2 \dots\ a_k).$$
If $M = 0$, then we're done. Otherwise, we can carry a non-zero element to the head of the row, scale it so that it equals 1, and then get rid of the rest of the row by subtracting suitable multiples of 1 from each element. Now, suppose $n > 1$; again, if $M = 0$, then we're done. So we can suppose that there is some $a = a_{ij} \neq 0$. Using row and column exchanges we can carry this a to the top-left corner, and scale the first row by a^{-1} to change it to 1. Then we can subtract a multiple of this from every other element of the matrix that's either on the first row or the first column to clear both the first row and column, and end up with a matrix that looks like this:
$$M' = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & a_{nk} \end{pmatrix}.$$
Now, the submatrix that we get when we remove the first row and column is an $(n-1) \times (k-1)$ matrix. So, by the induction hypothesis, this can be changed to a matrix with 0's off the diagonal and only 1's and 0's on the diagonal using row and column operations, which do not affect the 1 sitting by itself in the corner. □

Definition 2.5. Let M be an $n \times k$ matrix over K. Then

(1) The row rank $\operatorname{rk}_{row} M$ is the number $\dim R(M)$.
(2) The column rank $\operatorname{rk}_{col} M$ is the number $\dim C(M)$.
(3) The determinant rank $\operatorname{rk}_{\det} M$ is the largest number m, for which there exists an $m \times m$ submatrix, whose determinant is non-zero.

Lemma 2.6. *An $n \times n$ matrix M has non-zero determinant iff $\operatorname{rk}_{row} M = n$ iff $\operatorname{rk}_{col} M = n$.*

Proof. Using row and column operations we can get M into the special form as in the Lemma above. Now, $\operatorname{rk}_{row} M$ and $\operatorname{rk}_{col} M$ are both invariant under row and column operations. But in this special form that we have, it's immediate that $\operatorname{rk}_{row} M = \operatorname{rk}_{col} M$: they're both the number of 1's along the diagonal. Moreover, the determinant was non-zero to begin with iff it stays non-zero under all row and column operations. The matrix in this special form, however, has nonzero determinant iff all its diagonal entries are 1's, which can happen iff

$$\operatorname{rk}_{row} M = \operatorname{rk}_{col} M = n.$$

□

Theorem 2.7. *Let M be an $n \times k$ matrix over K. Then*

$$\operatorname{rk}_{row} M = \operatorname{rk}_{col} M = \operatorname{rk}_{\det} M.$$

Proof. We've shown that both $\operatorname{rk}_{row} M$ and $\operatorname{rk}_{col} M$ are invariant under row and column operations. Notice now that if we've converted our matrix M into the special matrix as in the Lemma above, which has only 1's and 0's on the diagonal, and 0's everywhere else, then this special matrix has the *same* row and column rank: they both equal the number of 1's along the diagonal. This shows that for every matrix M

$$\operatorname{rk}_{row} M = \operatorname{rk}_{col} M.$$

Also observe that the determinant rank of the special matrix is the same as its row and column ranks: it's again the number of 1's on the diagonal. So if we manage to show that the determinant rank is also invariant under row and column operations, then our proof will be done. We'll prove that the determinant rank does not decrease under column operations. This will prove invariance under column operations, just as before; the proof for invariance under row operations is strictly analogous.

So suppose $\operatorname{rk}_{\det} M = m$; then there is an $m \times m$ submatrix N of M, such that $\det N \neq 0$, and for $l > m$, we can't find any $l \times l$ submatrix which has non-zero determinant. Since row and column exchanges clearly do not change the determinant rank, we can assume that this submatrix N is in the top left corner. If the column operation doesn't affect any columns of N, or if they only move around columns within N without bringing any outside columns into play, then we're safe.

There are two ways in which there can be trouble: Either some column of N gets exchanged with a column outside of N, or some column outside of N is added to a column in N. In either case, consider the $m \times (m+1)$ submatrix P that's formed by N and the first m elements of the column that is going to be exchanged with (or added to) the poor column in N. Since N has non-zero determinant, we know by the previous Lemma that $\operatorname{rk}_{col} N = m$. This shows that $\operatorname{rk}_{col} P \geq m$; since the subspace $C(P)$ is invariant under column operations on P, even after we've changed P to some other matrix P' using these column operations, we still have $C(P') = C(P)$, and so

$$\mathrm{rk}_{col}\, P' = \dim C(P') = \dim C(P) \geq m.$$

But then P' must also have at least m linearly independent columns (in fact, it'll have exactly m linearly independent columns). Now, pick m linearly independent columns in P' and let N' be the $m \times m$ submatrix formed by these m linearly independent columns. Then, it follows from the Lemma above that $\det N' \neq 0$, which finishes our proof. $\square$