REU 2006 · Apprentice Program

Instructor: Miklos Abert
Scribe: Deepam Patel

July 6, 2006.
NOT FULLY PROOF-READ

1 Determinants and Linear equations.

Let $A = (a_{i,j})$ be an $n \times n$ matrix. Denote by $A_{i,j}$ the $(n-1) \times (n-1)$ matrix derived by deleting the i-th row and j-th column of A. Then we have the following formula for the determinant of A.

Theorem 1.1. *Given* $1 \leq i \leq n$,
$det(A) = \sum_{j=1}^{n} a_{i,j}(-1)^{i+j} det(A_{i,j})$.

Proof. Recall that by definition $det(A) = \sum_{\sigma}(-1)^{sgn(\sigma)} \prod_{i=1}^{n} a_{i,\sigma(i)}$. For a given $a_{i,j}$, group all terms with $a_{i,j}$ together and factor out the $a_{i,j}$. The remaining term will be denoted by $M_{i,j}$. Since every term in the above sum contains an element from the i-th row, for any fixed i, it follows that $det(A) = \sum_{j=1}^{n} a_{i,j} M_{i,j}$. In fact, one can see that $\mid det(A_{i,j}) \mid = \mid M_{i,j} \mid$. The theorem follows from the following exercise. □

Exercise 1.2. *Show that* $det(A_{i,j}) = (-1)^{i+j} M_{i,j}$.

Now consider, for any k, $\sum_{j=1}^{n} a_{k,j}(-1)^{i+j} det(A_{i,j})$. If $k = i$, then the theorem gives us the value of this sum. But, if $k \neq i$ then this sum is always zero. To see this, let B be the matrix with all rows the same as those of A except that we replace the i-th row by the k-th row. Then $det(B) = 0$. On the other hand, the theorem shows that $det(B) = \sum_{j=1}^{n} a_{k,j}(-1)^{i+j} det(A_{i,j}$. If we let $B = (b_{i,j} = (-1)^{i+j} det(A_{i,j}))$, then the above discussion shows that $AB^t = det(A)I$, where I is the identity matrix.

We shall use the the preceding discussion to solve linear equations. For example, consider following linear system of equations:

$$7x + 3y + z = 1 \tag{1}$$

$$-x - y + 7z = 2 \tag{2}$$

$$3x + 3y + 3z = 1 \tag{3}$$

We would like to determine all vectors (x, y, z) simultaneously solving our linear system. The above equation can be rewritten in matrix form as:

$$\begin{pmatrix} 7 & 3 & 1 \\ -1 & -1 & 7 \\ 3 & 3 & 3 \end{pmatrix}$$

Denote this matrix by A. Let $X = (x, y, z)$ and $Z = (1, 2, 1)$. Thinking of X and Z as column vectors the above system of equations can be represented as the matrix equation $AX = Z$. In general, we are given a $(n \times n)$ matrix A, a vector of n-variables X and a n-vector Z. We would like to determine when the the matrix equation $AX = Z$ has solutions. Suppose $Z = 0$. Suppose our variables are denoted by x_i for $1 \leq i \leq n$. Then $x_i = 0$ is always a solution. Furthermore, the set of solutions forms a vector space. If $det(A)$ is not 0 then 0 is the unique solution of this system. Now suppose $Z \neq 0$. One has a unique solution if and only if $det(A) \neq 0$. If $det(A) = 0$ then one needs Z to be in the span of the columns of A.

2 Linear transformations

Definition 2.1. *Let U, W be subspaces of V. Then $U + W = \{u + w | u \in U, w \in W\}$.*

Note that $U + W$ is a subspace of V and $U + W = Span < U \cup V >$.

Definition 2.2. *We write $V = U \oplus W$ if $V = U + W$ and $U \cap W = \{0\}$.*

Lemma 2.3. *$V = U \oplus W \Leftrightarrow$ all elements of $v \in V$ can be uniquely written in the form $u + w$ for $u \in U$ and $w \in W$.*

Proof. $\Rightarrow$ Since $V = U \oplus W$ there exist $u \in U$ and $w \in W$ such that $v = u + w$. Suppose this representation is not unique. Then there exist $u_1 \in U$ and $w_1 \in W$ such that $u_1 + w_1 = v = u + v$. Then $u_1 - u = w_1 - w = 0$ since $U \cap W = \{0\}$.

$\Leftarrow$ Exercise $\square$

Lemma 2.4. *$V = U \oplus W \Rightarrow dim(V) = dim(U) + dim(W)$.*

Proof. Let $B_U = \{u_1, \ldots, u_m\}$ and $B_W = \{w_1, \ldots, w_p\}$ be a basis for U and W respectively. It is enough to show that $B_U \cup B_W$ is a basis for V. It is clear that $B_U \cup B_W$ spans V. Suppose the elements of $B_U \cup B_W$ are not linearly independent. Then one has a relation $\sum_{i=1}^{m} a_i u_i + \sum_{j=1}^{p} b_j w_j = 0$. Then $W \cap U = \{0\}$ implies $\sum_{i=1}^{m} a_i u_i = \{0\}$ and $\sum_{j=1}^{p} b_j w_j = 0$. But, then all the a_i's and b_j's are 0. $\square$

Definition 2.5. *A linear transformation $A : V \to W$ is a map such that $A(v_1 + v_2) = A(v_1) + A(v_2)$ and $A(\lambda v) = \lambda A(v)$. Here v_i and v are vectors in V and λ is a scalar in the base field K.*

We denote the set of linear maps from V to W by $Hom_K(V, W)$.

Definition 2.6. *The kernel of a linear transformation A is $Ker(A) = \{v \in V, A(v) = 0\}$. The image of A is $Im(A) = \{A(v) \in W\}$.*

Note that $Ker(V)$ is a subspace of V and $Im(A)$ is a subspace of W.

Lemma 2.7. *$Span(X) = V \Rightarrow Span(< A(X) >) = Im(A)$*

Proof. Since A is a linear transformation $Span(< A(X) >) = A(Span(X))$. But, $A(Span(X)) = A(V) = Im(A)$. $\square$

Lemma 2.8. *Let B be a basis of V. Let $f : B \to W$ be any map. Then f extends uniquely to a linear map $f : V \to W$.*

Proof. Suppose $B = \{v_1, \ldots, v_n\}$. Let $v \in V$. Then $v = \sum_{i=1}^{n} a_i v_i$. Define $f(v) = \sum_{i=1}^{n} a_i f(v_i)$. $\square$

Lemma 2.9. *The following are equivalent.*

1: A is injective.

2: $Ker(A) = 0$

3: all independent subsets of V have independent images.

4: There exists a basis B of V such that A is injective restricted to B and $A(B)$ is linearly independent.

Proof. $1 \Rightarrow 2$: If $v \in Ker(A)$ then $A(v) = 0 = A(0)$. So if A is injective then $v = 0$.

$2 \Rightarrow 3$: Let X be a linearly independent subset of V. Suppose $Im(X)$ is linearly dependent. Then there exist $x_1, \ldots, x_p$ in X such that $\sum_{i=1}^{p} a_i A(x_i) = 0$. Then $A(\sum_{i=1}^{p} a_i x_i) = 0$. But, then $\sum_{i=1}^{p} a_i x_i = 0$, a contradiction.

$3 \Rightarrow 4$: Let $w_1, \ldots w_n$ be a basis of $Im(A)$. Choose $v_i \in V$ such that $A(v_i) = w_i$. Then v_i are linearly independent. Suppose v is not in the span of the v_i. Then $\{v, v_1, \ldots, v_n\}$ is a linearly independent set and therefore so is its image. But, this is a contradiction since $A(v)$ is in the span of the w_i. Therefore the v_i's give the required basis.

$4 \Rightarrow 1$: Clear $\square$

Theorem 2.10. *$V \backsimeq W \Leftrightarrow dim(V) = dim(W)$.*

Proof. $\Rightarrow$ Let B be a basis of V. Then surjectivity implies $Im(B)$ spans W and injectivity (previous Lemma) implies that $Im(B)$ is linearly independent. Hence it's a basis of W.

$\Leftarrow$ Choose a basis $B = \{v_1, \ldots, v_n\}$ of V and a basis $C = \{w_1, \ldots, w_n\}$ of W. Define A to be the linear transformation which send v_i to w_i.

$\square$

Note that the above theorem implies that any n-dimensional vector space V is isomorphic to K^n. Thus, to study linear transformations between an n-dimensional vector space and a k-dimensional vector space one can study the set $Hom_K(K^n, K^m)$. An element of this set is determined by its values on the standard basis vectors. In particular, it can be represented by a $m \times n$ matrix. The $j-$th column consists of the coefficients appearing in the representation of the image of the j-th basis vector under the given linear map as a linear combination of the standard basis vectors in K^m. This interpretation can be used to show matrix associativity since composition of linear maps is associative.

Theorem 2.11. *$dim(V) = dim(Ker(A) + dim(Im(A))$.*

Proof. Let B be a basis for the kernel. Let $C = \{c_1, \dots, c_n\}$ be a basis for the image. Let w_i be an inverse image of c_i. Then $B \cup \{w_1, \dots, w_n\}$ spans V.

$\square$

3 Derivative

Let f be a differentiable function $\mathbb{R} \to \mathbb{R}$. Then the derivative of $f(x)$ denoted $f'(x)$ is the $\lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$. It follows that $f(x+h) - f(x) = f'(x)h + r(h)$ where $r(h)$ is a small remainder term. In particular, we can think of the derivative of f at x as a linear transformation $\mathbb{R} \to \mathbb{R}$ which sends h to $f'(x)h$. This linear transformation approximates f near x. This allows us to define the derivative in general.

Consider the complex conjugation map from $\mathbb{C} \to \mathbb{C}$. This map is just a linear map when considered as a map of $\mathbb{R}^2$. Thus, it is differentiable as a function on $\mathbb{R}^2$. But, it is not complex differentiable. This is easily seen by computing the limits in the definition of the derivative by coming in from the imaginary axis and the real axis (the two limits are different).

Exercise 3.1. *Describe a condition on 2×2 matrices giving rise to complex differentiable maps.*