UChicago WOMP 2007 : Algebra I : Linear Algebra

September 10, 2007

Welcome to The University of Chicago! To get things started, let's do some linear algebra. The goal of this program is to help you hit your math groove before the semester starts. These pages will prove a few results, state some formulas, and (re)expose you to the sort of manipulations common to the practice of linear algebra in mathematical life. This will probably not be useful as a learning tool for first timers. (I assume that all of you have taken at least one course devoted to linear algebra.) Please help me be more clear by informing me of bad or mysterious notation (on the board and in my notes), confusing omissions, etc. The omitted proof of any statement should be understood to be an exercise.

Dimension Formulae

Let k be a field. We will discuss the theory of vector spaces over k. Every vector space over k is isomorphic to a vector space of the form $\bigoplus_I k$ with the natural vector space structure (I can have any cardinality). Let V be a vector space. V is *finite dimensional* if V is isomorphic to k^n for some positive integer n. The *dimension* $\dim(V)$ of V is this number n. Recall that $\dim(V)$ is both the size of a minimal spanning set and of a maximal linearly independent set. **For the remainder of these notes, assume that all explicitly mentioned vector spaces are finite dimensional.**

There are two important formulas that describe the behavior of the dimension.

Formula 1. *Let U be a subspace of a vector space V. Then* $\dim(U) \leq \dim(V)$.

Formula 2. *Let U and W be subspaces of a vector space V. Then*

$$\dim(U+W) = \dim(U) + \dim(W) - \dim(U \cap W).$$

Formula 3. ***(Rank-Nullity.)*** *Let $T : V \to W$ be a linear transformation with V, W vector spaces. Then*

$$\dim(\operatorname{im} T) + \dim(\ker T) = \dim(V).$$

All of these formulae can be verified using bases. We can immediately draw some conclusions. If $U, W \subset V$ and $\dim(U) + \dim(W) > \dim(V)$ then $\dim(U \cap W) > 0$ and thus $U \cap V \neq 0$. Furthermore if T is either injective or surjective and $\dim(V) = \dim(W)$ then T is an isomorphism.

Proposition 1. *Let $f(x) \in k[x]$ be a nonzero polynomial. Then there exists $g(x) \in k[x]$ such that $f(x)g(x) = \sum_{p\ prime} \alpha_p x^p$ for some $\alpha_p \in k$.*

Proof. Let $V_n = \{h(x) \in k[x] | \deg(h) \leq n\}$ and let

$$P_n = \{ \sum_{p \text{ prime}, p \leq n} \alpha_p x^p | \alpha_p \in k \}.$$

Multiplication by f is a linear tranformation $\mu : V_n \to V_{n+\deg(f)}$ which is injective. Therefore $\mathrm{im}(\mu) = \dim(V_n)$ and $\dim(V_n) = n+1$ because $\{1, x, \ldots, x^n\}$ is a basis for V_n. Since there are infinitely many primes, $\dim(P_n) \to \infty$ with n. Let N be so large that $\dim(P_{N'}) > \deg(f)$, where $N' = N + \deg(f)$. Then, viewing $\mathrm{im}(\mu)$ as a subspace of $V_{N'}$ we have

$$\dim(\mathrm{im}(\mu)) + \dim(P_{N'}) > N + 1 + \deg(f) = \dim(V_{N'}).$$

Hence $\mathrm{im}(\mu) \cap P_{N'} \neq 0$ so there is a $g(x) \in k[x]$ such that $f(x)g(x) \in P_{N'}$ as desired. □

Dual spaces, trace, and the transpose

If V, W are vector spaces denote by $\mathrm{Hom}(V, W)$ the vector space of linear transformations $V \to W$. Set $V^* = \mathrm{Hom}(V, k)$; this is the *dual space* to V. There is a function $V^* \times W \to \mathrm{Hom}(V, W)$ defined by

$$T_{(\phi,w)}(v) = \phi(v)w.$$

The image of this map spans $\mathrm{Hom}(V, W)$. In fact, given bases $\{\phi_i\}, \{w_j\}$ for $V*$ and W, respectively, $\{T_{(\phi_i,w_j)}\}$ is a basis for $\mathrm{Hom}(V, W)$.

Let $\langle , \rangle : V^* \times V \to k$ be defined by

$$\langle \phi, v \rangle = \phi(v).$$

Given a basis $\{v_i\}$ for V there is a dual basis $\{v_i^*\}$ satisfying $\langle v_i^*, v_j \rangle = \delta_{ij}$. (Now we know that $\dim(V) = \dim(V^*)$.)

Fix bases $\{v_i\}, \{w_j\}$ for V and W and consider $T = \sum \alpha_{ij} T_{(v_i^*,w_j)} \in \mathrm{Hom}(V, W)$. Plugging in basis vectors v_i we see that $T(v_i) = \sum_j \alpha_{ij} w_j$ so that the coefficients α_{ij} are just the coefficients of the matrix representing T with respect to these bases.

The trace and transpose maps can be described in a basis free way. Define $\mathrm{Tr}(T_{(\phi,v)}) = \phi(v)$ for any $\phi \in V^*$ and $v \in V$. We extend this linearly to obtain a map $\mathrm{Tr} : \mathrm{Hom}(V, V) \to k$. Fixing a basis $\{v_i\}$ for V we calculate

$$\mathrm{Tr}(\sum \alpha_{ij} T_{(v_i^*,v_j)}) = \sum \alpha_{ij} v_i^*(v_j) = \sum_i \alpha_{ii}$$

and verify that this does indeed give another description of the familiar trace function.

The transpose of a linear transformation $T \in \mathrm{Hom}(V, W)$ exists naturally in $\mathrm{Hom}(W^*, V^*)$. It is denoted ${}^tT$ and is defined by the formula

$$\langle {}^tT(w^*), v \rangle = \langle w^*, T(v) \rangle.$$

The form $\langle , \rangle$ supplies an isomorphism $V \to V^{**}$, just by sending v to the map $\phi \mapsto \langle \phi, v \rangle$. So with the usual bases, let us represent ${}^tT$ when $T = \sum \alpha_{ij} T_{(v_i^*,w_j)}$. Write ${}^tT = \sum \beta_{ij} T_{(w_i,v_j^*)}$ and calculate

$$\beta_{ij} = \langle \sum_l \beta_{il} v_l^*, v_j \rangle = \langle {}^tT(w_i^*), v_j \rangle = \langle w_i^*, T(v_j) \rangle = \langle w_i^*, \sum_l \alpha_{jl} w_l \rangle = \alpha_{ji}.$$

Bilinear forms

A function $(,) : V \times W \to k$ is *bilinear* if for all $v \in V$ and $w \in W$ the functions $(v, -) : W \to k$ and $(-, w) : V \to k$ are linear. Let's think about bilinear forms $(,) : V \times V \to k$. Such a form is *symmetric* if for all $u, v \in V$, $(u, v) = (v, u)$ and is called *alternating* if for all $v \in V$, $(v, v) = 0$. An alternating form satisfies $(u, v) = -(v, u)$, and this is how to think of one unless you happen to be working in characteristic 2.

A bilinear form is *nondegenerate* if the only vector $v \in V$ such that $(v, w) = 0$ for all $w \in V$ is $v = 0$. Bilinear forms are equivalent to linear transformations $\beta : V \to V^*$, the equivalence given simply by the relationship $\beta(v)(u) = (v, u)$. Given a bilinear form, let β_v be the functional $\beta_v(u) = (v, u)$. In this light, nondegenerate forms are equivalent to isomorphisms $\beta : V \to V^*$. Given a linear transformation $\beta : V \to V^*$ consider $W = \ker(\beta)$ and $U^* = \mathrm{im}(\beta)$. Let $p : V^* \to U^*$ be a projection and let $i : U \to V$ be the transpose (it is an injection since p is a surjection). Then $V = U \oplus W$ and $\beta|_U : U \to U^*$ is nondegenerate. Hence we can decompose a given bilinear form into a nondegenerate form plus the zero form.

Given a basis $\{v_i\}$ for V there is a matrix representation for $(,)$ given by $M = (m_{ij})$ where $m_{ij} = (v_i, v_j)$. If $u = \sum \alpha_i v_i$ and $v = \sum \beta_i v_i$ then

$$(u, v) = \sum \alpha_i m_{ij} \beta_j.$$

There are two very nice results about the matrix representations of nondenerate symmetric or alternating forms.

Proposition 2. *Let $(,)$ be a nondegenerate alternating form on V. Then $\dim(V) = 2n$ and there is a basis (called a* canonical basis*) such that the matrix representing $(,)$ in this basis is*

$$\begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$$

where I is the $n \times n$ identity matrix.

Proof. Let u and v be such that $(u, v) \neq 0$. u and v are linearly independent since $(,)$ vanishes along lines. Then replace v with $v/(u, v)$ to assume that $(u, v) = 1$. Let $W = \ker(\beta_v) \cap \ker(\beta_u)$. Let $w \in W$ and say $(w, x) \neq 0$ for some x. Let $x' = x + (v, x)u - (u, x)v$ and observe that $(w, x') = (w, x) \neq 0$ and $x' \in W$. Hence, $(,)$ restricts to a nondegenerate form on W. By induction W has even dimension and a basis of the required form. So $\dim(V) = \dim(W) + 2$ and adding u, v to the canonical basis for W gives a canonical basis for V up to permutation. □

Proposition 3. ***(Sylvester's Law.)*** *Let $(,)$ be a nondegenerate symmetric form on a real vector space V. Then there are nonnegative integers $n + m = \dim(V)$ and a basis for V such that the matrix of $(,)$ with respect to this basis is*

$$\begin{pmatrix} I_n & 0 \\ 0 & -I_m \end{pmatrix}$$

where I_n and I_m are the $n \times n$ and $m \times m$ identity matrices respectively.

Proof. Observe first that the restriction of $(,)$ to any subspace is nondegenerate. Suppose that there is a vector $v \in V$ such that $(v, v) \neq 0$. Then we may replace v with $v/\sqrt{|(v, v)|}$ and assume $(v, v) = \pm 1$. Let $W = \ker(\beta_v)$. Then $(,)$ restricted to W is nondegenerate so W has a basis of the desired form and adding v to it will give (up to permutation) the desired basis for V. So we just need to find this v.

Suppose that $(u, v) \neq 0$ then $(u + v, u + v) = (u, u) + (v, v) + 2(u, v)$. If $(u + v, u + v) = 0$ then either (u, u) or (v, v) is nonzero and we can proceed as above. If $(u + v, u + v) \neq 0$ then $u + v \neq 0$ will work. □

The difference $n - m$ is called the *signature* of $(,)$. Perhaps surprisingly, the signature has applications to the topology of manifolds! Note that we only used that assumption that $k = \mathbb{R}$ to take a square root (and so that $2(u, v) \neq 0$ if $(u, v) \neq 0$). The same argument proves that if $\mathrm{char}(k) \neq 2$, then a symmetric bilinear form can be represented by a diagonal matrix. Over $\mathbb{C}$, we can take a square root of any nonzero number so a nondegenerate, symmetric, bilinear form can always be represented by the identity matrix.

The determinant

There is a function $\det : \mathrm{Hom}(V,V) \to k$ called *the determinant* which has numerous wonderful properties. Given a matrix M representing the linear transformation $T \in \mathrm{Hom}(V,V)$, $\det(T)$ is a homogenous polynomial of degree $n = \dim(V)$ in the entries of M,

$$\det(T) = \det(M) := \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \prod_{i=1}^{n} m_{i\sigma(i)}.$$

From this formula it is evident that, over $\mathbb{R}$ or $\mathbb{C}$, the determinant is a smooth function on $\mathrm{Hom}(V,V)$. There is more. If $\det(T) \neq 0$ then T is invertible and T^{-1} is represented by the matrix $\det(T)^{-1}(\mathrm{adj}(j|i))_{ij}$ where $\mathrm{adj}(j|i)$ is the determinant of the *adjugate* matrix, obtained from M by deleting row j and column i. Therefore, (again over $\mathbb{R}$ or $\mathbb{C}$) inversion is a smooth operation on invertible linear transformations $V \to V$. Anyway, one of the tenets of linear algebra is that a transformation is invertible if and only if its determinant is nonzero.

The formula above also makes it easy to show that if $A = (A_{ij})$ is a block diagonal matrix, with A_{ij} an $n \times n$ matrix, then

$$\det(A) = \det((\det(A_{ij}))_{ij}).$$

In particular if $A = \left(\begin{smallmatrix} B & 0 \\ 0 & B \end{smallmatrix}\right)$ is a real matrix then $\det(A) \geq 0$.

Jordan canonical form

Suppose now that k is algebraically closed. Let $T \in \mathrm{Hom}(V,V)$. Associated to T is a finest decomposition into invariant subspaces $V = V_1 \oplus \cdots \oplus V_m$ where $T(V_i) \subset V_i$. The simplest situation is when the restriction of T_i to V_i is just multiplication by a scalar. Not every transformation can be so diagonalized. However it is always the case that $T = D + N$ where D is diagonal and $N^{\max \dim(V_i)} = 0$. Explicitly, T has exactly one (up to permuting the blocks) matrix representative of the form

$$\begin{pmatrix} J_{\lambda_1,d_1} & 0 & \cdots & 0 \\ 0 & J_{\lambda_2,d_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{\lambda_r,d_r} \end{pmatrix}$$

where J_{λ_i,d_i} is the $d_i \times d_i$ matrix, called a *Jordan block*,

$$\begin{pmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 \\ 0 & 0 & \lambda_i & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i \end{pmatrix}.$$

Both D and N can be expressed as polynomials in T (with coefficients in k).

Here is one application of Jordan canonical form. Let P be the matrix representing the permutation $\sigma(i) = s - i$. Then $P = P^{-1}$ and

$$PJ_{\lambda,s}P = {}^tJ_{\lambda,s}.$$

So if M is a matrix and AMA^{-1} is in Jordan form, then there is a matrix $Q = Q^{-1}$ such that $QAMA^{-1}Q = {}^tA^{-1}{}^tM{}^tA$. Rearranging this, ${}^tM = SMS^{-1}$ where $S = {}^tAQA$ so that M and ${}^tM$ are similar.