Apprentice Linear Algebra, 3rd day, 07/06/05

REU 2005

Instructor: László Babai
Scribe: Mohammed Abouzaid

Notes revised: July 8, 2005

3 Linear Maps

Definition 3.1. Given two vector spaces V and W we say that a function $f : V \to W$ is *linear* if:

$$(\forall \mathbf{a}, \mathbf{b} \in V)(f(\mathbf{a} + \mathbf{b}) = f(\mathbf{a}) + f(\mathbf{b}))$$

$$(\forall \mathbf{a} \in V)(\forall \lambda \in \mathbb{R})(f(\lambda \mathbf{a}) = \lambda f(\mathbf{a})).$$

Equivalently we may combine these two conditions into one:

$$(\forall \mathbf{a}_1, \dots, \mathbf{a}_n \in V)(\forall \lambda_1, \dots, \lambda_n \in \mathbb{R}) \left(f\left(\sum_{i=1}^{n} \lambda_i \mathbf{a}_i \right) = \sum_{i=1}^{n} \lambda_i f(\mathbf{a}_i) \right)$$

Such a function is called a *linear map* or a *homomorphism.* If $V = W$, we may also call f a *linear transformation* or a *linear operator.* As an immediate Corollary, we obtain:

Corollary 3.2.

$$f(\mathbf{0}) = \mathbf{0}$$

Proof:

$$f(\mathbf{0}) = f(0 \cdot \mathbf{0}) = 0 \cdot f(\mathbf{0}) = f(\mathbf{0}).$$

Regarding the entire discussion of vector spaces and their properties, we may replace $\mathbb{R}$ by any field F, and define vector space "over F" for which F is the domain of scalars. Note that it only makes sense to consider linear maps between vector spaces over the same field.

We are already familiar with the field of real numbers $\mathbb{R}$, but also with rational numbers $\mathbb{Q}$, complex numbers $\mathbb{C}$, and the finite fields $\mathbb{F}_p$ (integers $\mod p$) for primes p.

Example 3.3. $\mathbb{F}_p^n$ is a vector space over $\mathbb{F}_p$ with p^n elements.

We will now consider some examples of linear maps.

Example 3.4. If $F[x]$ is the space of polynomials over F (considered as a vector space over F) the differentiation operator:

$$D = \frac{d}{dx} : f \mapsto f'$$

is a linear map (Check this).

3.1 Geometric Transformations

Geometric transformations also provide examples of linear maps. Since a linear map must fix the origin, we will only consider geometric transformations with this property:

Example 3.5. Isometries of the plane (congruences) which fix the origin are linear maps. Examples include reflection across an axis that passes through the origin and rotation by a prescribed angle θ about the origin.

Exercise 3.6. Prove that every congruence of the plane fixing the origin is either a rotation or a reflection.

Note that since rotations preserve orientation, while reflections reverse it, these two possibilities are mutually exclusive. We will call transformations which preserve orientation *sense preserving,* while those that do not preserve orientation will be called *sense reversing.* Use the following Lemma to give an elegant solution to the previous exercise.

Lemma 3.7. *If r_1 and r_2 are reflections about two axes that are angle θ apart, then $r_1 \circ r_2$ is a rotation by 2θ.*

We would like to extend these results to congruence in 3-dimensional space. Note that we have rotations about axes, and reflections across planes. (Note: Reflection through a line corresponds to 180 degree rotation.) Note that these reflections fix all points of a plane, while the rotations fix all points of a line. We have the following Lemma which extends the previous result from dimension 2 to 3:

Lemma 3.8. *The composition of reflections in two planes which are at an angle θ is the rotation by 2θ about their line of intersection.*

Unlike the 2-dimensional case, these transformations do not exhaust the isometries of space. Indeed, consider the transformation

$$\begin{aligned} C : \mathbb{R}^3 &\rightarrow \mathbb{R}^3 \\ x &\mapsto -x \end{aligned}$$

which correspond geometrically to reflection through the origin. Since the only fixed point of this transformation is the origin, it cannot be the same as one of the above examples. It is clear geometrically that this transformation is sense reversing. In fact, we can obtain is as the composition of 3 reflection through mutually orthogonal planes. More generally:

Exercise 3.9. (a) Every sense reversing congruence of the 3-dimensional space that fixes the origin is the composition of 3 reflections in planes that contain the origin. (b) Every sense reversing congruence of the 3-dimensional space is the composition of 3 reflections. (Note that no fixed points are assumed.)

You should use the following Theorem in order to prove part (a) of the the above exercise:

Theorem 3.10. *Every sense-preserving congruence of* $\mathbb{R}^3$ *which fixes the origin is a rotation.*

We will see in a future class how this Theorem follows from the fact that every real polynomial of degree 3 has a real root. In fact, we have the more general result:

Lemma 3.11. *Every odd degree real polynomial has a real root.*

Proof: Use the continuity of polynomials, and the fact that the limit in one direction is $+\infty$ and $-\infty$ in the other.

Definition 3.12. A *rotational reflection* is the composition of a rotation and the central reflection C.

The the central reflection is given by multiplication by -1, it commutes with every linear transformation, hence the order of composition in the previous definition is not an issue. An alternative definition goes as follows:

Consider a plane P and a line L perpendicular to it. A *rotational reflection* is the composition of a rotation about the line with a reflection through the plane.

To see that these two notions are the same, we express the central reflection as the composition of reflections through mutually orthogonal planes:

$$C = r_1 \circ r_2 \circ r_3.$$

We may choose r_3 to be the reflection through P, while r_2 and r_1 are reflection through planes P_2 and P_1 whose intersection is L. In particular, using Lemma 3.8, we may express the rotation by angle θ about the line L as the composition

$$r_0 \circ r_1$$

where r_0 is reflection through some plane P_0 which passes through L and such that the angle between P_0 and P_1 is $\frac{\theta}{2}$. Our original definition of a rotational reflection is a transformation of the form:

$$(r_0 \circ r_1) \circ C = r_0 \circ r_1 \circ r_1 \circ r_2 \circ r_3 = r_0 \circ r_2 \circ r_3.$$

Note that $r_0 \circ r_2$ is a rotation (about the line L) with angle $\pi + \theta$ so the expression

$$(r_0 \circ r_2) \circ r_3$$

is exactly that of a rotational reflection according to our second definition. We can now go through the entire argument "backward" by inserting $r_1 \circ r_1$ in order to conclude that the two notions are indeed equivalent.

Note that the trick in the above argument is two choose the expression for a rotation as a product of reflection "appropriately."

Exercise 3.13. (Use Theorem 3.10) All sense-reversing congruences of $\mathbb{R}^3$ which fix the origin are rotational reflections.

Note that rotational reflections include reflections in a plane when the rotation angle is 0.

Examples 3.14. We list further geometric examples of linear transformations.

- Orthogonal projection onto a line.
- Skew projection onto a line. (This means choosing a direction along which the projection occurs; this direction is not necessarily perpendicular to the chosen line.)
- Shearing. In coordinates, this is given by the transformation

  $$(x, y) \mapsto (x + \alpha y, y),$$

  where α is a given real number. This transformation fixes the x-axis and moves a point in the x-direction by an amount proportional to its y coordinate (like tilting a deck of cards).
- Vertical stretching / shrinking, which is given in coordinates by:

  $$(x, y) \mapsto (x, \alpha y)$$

  for some real number α. This transformation will map circles to ellipses and axis-parallel squares to rectangles. One may also compose vertical stretching/shrinking with horizontal stretching/shrinking.

We will now return to an abstract point of view:

3.2 Linear Maps: Rank, Kernel

Theorem 3.15. *If $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is a basis of V and $\mathbf{w}_1, \ldots, \mathbf{w}_n$ are arbitrary elements of W then there exists a unique linear map $f : V \to W$ such that $f(\mathbf{e}_i) = \mathbf{w}_i$.*

We emphasise the fact that the choice of $\mathbf{w}_1, \ldots, \mathbf{w}_n$ is arbitrary, and that V and W do not necessarily have the same dimension.

Proof: We first prove uniqueness:

Given $\mathbf{v} \in V$ we may write $\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{e}_i$ for a unique choice α_i of scalars. If f is a linear function satisfying the conditions of the theorem, we can compute that:

$$f(\mathbf{v}) = f\left(\sum_{i=1}^{n} \alpha_i \mathbf{e}_i\right) = \sum_{i=1}^{n} \alpha_i f(\mathbf{e}_i) = \sum_{i=1}^{n} \alpha_i \mathbf{w}_i,$$

which is indeed a uniquely determined value.

In order to prove existence, we use our result for uniqueness in order to define:

$$f(\mathbf{v}) := \sum_{i=1}^{n} \alpha_i \mathbf{w}_i,$$

where the scalars α_i are uniquely determined by the equation

$$\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{e}_i$$

Exercise 3.16. Complete the proof of existence (i.e: Check the above map f is linear).

Definition 3.17. The *image of f* is the subset of W:

$$\mathrm{Im}(f) = \{f(\mathbf{v}) : \mathbf{v} \in V\}.$$

Exercise 3.18.

$$\mathrm{Im}(f) \leq W$$

Definition 3.19. The *rank of f* is

$$\mathrm{rk}(f) = \dim(\mathrm{Im}(f)).$$

Exercise 3.20.

$$\mathrm{rk}(f) \leq \dim V$$

Proof: Since

$$\mathrm{Im}(f) = \mathrm{Span}\{f(\mathbf{e}_i) \,|\, i = 1, \ldots, n\},$$

we know that $\mathrm{Im}(f)$ has a spanning set which consists of n elements. Therefore, its dimension is less than or equal to n by Magic #1.

Definition 3.21. The *kernel of f* is:

$$\ker(f) = f^{-1}(\mathbf{0}).$$

Exercise 3.22.

$$\ker(f) \leq V$$

Example 3.23. The kernel of the differentiation map on $F[x]$ consists of the subspace of constant polynomials.

One of the main theorems of this theory is:

Theorem 3.24.

$$\dim \ker(f) + \mathrm{rk}(f) = \dim V$$

Proof: Let k be the dimension of $\ker(f)$, and choose $\mathbf{e}_1, \ldots, \mathbf{e}_k$ a basis of $\ker(f)$. Note that these vectors are linearly independent in V (though they may not span it), so we may extend them to a basis (i.e: choose $\mathbf{e}_{k+1}, \ldots, \mathbf{e}_n$ vectors in V such that $\mathbf{e}_1, \ldots, \mathbf{e}_k, \mathbf{e}_{k+1}, \ldots, \mathbf{e}_n$ is a basis of V).

Claim 3.25. $f(\mathbf{e}_{k+1}), \ldots, f(\mathbf{e}_n)$ *is a basis of* $\mathrm{Im}(f)$.

Note that an immediate consequence of this claim is the Corollary

Corollary 3.26. $\dim(\mathrm{Im}(f)) = n - k$

from which the theorem immediately follows. So it suffices to prove the claim. This is itself done in two parts:

1. The vectors span $\mathrm{Im}(f)$: We know that $\mathrm{Im}(f)$ is spanned by $\{f(\mathbf{e}_1), \ldots, f(\mathbf{e}_n)\}$. Since $f(\mathbf{e}_1) = \ldots = f(\mathbf{e}_k) = \mathbf{0}$, we conclude that:

$$\mathrm{Im}(f) = \mathrm{Span}\{f(\mathbf{e}_1), \ldots, f(\mathbf{e}_n)\} = \mathrm{Span}\{f(\mathbf{e}_{k+1}), \ldots, f(\mathbf{e}_n)\}.$$

2. Linear independence: Suppose

$$\sum_{i=k+1}^{n} \alpha_i f(\mathbf{e}_i) = \mathbf{0},$$

and recall that the desired conclusion is that all scalars α_i vanish. Let $\mathbf{g} = \sum_{i=k+1}^{n} \alpha_i \mathbf{e}_i$. Since f is linear, we have

$$f(\mathbf{g}) = f\left(\sum_{i=k+1}^{n} \alpha_i \mathbf{e}_i\right) = \sum_{i=k+1}^{n} \alpha_i f(\mathbf{e}_i) = \mathbf{0},$$

so that $\mathbf{g} \in \ker(f)$. But $\mathbf{e}_1, \ldots \mathbf{e}_k$ span the kernel of f, so there are scalars β_j such that

$$\mathbf{g} = \sum_{j=1}^{k} \beta_j \mathbf{e}_j.$$

Equating our two expressions for g we conclude that

$$\sum_{i=k+1}^{n} \alpha_i \mathbf{e}_i - \sum_{j=1}^{k} \beta_j \mathbf{e}_j = \mathbf{0}.$$

Since the set $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is linearly independent, we conclude that all scalar coefficients vanish, and, in particular, that all the scalars α_i are zero, which was the desired conclusion.

3.3 Systems of Linear Equations

We shall now explore some consequences of this theorem. Let A be a $k \times \ell$ matrix over a field F. Given a column vector $\mathbf{x} \in F^\ell$ we can use the matrix A to produce

$$A\mathbf{x} = \mathbf{y} \in F^k.$$

The map

$$\begin{aligned} f : F^\ell &\to F^k \\ \mathbf{x} &\mapsto A\mathbf{x} \end{aligned}$$

is a linear map by the distributivity of matrix multiplication.

Exercise 3.27. Prove that the rank of f is the rank of the matrix A (Hint: Use the column rank of A, and try to show that the image of f is the span of the column vectors of A).

One can do the previous exercise explicitly by writing $A = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_\ell \end{bmatrix}$ where $\mathbf{a}_i$ is the ith column of A, and hence an element of F^k. Using this notation, we may express the product:

$$A\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_\ell \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_\ell \end{bmatrix} = \sum_{i=1}^{\ell} x_i \mathbf{a}_i$$

We will now relate this to the theory of systems of linear equations. Express $A = (\alpha_{ij})_{\substack{i=1,\ldots,k \\ j=1,\ldots,\ell}}$, and $\mathbf{x}$ as a column vector as above. Consider the equation

$$A\mathbf{x} = \mathbf{0} \tag{1}$$

where A is a $k \times \ell$ matrix and $\mathbf{x} \in F^\ell$ is an unknown vector. We may expand this equation as the system of homogeneous linear equations ("homogoeneous" refers to the right hand side being zero):

$$\begin{aligned} \alpha_{11}x_1 + \alpha_{12}x_2 + \cdots + \alpha_{1\ell}x_\ell &= 0 \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \cdots + \alpha_{2\ell}x_\ell &= 0 \\ \vdots \qquad\qquad &\;\vdots\;\; \vdots \\ \alpha_{k1}x_1 + \alpha_{k2}x_2 + \cdots + \alpha_{k\ell}x_\ell &= 0 \end{aligned}$$

Note that $U = \ker(f) = \{\mathbf{x} \in F^\ell : A\mathbf{x} = \mathbf{0}\}$ is exactly the set of solutions to this homogeneous system of k linear equations. This system of linear equations is succinctly contained in the matrix equation $A\mathbf{x} = \mathbf{0}$.

Geometric Examples: Note that if we have a system of equations in 3 variables over $\mathbb{R}$, then the solutions will form a plane if there is only one equation, a line if there are two, and a single point (the origin) if there are three or more (Unless some of the equations are redundant). We now generalize this to higher dimensions:

Claim 3.28. *The set U of solutions to a system of homgeneous linear equations is a subspace: $U \leq F^\ell$.*

Proof:

1. $\mathbf{0} \in U$ is clear since choosing all coordinates x_i to be zero yields a solutions to any homogeneous linear equations.
2. U is closed under addition since $A\mathbf{x} = \mathbf{0}$ and $A\mathbf{y} = \mathbf{0}$ imply that $A(\mathbf{x}+\mathbf{y}) = A\mathbf{x}+A\mathbf{y} = \mathbf{0}$.
3. $\forall \lambda \in F$, we have $A(\lambda\mathbf{x}) = \lambda A\mathbf{x} = \mathbf{0}$ if $A\mathbf{x} = \mathbf{0}$.

We are now ready to state the fundamental theorem of the theory of systems of linear equations:

Theorem 3.29. *The dimension of the set of solutions of a system of homogeneous linear equations is the number of variables minus the rank of the system:* $\dim(U) = \ell - \mathrm{rk}(A)$.

Proof: Consider the linear transformation $f : \mathbf{x} \mapsto A\mathbf{x}$. We have defined U to be the kernel of f. By Theorem 3.24,

$$\dim U = \dim V - \mathrm{rk}(f) = \ell - \mathrm{rk}(f).$$

but Exercise 3.27 implies that $\mathrm{rk}(A) = \mathrm{rk}(f)$, which yields the desired formula

We can interpret this to say that each non-redundant equation reduces the dimension of the space of solutions by 1. Physicists would say that we lose a degree of freedom which each nonredundant equation.

3.4 Orthogonality

We will now apply this theorem to determine the maximum number of clubs in Eventown.

Our tool will be a set of geometric concepts over an arbitrary field F ("geometric algebra").

First, we introduce the dot product on F^n by the expression:

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} x_i y_i,$$

where $(x_i)_{i=1}^n$ and $(y_i)_{i=1}^n$ are the coordinates of $\mathbf{x}$ and $\mathbf{y}$, respectively.

Definition 3.30. $\mathbf{x}$ is *perpendicular* to $\mathbf{y}$ (denoted $\mathbf{x}\perp\mathbf{y}$) if $\mathbf{x}\cdot\mathbf{y}=0$.

Definition 3.31. A vector $\mathbf{x}$ is *isotropic* if $\mathbf{x}\cdot\mathbf{x}=0$.

Example 3.32. The vector

$$\begin{pmatrix} 1 \\ i \end{pmatrix}$$

an isotropic vector in $\mathbb{C}^2$ since $1^2+i^2=0$.

There are no isotropic vectors in $\mathbb{R}^n$. But there are finite fields $\mathbb{F}_p$ for which there are isotropic vectors in $\mathbb{F}_p^2$.

Exercise 3.33. For what primes p does there exist isotropic vectors in $\mathbb{F}_p^2$? (Hint: Experiment with small primes. Expect an answer of striking simplicity.)

Definition 3.34. Given two subsets A and B is F^n, we say that A is *perpendicular* to B (denoted $A\perp B$) if

$$(\forall\mathbf{a}\in A)(\forall\mathbf{b}\in B)(\mathbf{a}\perp\mathbf{b})$$

Definition 3.35. Given $A\subseteq F^n$, we define $A^\perp$ (pronounced A-perp):

$$A^\perp := \{\mathbf{b}\in F^n : \mathbf{b}\perp A\}$$

Note that $\emptyset^\perp = F^n$.

Exercise 3.36. $A^\perp \le F^n$ (Hint: Use distributivity of the dot product).

Theorem 3.37. *If $S\subseteq F^n$, then*

$$\dim(S^\perp) = n - \operatorname{rk}(S).$$

Proof: Note that if $r=\operatorname{rk}(S)$, then we may choose $\mathbf{b}_1,\dots,\mathbf{b}_r$ such that

$$\operatorname{Span}(S) = \operatorname{Span}\{\mathbf{b}_1,\dots,\mathbf{b}_r\},$$

and that $\mathbf{x}\perp S$ if and only if $\mathbf{b}_i\cdot\mathbf{x}=0$ for all $i=1,\dots r$, which yields a system of linear equations. Since the vectors $\{\mathbf{b}_i\}_{i=1}^r$ are linearly independent, the rank of this system is r, so the dimension of the space of solutions is $n-r$ by the fundamental theorem of the theory of systems of linear equations. But the space of solutions is precisely $S^\perp$.

Corollary 3.38. *If $U\le F^n$, then $\dim U + \dim U^\perp = n$.*

Proof:

$$\dim U^\perp = n - \dim\operatorname{Span} U = n - \dim U$$

Corollary 3.39. *If $U\le F^n$, then $U^{\perp\perp}=U$.*

Proof: As usual, a proof of equality between two sets has two parts:

- $U \subseteq U^{\perp\perp}$ is obvious from the definition of $\perp$.
- To check that a subspace is equal to the full vector space, it suffices to check that they have the same dimension. Therefore, we compute:

$$\dim U^{\perp\perp} = n - \dim U^{\perp} = n - (n - \dim U) = \dim U.$$

Definition 3.40. U is a *totally isotropic* subspace of F^n if $U \perp U$.

Example 3.41. The 1-dimensional subspace generated by the vector

$$\begin{pmatrix} 1 \\ i \end{pmatrix}$$

in $\mathbb{C}^2$ is isotropic.

Corollary 3.42. *If U is a totally isotropic subspace of bF^n then*

$$\dim U \leq \lfloor n/2 \rfloor$$

Proof: Observe that U is a totally isotropic if and only if $U \subseteq U^{\perp}$, which implies that $\dim U \leq \dim U^{\perp}$. Since $\dim U + \dim U^{\perp} = n$, we conclude that $2 \dim U \leq n$.

Corollary 3.43. *If $S \subseteq F^n$ and $S \perp S$, then*

$$\mathrm{rk}(S) \leq \lfloor n/2 \rfloor.$$

Proof: Let $U = \mathrm{Span}(S)$. Since $S \perp S$ implies that $U \perp U$, we may use the previous Corollary to conclude that $\mathrm{rk}\, S = \dim U \leq \lfloor n/2 \rfloor$.

Corollary 3.44. *If $S \subseteq \mathbb{F}_p^n$ and $S \perp S$, then*

$$\mathrm{rk}(S) \leq p^{\lfloor n/2 \rfloor}.$$

We now return to Eventown. Recall that to each club C (which is a subset of $\{1, \ldots, n\}$) we assigned an incidence vector $v_C = (\alpha_1, \ldots, \alpha_n)$ where

$$\alpha_i = \begin{cases} 1 & i \in C \\ 0 & i \notin C \end{cases}$$

Note that $v_A \cdot v_B = |A \cap B|$, and that, in particular, $v_A \cdot v_A = |A|$. This suggests that we consider our incidence vectors in a vector space over $\mathbb{F}_2$, so that the rules of Eventown become:

1. $v_A \cdot v_A = 0$ for any club A, and
2. $v_A \cdot v_B = 0$ for any two clubs A and B.

In other words, if $S \subseteq \mathbb{F}_2^n$ is the set of incidence vectors of Eventown clubs, then $S \perp S$. We can now apply the previous results (Corollary 3.44) to conclude:

Corollary 3.45.

$$|S| \leq 2^{\lfloor n/2 \rfloor}.$$

We observe that every maximal system of Eventown clubs corresponds to a totally isotropic subspace of $\mathbb{F}_2^n$.

Exercise 3.46. Prove that if $U \leq \mathbb{F}_2^n$ is totally isotropic then there exists $W \leq \mathbb{F}_2^n$ which is also totally isotropic such that $U \leq W$ and $\dim W = \lfloor n/2 \rfloor$.

Corollary 3.47. *Every maximal Eventown system of clubs is* maximum *(i. e., it consists of $2^{\lfloor n/2 \rfloor}$ clubs).*

Notice the contrast with the result we proved for Oddtown.

Exercise 3.48. Using the previous Corollary, prove that there exists a maximal Eventown system of clubs which is not the "married couple" solution.