Linear Algebra, 4th day, Thursday 7/1/04
REU 2004. Info:
http://people.cs.uchicago.edu/∼laci/reu04.

Instructor: László Babai
Scribe: Nick Gurski

1 Linear maps

We shall study the notion of maps between vector spaces, called linear maps or homomorphisms. A function $f : V \to W$ between vector spaces defined over the same field $\mathbb{F}$ is a **linear map** if

$$\begin{aligned} &(1) \quad f(\underline{x}+\underline{y}) = f(\underline{x}) + f(\underline{y}) \\ &(2) \quad f(\alpha \underline{x}) = \alpha f(\underline{x}) \end{aligned}$$

for all $\underline{x}, \underline{y} \in V$ and all $\alpha \in \mathbb{F}$; these two conditions are equivalent to the single condition that f distributes over linear combinations

$$f(\sum_i \alpha_i \underline{x}_i) = \sum_i \alpha_i f(\underline{x}_i).$$

Now assume that we have a basis

$$\mathbf{e} = \{\underline{e}_1, \underline{e}_2, \ldots, \underline{e}_n\}$$

for V and a basis $\mathbf{e}'$ for W. We would like to assign a matrix to the linear map f which depends on the two bases we have chosen and represents the function f completely. We do this by defining the matrix $[f]_{(\mathbf{e},\mathbf{e}')}$ to have as its i-th column the vector $[f(\underline{e}_i)]_{\mathbf{e}'}$. Thus we can write this matrix as

$$[f]_{(\mathbf{e},\mathbf{e}')} = [[f(\underline{e}_1)]_{\mathbf{e}'} \cdots [f(\underline{e}_n)]_{\mathbf{e}'}];$$

if W is of dimension k, this matrix is a $k \times n$ matrix. Our next goal is to show that this matrix $[f]_{(\mathbf{e},\mathbf{e}')}$ gives the same geometric information as the linear map f.

We begin this task with a theorem.

Theorem 4.1. *Let* $\mathbf{e} = \{\underline{e}_1, \cdots, \underline{e}_n\}$ *be a basis for* V*. Then for any set of vectors* $\underline{w}_1, \cdots, \underline{w}_n \in W$*, there exists a unique linear map* $f : V \to W$ *with the property that* $f(\underline{e}_i) = \underline{w}_i$ *for all* i*.*

Proof: We first show uniqueness, and existence follows easily afterwards. Assuming there is such an f, then its value on a vector $\underline{x} \in V$ is determined since $\mathbf{e}$ is a basis. Writing $\underline{x} = \sum_{i=1}^{n} \alpha_i \underline{e}_i$, we have that

$$f(\underline{x}) = f(\sum_{i=1}^{n} \alpha_i \underline{e}_i) = \sum_{i=1}^{n} \alpha_i f(\underline{e}_i);$$

this shows uniqueness.

To show the existence of such a map, we define

$$f(\underline{x}) = \sum_{i=1}^{n} \alpha_i \underline{w}_i$$

and check that this is a linear map. This is left as a simple exercise.

The theorem means that the matrix $[f]$ is determined by the images $f(\underline{e}_i)$ of each of the basis vectors in V, and by the definition of basis this determines the linear map f as well. Conversely, if we know where f sends each basis vector, then there is a unique linear map that agrees with f on $\mathbf{e}$. Thus the matrix $[f]$ and the linear map f give the same information.

Example 4.2 ((The derivative of polynomials)). Remember that the collection of all polynomials of degree n, denoted P_n, is a vector space of dimension $n+1$ which is spanned by the basis $1, x, x^2, \cdots, x^n$. There is a linear map D which sends the polynomial $p(x)$ to its derivative polynomial $p'(x)$; thus we have a linear map $D : P_n \to P_{n-1}$. We shall compute the matrix $[D]$ with respect to the natural bases. We know that $D(x^k) = kx^{k-1}$, so $D(\underline{e}_k) = k\underline{e}_{k-1}$ where k ranges from 0 to n. This shows that the matrix $[D]$ has, as its k-th column the vector which is all zeros except that its $(k-1)$-st entry is $(k-1)$; this is an $n \times (n+1)$ matrix. For $n = 3$, the matrix for $D : P_3 \to P_2$ is shown below.

$$[D] = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

Example 4.3. Projection onto the plane:

We have a linear map $\pi : \mathbb{R}^3 \to \mathbb{R}^2$ which forgets about the z-coordinate. Using the standard basis, we have

$$\begin{aligned} \pi(\underline{e}_1) &= \underline{e}_1 \\ \pi(\underline{e}_2) &= \underline{e}_2 \\ \pi(\underline{e}_3) &= 0 \end{aligned}$$

and so the 2×3 matrix $[\pi]$ is

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Example 4.4. Rotation of the plane by the angle α: We denote by ρ_α the map which rotates $\mathbb{R}^2$ by the fixed angle α. Using the standard basis, we find that

$$[\rho_\alpha] = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}.$$

If we use a different basis, however, the matrix for ρ_α changes. Let $\underset{\sim}{e}_1$ be the vector $(1,0)$ as before, but now let $\underset{\sim}{e}_3$ be the vector obtained by rotating $\underset{\sim}{e}_1$ by α. In general, these vectors are linearly independent, and thus form a basis for $\mathbb{R}^2$. Now we will compute $[\rho_\alpha]$ with respect to this basis. By definition, $\rho_\alpha(\underset{\sim}{e}_1) = \underset{\sim}{e}_3$; using some basic geometry we can determine $\rho_\alpha(\underset{\sim}{e}_3)$ and thus find that the matrix $[\rho_\alpha]$ is now

$$\begin{bmatrix} 0 & -1 \\ 1 & 2\cos\alpha \end{bmatrix}.$$

This examples shows that when computing the matrix associated to a linear map, it is important to remember what bases are involved.

We can now make two definitions that will be useful later.

Definition 4.5. Let A be an $n \times n$ matrix (it is important that A is a square matrix here). Then we define the **trace** of A, denoted tr$A$, to be the sum of the diagonal entries of A. If A is the 2×2 matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

then we define the **determinant** of A to be

$$\det A = ad - bc.$$

We will later see how to define determinant for $n \times n$ matrices. If we are given a linear map $f : V \to V$, then the trace and determinant of the matrix $[f]$ do not depend on what basis we have chosen. In the examples above of the two matrices associated to the rotation map, it is easy to check that both matrices have trace $2\cos\alpha$ and determinant 1.

Assume that we are given two matrices, $A = (a_{ij})$ and $B = (b_{jk})$ where A is an $r \times s$ matrix and B is an $s \times t$ matrix. Then we can multiply them to get a matrix $C = AB$, where $C = (c_{ik})$ and the entry c_{ik} is equal to

$$\sum_{j=1}^{s} a_{ij} b_{jk}.$$

Note that A must have the same number of rows as B has columns or the indices will not match as they should.

Let $f : V \to W$ be a linear map, with bases $\mathbf{e}$ for V and $\mathbf{e}'$ for W. Given a vector $\underset{\sim}{x} \in V$, we would like to determine the coordinates of the vector $[f(\underset{\sim}{x})]$ with respect to $\mathbf{e}'$ in terms of the corrdinates of $\underset{\sim}{x}$ with respect to $\mathbf{e}$ and the matrix $[f]_{(\mathbf{e},\mathbf{e}')}$.

Exercise 4.6. Using the definition of matrix multiplication above, we have the formula

$$[f]_{(\mathbf{e},\mathbf{e}')}[\underline{x}]_{\mathbf{e}} = [f(\underline{x})]_{\mathbf{e}'}.$$

We can use this exercise to study the relationship between the composite of two linear maps and the product of their matrices. Let $f : V \to W$ and $g : W \to Z$ be two linear maps, and let V, W, Z have bases $\mathbf{e}, \mathbf{e}', \mathbf{e}''$, respectively. We have the composite map gf which is defined on vectors as $(gf)(\underline{x}) = g(f(\underline{x}))$; we also have matrices $[f]_{(\mathbf{e},\mathbf{e}')}$ and $[g]_{(\mathbf{e}',\mathbf{e}'')}$, as well as a matrix for the composite gf with respect to $\mathbf{e}$ and $\mathbf{e}''$. The goal is to prove that

$$[gf]_{(\mathbf{e},\mathbf{e}'')} = [g]_{(\mathbf{e}',\mathbf{e}'')}[f]_{(\mathbf{e},\mathbf{e}')}.$$

To do this, a lemma is required.

Lemma 4.7. *If A is a $k \times n$ matrix, then $A = 0$ if and only if for all $\underline{x} \in \mathbb{F}^n$, $A\underline{x} = 0$.*

Proof: If $A = 0$, it is obvious that $A\underline{x} = 0$ for all $\underline{x}$. Assume that $A \neq 0$; we will produce a vector $\underline{x}$ such that $A\underline{x} \neq 0$, and that will complete the proof. Since A is nonzero, it has a nonzero entry; call it a_{ij}. Then $A\underline{e}_j = \sum_l a_{lj}\underline{e}_l$, where the vector $\underline{e}_l$ is zero except for entry l which is 1. This vector is nonzero, since a_{ij} is nonzero.

Corollary 4.8. *If for all vectors $\underline{x}$, $A\underline{x} = B\underline{x}$, then $A = B$ as matrices.*

Proof: If $A\underline{x} = B\underline{x}$, then $(A - B)\underline{x} = 0$ for all $\underline{x}$ and therefore $A - B = 0$ by the above.

Now we can show that $[gf] = [g][f]$ by showing that these two matrices are equal when applied to any vector $\underline{x}$. On the left side, we have

$$[gf][\underline{x}] = [(gf)(\underline{x})]$$

by the exercise above. On the right, we have

$$[g][f][\underline{x}] = [g][f(\underline{x})] = [g(f(\underline{x}))],$$

but $[(gf)(\underline{x})] = [g(f(\underline{x}))]$ since that is exactly how the composition of two functions is defined. This proves that $[gf] = [g][f]$.

Returning to our example of the linear map which rotates by α, we can use that $[\rho_\alpha][\rho_\beta] = [\rho_\alpha\rho_\beta]$ to prove the angle-sum formula from trigonometry. Since rotating by β and then rotating by α is the same as rotating by $\alpha + \beta$, we get that

$$[\rho_\alpha\rho_\beta] = [\rho_{\alpha+\beta}] = \begin{bmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta) \\ \sin(\alpha+\beta) & \cos(\alpha+\beta) \end{bmatrix}.$$

Multiplying the matrices for ρ_α and ρ_β together, the result is

$$\begin{bmatrix} \cos\alpha\cos\beta - \sin\alpha\sin\beta & -(\sin\alpha\cos\beta + \cos\alpha\sin\beta) \\ \sin\alpha\cos\beta + \cos\alpha\sin\beta & \cos\alpha\cos\beta - \sin\alpha\sin\beta \end{bmatrix};$$

comparing these two matrices gives the desired identity.

Now we turn to using matrices to solve systems of linear equations. If $A = [\underline{a}_1 \underline{a}_2 \cdots \underline{a}_n]$ is a $k \times n$ matrix with columns $\underline{a}_i$, and $\underline{x}$ is any vector, then $A\underline{x} = x_1\underline{a}_1 + x_2\underline{a}_2 + \cdots + x_n\underline{a}_n$. Therefore the equation $A\underline{x} = 0$, where $\underline{x}$ is the variable, always has the trivial solution $\underline{x} = \underline{0}$.

Exercise 4.9. $A\underline{x} = 0$ has a nontrivial solution if and only if the columns of A are linearly dependent if and only if the rank of A is less than n.

Exercise 4.10. If A is a $k \times n$ matrix, then $A\underline{x} = 0$ has a nontrivial solution if $k < n$ (HINT: use Exercise **??**).

Exercise 4.11 (Showing the existence of a nontrivial solution). Using the corollary above, we can show that the system of equations

$$\begin{aligned} 2x + 3y - z &= 0 \\ 7x + 8y + 5z &= 0 \end{aligned}$$

has a nontrivial solution. Solving this system is equivalent to solving $A\underline{x} = 0$ where

$$A = \begin{bmatrix} 2 & 3 & -1 \\ 7 & 8 & 5 \end{bmatrix}$$

and

$$\underline{x} = \begin{bmatrix} 7 \\ 8 \\ 5 \end{bmatrix}.$$

In this case, $k = 2$ and $n = 3$, so there must be a nontrivial solution.

If we try to solve the general equation $A\underline{x} = \underline{b}$, this is equivalent to solving

$$\sum_i x_i \underline{a}_i = \underline{b},$$

or finding out if $\underline{b}$ is in the span of the vectors $\underline{a}_i$.

Exercise 4.12. The equation $A\underline{x} = \underline{b}$ is solvable if and only if $\underline{b} \in \mathrm{Span}\{\underline{a}_i\}$ if and only if the rank of A is the same as the rank of $[A|\underline{b}]$.

Let $U = \{\underline{x} : A\underline{x} = 0\}$; this is called the solution space. It is easy to see that U is a subspace of $\mathbb{F}^n$. It is obvious that U contains the zero vector. If $\underline{x}, \underline{y} \in U$, then $(\underline{x} + \underline{y}) \in U$ since

$$A(\underline{x} + \underline{y}) = A\underline{x} + A\underline{y} = 0 + 0 = 0.$$

Similarly, if $\underline{x} \in U$, then $\alpha\underline{x} \in U$.

Theorem 4.13. *The dimension of U is $n - \mathrm{rk}(A)$.*

Proof: Define $\varphi : \mathbb{F}^n \to \mathbb{F}^k$ by $\varphi(\underline{x}) = A\underline{x}$; φ is linear because A is a matrix. The kernel of φ is exactly the subspace U defined above. We know that $\dim(\ker\varphi) = n - \dim(\mathrm{im}\varphi)$, so we only need to show that $\dim(\mathrm{im}\varphi) = \mathrm{rk}(A)$. But the image of φ is the set of all vectors of the form $A\underline{x}$, and this subspace is spanned by the columns of A, hence the dimension of the image is necessarily the rank of A.

It is important to note that $A\underline{x} = \underline{b}$ may or may not have a solution. Assume that it does have a solution, and call that solution $\underline{x}_0$. To find all solutions to $A\underline{x} = \underline{b}$, we can subtract

$$\begin{array}{ccc} & A\underline{x} & = & \underline{b} \\ - & A\underline{x}_0 & = & \underline{b} \\ \hline & A(\underline{x} - \underline{x}_0) & = & 0 \end{array}$$

to find that all solutions of $A\underline{x} = \underline{b}$ are translates of the vector $\underline{x}_0$ by elements of U. Another way to say this is that $\underline{x} - \underline{x}_0$ is a solution to $A\underline{x} = 0$ (where $\underline{x}$ really means two different things in this sentence) , so $\underline{x} - \underline{x}_0 \in U$; this is the same as saying that $\underline{x} \in \underline{x}_0 + U$ or that $\underline{x}$ has the form $\underline{x}_0 + \underline{w}$ for $\underline{w} \in U$.