Apprentice Program: Linear Algebra

Instructor: Miklós Abért
Notes taken by Matt Holden and Kate Ponto

June 26,2006

1 Matrices

An $n \times k$ *matrix* A over a ring R is a collection of nk elements of R, arranged in n rows and k columns. The element of A in row i and column j is denoted $A_{ij} \in R$.

The *transpose* A^T of A is defined by $A^T_{ij} = A_{ji}$. Addition of matrices is defined by $(A + B)_{ij} = A_{ij} + B_{ij}$.

The product of A and B is only defined if the number of columns of A equals the number of rows of B. If A is an $n \times k$ matrix and B is a $k \times m$ matrix then the product AB is defined by

$$(AB)_{ij} = \sum_{l=1}^{k} A_{il} B_{lj},$$

so the ij entry of AB is the dot product of the ith row of A with the jth column of B. In particular, if $n = 1$ and $m = 1$ (i.e., A is a row vector and B is a column vector) then AB is simply the dot product of A and B.

The set of all $n \times n$ matrices over R is denoted by $M_n(R)$.

Claim 1. *$M_n(R)$ is a ring.*

As an exercise, you should check, for instance, that $A(B+C) = AB+AC$ and $A(BC) = (AB)C$.

The zero element of this ring is the *zero matrix*, every entry of which is $0 \in R$. The unit is the *identity matrix*

$$I = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Notice that the ring $M_n(R)$ is *not* commutative for $n \geq 2$, even when R is commutative. For example, in $M_2(\mathbb{Z})$ we have

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

This example also shows that a product of nonzero matrices can equal zero, so $M_n(R)$ is said to have nonzero *zero divisors*. This is in sharp contrast to other familiar rings such as $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$, which have no nonzero zero divisors (so $xy = 0$ in these rings only if $x = 0$ or $y = 0$).

2 Permutations

A *permutation* of the set $\{1, \dots, n\}$ is simply a bijection from this set to itself. Notice that a composition of permutations is again a permutation, and each permutation has an inverse (since bijections are invertible). Thus, it is easy to see that the set $\mathrm{Sym}(n)$ of all permutations of $\{1, \dots, n\}$ forms a group, called the *symmetric group* on this set. For example, $\mathrm{Sym}(7)$ contains the permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 4 & 1 & 3 & 5 & 7 & 6 \end{pmatrix}$$

which maps $1 \mapsto 2, \quad 2 \mapsto 4, \quad 3 \mapsto 1$, etc. We can write this same permutation more efficiently using *cycle notation* as (1 2 4 3)(5)(6 7). We typically omit the fixed points of a permutation in cycle notation, so we would write simply (1 2 4 3)(6 7).

A permutation that interchanges two numbers $i \neq j$ and fixes every other number (written $(i\ j)$ in cycle notation) is called a *transposition*.

Consider now the *bubble sort* algorithm for sorting a sequence $a_1, \dots, a_n$ of n numbers. For $i = 1, \dots, n-1$, if $a_i > a_{i+1}$ then we interchange a_i and a_{i+1}. After traversing the sequence once, the largest number in the

sequence must be in the last position. After traversing it k times, the largest k numbers are in correct position (try proving this rigorously by induction). Thus, we obtain the sorted sequence $1, 2, \ldots, n$ after at most n traversals of the sequence, and each traversal uses at most $n-1$ transpositions, so bubble sorting n numbers takes at most $n(n-1)$ steps. (In fact, a more careful analysis shows that it takes at most $n(n-1)/2$ steps, and this bound is the best possible.)

Thus, if $a_1, \ldots, a_n$ represents a permutation π of $\{1, \ldots, n\}$, then we can find finitely many transpositions $t_1, \ldots, t_k$ such that $\pi t_1 \ldots t_k = \mathrm{id}$, hence

$$\pi = t_k^{-1} \ldots t_1^{-1} = t_k \ldots t_1.$$

We have proved the following:

Proposition 2. *Every permutation is a product of transpositions.*

For the example permutation given above, bubble sort takes only two traversals to sort the sequence, with intermediate steps:

$$\begin{array}{c} 2\ 4\ 1\ 3\ 5\ 7\ 6 \\ 2\ 1\ 4\ 3\ 5\ 7\ 6 \\ 2\ 1\ 3\ 4\ 5\ 7\ 6 \\ 2\ 1\ 3\ 4\ 5\ 6\ 7 \\ 1\ 2\ 3\ 4\ 5\ 6\ 7 \end{array}$$

Reading the transpositions from these intermediate steps, we can write (using cycle notation):

$$(1\ 2\ 4\ 3)(6\ 7) = (1\ 2)(6\ 7)(3\ 4)(2\ 3).$$

We define the *sign* of a permutation π to be

$$\mathrm{sgn}(\pi) = \begin{cases} +1 & , \quad \pi \text{ is a product of an even number of transpositions} \\ -1 & , \quad \pi \text{ is a product of an odd number of transpositions.} \end{cases}$$

In order to show that $\mathrm{sgn}(\pi)$ is well-defined, we must prove the following

Lemma 3. *If $t_1, \ldots, t_k$ and $r_1, \ldots, r_l$ are transpositions such that $t_1 \cdots t_k = r_1 \cdots r_l$ then $k \equiv l \pmod 2$.*

Proof. Notice that it is enough to prove that a product of an odd number of transpositions cannot equal the identity. To prove this, consider the function $s : \mathrm{Sym}(n) \to \{\pm 1\}$ defined by

$$s(\pi) = \frac{\prod_{i<j}(x_{\pi(i)} - x_{\pi(j)})}{\prod_{i<j}(x_i - x_j)}.$$

As an exercise, you should show that $s(\pi\tau) = -s(\pi)$ whenever τ is a transposition. It follows that if $t_1 \dots t_k = \mathrm{id}$ then $(-1)^k = s(\mathrm{id}) = 1$, so k is even, as claimed. □

We say a permutation π is *even* if $\mathrm{sgn}(\pi) = 1$ and *odd* if $\mathrm{sgn}(\pi) = -1$. Notice that cycles of even length are odd and vice versa. In fact, we can determine the sign of π from its cycle structure: if we write π as a product of *disjoint* cycles then

$$\mathrm{sgn}(\pi) = (-1)^{\#\text{ of even length cycles in } \pi}.$$

We showed above that (1 2 4 3)(6 7) is a product of four transpositions, so it's an even permutation. We can also see this from the fact that its disjoint cycle decomposition contains two even length cycles.

3 Determinants

Question: how many non-attacking rooks can be placed on an $n \times n$ chessboard?

We can place at most one in each row, so we can place at most n on the board. On the other hand, if there are less than n rooks on the board then there is some row and some column without rooks, and we are free to place a rook at the intersection of this row and column. Thus, n is the largest number of rooks we can place on the board. Call such a placement of n rooks a *rook configuration*. Then there is a natural one-one correspondence between rook configurations on an $n \times n$ board and permutations of $\{1, \dots, n\}$. In particular, there are $n!$ different rook configurations.

Given $A \in M_n(R)$, we define

$$\det A = \sum_{\pi \in \mathrm{Sym}(n)} \mathrm{sgn}(\pi) \prod_{k=1}^{n} A_{k,\pi(k)}.$$

As an exercise, you should prove the following basic properties:

- if A has a zero row or a zero column then $\det A = 0$,
- $\det A^T = \det A$.

Theorem 4 (Properties of the determinant). *Let $A \in M_n(R)$.*

1. *If A has an entire row or column of zeros, then $det(A) = 0$.*
2. *If A'=multiply the i-th row of A by c, then $det(A') = c \cdot det(A)$.*
3. *If A'=change the i-th and j-th rows of A, then $det(A') = -det(A)$.*
4. *If the i-th row of A =c·(j-th row of A), then $det(A) = 0$*
5. *If $A = A' = A''$ except in the i-th row and the i-th row of A is the sum of the i-th rows of A' and A'', then $det(A) = det(A') + det(A'')$.*
6. *If A' = add c·(i-th row of A) to the j-th row of A, then $det(A') = det(A)$.*

In the statements above 'row' can be replaced by 'column'.

Exercise 5. Let $A \in M_n(\mathbb{Z})$ such that every row of A has sum divisible by 7. Then $det(A)$ is divisible by 7.

Theorem 6. *For all $A, B \in M_n(R)$, $det(A \cdot B) = det(A)det(B)$.*

Proof. To compute the determinant of the matrix

$$\begin{pmatrix} A & 0 \\ * & B \end{pmatrix}$$

we only need to consider rook configurations that do not have any rooks in the top right corner. These rook configurations are exactly the rook configurations that have all rooks contained in A and B. So

$$det \begin{pmatrix} A & 0 \\ * & B \end{pmatrix} = det(A)det(B)$$

If we take $*$ to be $-I$, then we have the matrix

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

Without changing the determinant, we can convert this to a matrix having all zeros in the bottom right $n \times n$ matrix. We do this by adding multiples of the first n columns in the matrix to the last n columns. (To eliminate the i, j entry of B, add B_{ij} times the j-th column to the $j+n$-th column.) The new matrix is

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

and

$$det \begin{pmatrix} A & AB \\ -I & 0 \end{pmatrix} = (-1)^n det \begin{pmatrix} AB & A \\ 0 & -I \end{pmatrix} = (-1)^n det(A \cdot B) det(-I) = det(AB)$$

□

Theorem 7. *Let $a, b \in \mathbb{Z}$. If A is the $n \times n$ matrix with a's on the diagonal and all other entires b, then*

$$det(A) = (a-b)^{n-1}(a + b(n-1))$$

Proof. exercise (*hint:* Use row and column transformations and try to get lots of zeros.) □

Theorem 8. *If n people form c clubs such that:*

- *every club has the same number of members, (call this number a)*
- *every two clubs have the same number of common members, (call this number b)*
- *$a, b > 0$ and $a \neq b$,*

then $c \leq n$.

Proof. Assume $c > n$. Let D be the $c \times n$ matrix such that D_{ij} is 0 if person j is *not* a member of club i, D_{ij} is 1 if person j is a member of club i.

Since every club has the same number of members, the row sum is the same for all rows of D (and is a). Since every two clubs have the same number of common members, the dot product of any two different rows in D is the same (and is b).

Extend the matrix D to a $c \times c$ matrix A by adding columns of zeros. Like for D, the dot product of any row in A with itself is a and the dot product of any two different rows in A is b. Then is the same as saying that

$$AA^t = \begin{pmatrix} a & b & \dots & b \\ b & a & & b \\ \vdots & & \ddots & \vdots \\ b & \dots & b & a \end{pmatrix}$$

By Theorem 7, $det(AA^t) \neq 0$, but $det(A) = 0$. A contradiction. □

Exercise 9.

$$det \begin{pmatrix} 1 & a_1 & a_1^2 & a_1^3 & \dots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & a_2^3 & \dots & a_2^{n-1} \\ \vdots & & & & & \vdots \\ 1 & a_n & a_n^2 & a_n^3 & \dots & a_n^{n-1} \end{pmatrix} = ?$$

The determinant of the matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is the area of the parallelogram with two sides given by the vectors (a, b) and (c, d). The determinant of a 3×3 matrix is the volume of a parallelepiped given by the three rows. This parallelepiped will have volume 0 only when all three of the vectors are contained in a plane containing the origin in $\mathbb{R}^3$.