Miscellaneous Mathematical Concepts II

1. Matrix Lore

1.1. Matrix multiplication and dot products

Notice that a row vector times a column vector is really a dot product: If we write all vectors as column vectors, we have $v \cdot w = v^t w$. (Here we treat a 1×1 matrix as a scalar.)

A matrix product $C = AB$ is really a matrix of dot products:

C_{ij} is the dot product of row i of A with column j of B.

Problem V-1. Suppose the rows of an $n \times n$ matrix P are orthonormal, meaning that they are all of length 1 and that any two are orthogonal (perpendicular). Show that $PP^t = I$.

(Method: The length of a vector v is $\sqrt{v \cdot v}$, so v has length 1 when $v \cdot v = 1$. Also, two vectors are perpendicular when their dot product is 0.)

1.2. What affects entries in a product

From §1.1 you can see this:

Proposition. In a matrix product AB,

(i) row i of AB is affected only by row i of A; in fact, it's that row times B.

(ii) column j of AB is affected only by column j of B; in fact, it's A times that column.

Problem V-2. (a) For $C = AB$, fill in all entries of C that you know if $A = \begin{bmatrix} \cdot & \cdot & \cdot \\ 2 & 3 & 5 \\ \cdot & \cdot & \cdot \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 & 0 \\ 10 & 10 & 10 \\ 1 & 0 & 0 \end{bmatrix}$.

(b) The middle row of C is a linear combination of the rows of B. With what coefficients?

Problem V-3. If you need to transform 1000 points of $\mathbb{R}^2$ by τ_M for some 2×2 matrix M, instead of doing 100 separate matrix multiplications, you can make them the columns of a matrix P and find MP. Explain why that works.

1.3. Multiplying matrices on the left

Multiplying on the left by an elementary matrix is part of the following general principle:

Proposition. If you start with a matrix A and multiply on the left by M, then in MA every row is a linear combination of rows of A. Moreover, just *which* linear combinations depends on M and not on A.

To invent an M so that MA produces particular linear combinations of the rows of A, look at the case $A = I$, which gives $MI = M$.

For example, if you want a matrix M whose effect when multiplied on the left is to subtract 2 times row 1 from row 2, then make M by starting with the identity matrix of the right size and subtracting 2 times row 1 from row 2. In the 3×3 case this would be $M = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

In other words, to get M, do to the identity matrix what you want M to do to other matrices, when M is used as a multiplier on the left.

As you know, many religions and cultures have a version of the Golden Rule, which in older English is "Do unto others what you would have others do unto you." So it is said that the Golden Rule for matrix multiplication (on the left) is

> "Do unto others what you do unto I."

Problem V-4. (a) Suppose you want to scale the rows of a 3×3 matrix A by 10, 5, and 2 respectively. By what matrix should you multiply A on the left to accomplish this?

(b) What should you do if instead you want to scale the *columns* of A by the same three factors?

Problem V-5. A *permutation matrix* is a matrix obtained by permuting the rows of an identity matrix, or equivalently, a matrix that has zero entries except for a single 1 in each row and column. An example is $P = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$. What does τ_P do to $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$?

(Method: Use the Golden Rule. Of course, this can be computed directly too; the point is that you can tell what happens more intuitively, without doing a computation.)

Problem V-6. For $N = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$,

(a) Calculate NH where $H = \begin{bmatrix} H & E & L & P & I & A & M \\ S & I & N & K & I & N & G \end{bmatrix}$.

(b) If you wanted to produce that effect on H by multiplying on the left, how would you know to use N?

(c) Find N^2 in your head, without writing anything down except the answer.

(d) Find N_4^4, where $N_4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$.

Note. If M is a square matrix such that $M^k = 0$ for some k, then M is said to be **nilpotent** (meaning "nothing-power"). Nilpotent matrices turn out to be important building blocks in the theory of "similar" matrices.

$N = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ and its transpose are the simplest nilpotent matrices other than the zero matrix. They often provide counterexamples to statements you might think should be true.

Example: From experience with numbers, you might think $A^2 = 0 \Rightarrow A = 0$, but that's not true for matrices and N is a counterexample.

Problem V-7. Show that if v and w are orthogonal column vectors in $\mathbb{R}^n$ (see Problem V-1), then the $n \times n$ matrix wv^t is nilpotent.

Proposition. Any invertible matrix M is a product of elementary matrices.

Proof. Starting with an $n \times n$ invertible matrix M and row-reducing to row-reduced echelon form (RREF), since M is invertible and therefore has rank n, the RREF must also have rank n and so is I. Therefore the net effect is $E_k \dots E_1 M = I$. Then $M = E_1^{-1} \dots E_k^{-1}$, and inverses of elementary matrices are elementary.

Corollary. Any product MA with M invertible can be achieved by starting with A and performing elementary row operations.

Corollary. To invert A (assuming it *is* invertible), start with $[A|I]$ and row-reduce until you reach the form $[I|M]$. Then $M = A^{-1}$.

Proof. Row-reducing by a product of elementary matrices equalling some M results in $[MA|MI]$, so if $MA = I$ then the second half of the matrix is $M = A^{-1}$.

Problem V-8. Find the inverse of $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix}$.

Here is a handy shortcut for finding $A^{-1}B$, something that often occurs.

Proposition. If A is an invertible $n \times n$ matrix and B is an $m \times n$ matrix, the row-reduced echelon form of the augmented matrix $[A|B]$ is $[I|A^{-1}B]$.

Proof. Going from A to I means the effect of the row-reduction is to multiply by A^{-1}, and the Golden Rule of matrix multiplication says B is affected the same way.

1.4. Inverses of 2×2 matrices

Problem V-9. For the general 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, show that $AB = (\det A)I$, where $B = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

This gives a handy formula for the inverse of a 2×2 matrix, so learn it:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{\Delta} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}, \text{ where } \Delta = ad - bc, \text{ if } \Delta \neq 0.$$

Notice that the inverse is a scaled version of the 2×2 matrix obtained by *switching* the diagonal entries of the original matrix and and *negating* the off-diagonal entries (while *not* switching them).

Problem V-10. Re-do Problem V-8 using this formula.

2. Permutations

2.1. Definitions

Definition. A *permutation* of a set X is a one-to-one correspondence on $X \to X$.

Sometimes we are interested in permutations of a specific set, such as rows of a matrix, as in Problem V-5. It is important, though, to understand the structure of various permutations on n symbols for each specific value of n, and for that it makes little difference what the set is. Therefore we often use $X = \{1, 2, \ldots, n\}$ as a set of symbols to permute.

The set of permutations on $\{1, \ldots, n\}$ is denoted S_n, the *symmetric group* on n symbols[1]. For example, there are six permutations on three symbols, so S_3 has six elements. In general, S_n has $n!$ elements.

Temporarily, let's describe permutations by writing the symbols each with its image beneath. For example,

$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ means the permutation for which $1 \mapsto 2, 2 \mapsto 3, 3 \mapsto 1$.

Definition. Permutations are multiplied by taking their composition.

Problem V-11. What is $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$?

(Method: We are composing functions f and g. Find $f(g(1)), f(g(2)), f(g(3))$ and then write the answer in the matrix-like form[2].)

2.2. Cycles

A permutation is a **cycle** if some symbols go in a single cycle when the permutation is repeated, while the other symbols stay fixed.

Example: $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$ is a cycle since we have $1 \mapsto 3 \mapsto 2 \mapsto 1$.

For short such a cycle is written $\begin{pmatrix} 1 & 3 & 2 \end{pmatrix}$, again meaning $1 \mapsto 3 \mapsto 2 \mapsto 1$.

On five symbols, one cycle is $\begin{pmatrix} 1 & 4 & 2 & 5 \end{pmatrix}$, meaning $1 \mapsto 4 \mapsto 2 \mapsto 5 \mapsto 1$ and $3 \mapsto 3$. A cycle of length k is called a k-cycle. A 2-cycle is called a **transposition**.

Proposition. Any permutation on n symbols can be written as a product of disjoint (non-overlapping) cycles. For example, one permutation in S_4 is $(13)(24)$.

Since disjoint cycles commute, the order doesn't make a difference.

Problem V-12. Do (12) and (23) commute?

---

[1] A *group* is a set with an operation obeying some familiar rules; in the case of permutations the operation is multiplication as defined below.

[2] Notice that this means following the effect of the permutations on each symbol while moving right to left. Some texts like to multiply permutations following symbols while moving left to right.

Problem V-13. Write $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 4 & 3 & 7 & 6 & 5 & 1 \end{pmatrix}$ as a product of disjoint cycles. (Omit 1-cycles.)

Problem V-14. Write out the multiplication table of S_3 using cycle notation. The elements are **1** (meaning the identity transformation), 3-cycles, and 2-cycles (transpositions).

3. Equivalence relations

Definition. A **binary relation** on a set X is any rule by which some pairs are declared to be related and others not related[3]. Some examples for $X = \mathbb{R}$ are (i) the relation $<$, (ii) the relation $\geq$, (iii) the equality relation (where elements are related only to themselves), and (iv) the $\neq$ relation. If the name of the relation is ρ, we can write $x \rho y$ to mean x and y are related by ρ.

Some useful properties that a relation ρ can have are as follows. (As usual, it is understood that they mean for all x, or for all x, y, etc.)

(1) $x \rho x$	"ρ is reflexive"
(2) $x \rho y \Rightarrow y \rho x$	"ρ is symmetric"
(3) $x \rho y$ and $y \rho z \Rightarrow x \rho z$	"ρ is transitive"

Problem V-15. For each of examples (i) through (iv), which of properties (1), (2), (3) hold?

Definition. A relation ρ is an **equivalence relation** if all three of (1), (2), (3) hold.

Problem V-16. Which of the following examples (i)-(xi) are equivalence relations on X? (Explain why or why not.)

(v) $X = \mathbb{Z}$ and $x \rho y$ means $x - y$ is even.

(vi) X is any set, Y is any set, $f : X \to Y$ is any function, and $x_1 \rho x_2$ means $f(x_1) = f(x_2)$.

(vii) $X = \mathbb{R}^2$, there is a subspace W of X, and $x \rho y$ means $x - y \in W$.

(viii) $X = \text{Mats}(m \times n)$ and $A \rho B$ means that A is row-equivalent to B; in other words, A and B have the same row space, which is the same as saying they have the same row-reduced echelon form.

(ix) $X = \text{Mats}(n \times n)$ (square!) and ρ is "similarity" $\sim$, where $A \sim B$ means $P^{-1}AP = B$ for some invertible $n \times n$ matrix P

[3]Technically a relation is just a set of some pairs from X, meaning the set of pairs of elements that are declared to be related. In other words, $\rho \subseteq X \times X$.

(x) X is any set and every two elements are related by ρ.

(xi) X is any set, subsets B_i partition X, and $x\rho y$ means that x and y are in the same block.

Problem V-17. What is wrong with the following "proof" that a symmetric, transitive relation is reflexive? "If $x\rho y$ then also $y\rho x$ by symmetry, and then $x\rho y$ and $y\rho x$ together imply $x\rho x$ by transitivity".

The real purpose of equivalence relations is to be able to describe partitions easily, as in (xi) above:

Theorem. Suppose ρ is an equivalence relation on a set X. For each $x \in X$, define the block of x to be $B_x = \{y \in X | x\rho y\}$. Then (i) the distinct blocks of this kind form a partition of X, and (ii) the equivalence relation derived from this partition as in (xi) above is ρ.

Thus each equivalence relation corresponds to a partition and vice-versa. Notice that in the theorem, two different elements may give the same block, but we use just one copy of each block.

A common terminology is to call the blocks "equivalence classes", but this conflicts with the word "classes" as it is used in set theory.

Problem V-18. Prove the theorem.

Problem V-19. Where possible, tie the examples of equivalence relations given above to the examples of partitions from Problem Q-2.

Problem V-20. Prove that for any integer $n > 0$, "congruence modulo n" is an equivalence relation. Specifically, $x \equiv y \pmod{n}$ means that $n|(y - x)$ (i.e., n divides $y - x$, meaning that there is some $q \in \mathbb{Z}$ with $y - x = qn$). Use any reasonable properties of integer division. Remember, you can divide *into* 0 but not *by* 0.