Math Club Notes: 2006 May 16

[Actually these are notes for the May 9 meeting too.]

1 A broken probability argument

Stuart and I discussed problem A6 from the 2005 Putnam. You can find it (along with a solution) at http://www.unl.edu/amc/a-activities/a7-problems/putnam/. Our discussion produced the following question.

Toss 3 fair coins; what is the probability that they come up 2 one way and 1 the other?

It is easy to make a correct argument here. For example: note that the only other possibility is that all 3 coins come up the same, that is, all heads or all tails; so we get 2 coins one way and 1 the other in all but 2 cases, out of 2^3 possibilities total. So the probability is $1 - \frac{2}{8} = \frac{3}{4}$.

Of more interest is the following incorrect argument. If the coins come up 2 one way and 1 the other, one of the coins is the odd one out. Each of the other two coins must be of the opposite type (heads if the odd-one-out is tails, and vice versa). Each of those other two coins has a probability of $\frac{1}{2}$ of coming up opposite to the odd-one-out, so the probability of them both coming up right is $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$.

This argument is clearly wrong... but what exactly is the problem with it?

2 Rusin's exam question about matrix traces

The question was:

> Use the inner product $\langle A, B\rangle = \operatorname{tr}(A^T B)$ and the identity matrix for B to relate $\operatorname{tr}(A)$ and $\operatorname{tr}(A^2)$ when A is symmetric.

Some things Rusin didn't bother to say (since he was speaking in sci.math, hence with an audience of professional mathematicians who don't need this kind of thing said): Fix a positive integer n. The set of $n \times n$ matrices then forms a vector space under the usual operations of matrix addition and scalar multiplication. (It also has an operation of matrix multiplication; we don't need that for a vector space.) Rusin is saying that the function $\langle \cdot, \cdot \rangle$ defined by

$$\langle A, B\rangle = \operatorname{tr}(A^T B)$$

is an inner product on this space, that is, a real-valued binary function having the properties that, for any $n \times n$ matrices A, B, and C, and any scalar α,

$$\langle A, B\rangle = \langle B, A\rangle$$
$$\langle A, B + C\rangle = \langle A, B\rangle + \langle A, C\rangle$$
$$\langle \alpha A, B\rangle = \alpha\langle A, B\rangle$$
$$\langle A, A\rangle \geq 0 \text{ , and } \langle A, A\rangle = 0 \text{ iff } A = O$$

If you weren't familiar with this inner product (as I wasn't), it's a worthwhile exercise to verify that it has these properties. (You might want to brush up on the properties of the matrix trace function.)

One of the important things about inner products is that they induce norms; given an inner product $\langle\cdot,\cdot\rangle$ we can define a norm function $\|\cdot\|$ by

$$\|A\| = \sqrt{\langle A, A\rangle} \ .$$

A norm defined this way has all the usual properties; of particular relevance to us is the Cauchy-Schwarz inequality:

$$\langle A, B\rangle^2 \leq \|A\|^2\|B\|^2 \ .$$

(We looked at this inequality briefly once or twice before, in the more familiar context of the usual dot product on $\mathbb{R}^n$; see our notes for 2005 May 16 and May 22. Point is, it can be proven using only the properties of inner products, so it holds for any inner product, not just the dot product.)

Now for Rusin's question. He tells us what to do — take the identity matrix for B. Doing that, we get

$$\langle A, I\rangle = \operatorname{tr}(A^T I) = \operatorname{tr}(A^T) = \operatorname{tr}(A) \ .$$

(The last step is a property of the trace; besides, we're told A is symmetric, so $A^T = A$.)

$\operatorname{tr}(A)$ has appeared. We want to relate it (somehow) to $\operatorname{tr}(A^2)$. How can we make $\operatorname{tr}(A^2)$ appear? We know that $A^T = A$; it's fairly easy to spot the fact that

$$\operatorname{tr}(A^2) = \operatorname{tr}(A^T A) = \langle A, A\rangle \ .$$

So, in terms of $\langle\cdot,\cdot\rangle$, we're supposed to relate $\langle A, I\rangle$ and $\langle A, A\rangle$. Cue the Cauchy-Schwarz inequality:

$$\langle A, I\rangle^2 \leq \|A\|^2\|I\|^2 = \langle A, A\rangle\langle I, I\rangle \ .$$

Putting it all together, and computing that $\langle I, I\rangle = \operatorname{tr}(I^T I) = \operatorname{tr}(I) = n$, we obtain the following relationship:

$$(\operatorname{tr}(A))^2 \leq n \operatorname{tr}(A^2) \ .$$

A nifty fact, even if it only holds for symmetric A.

3 Counting beats case analysis

Vish and I looked at three examples of problems which can be solved both by counting and (at least in principle) by case analysis. Counting makes for much more appealing solutions, though it can require some cleverness.

3.1 Girls and boys at a circular table

The problem:

> 25 boys and 25 girls are sitting around a circular table. Show that there is a person at the table both of whose neighbours are girls.[1]

Label the seats $1, 2, \ldots, 50$. Split them into two groups: the odd-numbered and the even-numbered, that is, the seats $1, 3, 5, \ldots, 49$ and the seats $2, 4, 6, \ldots, 50$. Since there are 25 girls, one of these groups must have at least 13. Without loss of generality, the even-numbered seats have at least 13 girls. (If it's the odd-numbered seats, renumber them, starting one seat over.)

So there are 13 girls among the 25 seats numbered $2, 4, 6, \ldots, 50$. Now, each odd-numbered seat has either 0, 1, or 2 girl neighbours. Go through the odd-numbered seats and add up the number of girl neighbours each has. This procedure counts every girl in the even-numbered seats exactly twice — once as a neighbour of the seat to her left, and once as a neighbour of the seat to her right. So the total is $2 \times 13 = 26$. There are only 25 odd-numbered seats, so at least one of them has more than one girl neighbour, which concludes the proof.

Exercise: Is it possible to arrange 24 girls and 26 boys around a circular table so that nobody has two girls as neighbours? (In other words, is the result given in the problem the best possible result?)

3.2 Bites out of a chessboard

A very well-known puzzle.

Take an 8×8 grid of squares (like a chessboard) and remove the upper-left and lower-right corner squares. Is it possible to cover the rest of the board with 2×1 dominos? (Of course, we mean without overlapping.)

Turns out no. Colour the squares as they are in a chessboard. A domino in any placement and orientation covers one black square and one white square. But the two squares removed are the same colour, say, white; so there are 32 black squares and 30 white squares to be covered.

[1] Problem #6 from Titu Andreescu and Zuming Feng, *102 Combinatorial Problems*, Birkhäuser, 2002.

3.3 Monochromatic triangles

The problem: Take six points. Join every pair of points with a line segment. Colour each line segment either red or blue. A "monochromatic triangle" in the resulting figure is a set of three points whose three connecting line segments are all the same colour. Prove that there are at least two monochromatic triangles.

Proving there's at least one is fairly straightforward, even by case analysis. (It starts off like this: Pick three points at random. If they're all the same colour, we're done. If not, they're two of one colour and one of the other, say, two red and one blue. Take a fourth point, and consider the edges joining it to the three we already have. Etc.)

It's possible to prove there's at least two by similar case analysis, but it's a pain. Happily, there's a much nicer proof by counting, which I learned about in Dijkstra: see http://www.cs.utexas.edu/users/EWD/ewd07xx/EWD771.PDF. (I don't think the argument is original to Dijkstra.) Also see his follow-up, twenty years later, on upper bounds for the number of such triangles: http://www.cs.utexas.edu/users/EWD/ewd13xx/EWD1302.PDF.

4 Questions from a podcast

Vish brought a couple fun little problems from a podcast he listens to:

1. You have a stick which is 12 inches long. Put some marks on it so that you can measure $1, 2, \ldots, 12$ inches by measuring from mark to mark in various ways (and from marks to the end of the stick). E.g., with marks at 3 and 7, you could measure $3, 4, 5, 7, 9$. It's possible with 4 marks, in two ways. (Four ways, if you count reflections.) It's not possible with fewer. (Proof?)

2. Consider the following silly game. You have a pile of cards; so does your opponent. Each card has a distinct number on it. You and your opponent each draw a card from your respective piles at random. The higher card wins. Now, say that pile A "beats" pile B if the expected outcome of this game is for the owner of pile A to win. The question: can we distribute cards labelled $1, 2, \ldots, 9$ into three piles A, B, C of three cards each, such that A beats B, B beats C, and C beats A?