Well, this week's problems were kind of a dud because I didn't realize how many of you had not had a linear algebra course. It's a little hard to explain what the heck I'm doing in some of the problems to someone without much background, so just take this as an incentive to go out and learn more linear algebra!

I'm still in the market for pretty answers to problems 1, 3, and 4.

1. Suppose $A, B \in M_4(\mathbf{R})$ commute, and $\det(A^2 + AB + B^2) = 0$. Prove that

$$\det(A+B) + 3\det(A-B) = 6\det(A) + 6\det(B).$$

ANSWER: I'm not sure how to do this. Roughly speaking, commuting matrices are (generically) simultaneously diagonalizable, so you should think about a (complex) basis of eigenvalues; if $Av = \lambda v$ then we expect $Bv = \omega\lambda v$ where ω is a root of $X^2 + X + 1$, i.e. a cube root of unit. But I'm not sure where to go from there.

(Perhaps more concretely: let v be in the kernel of the singular matrix A^2+AB+B^2; then $0 = (A-B)(A^2+AB+B^2)v = (A^3-B^3)v$, so that A^3v and B^3v are the same vector. But I'm not sure what that gets us, and I feel like I've already used all the information given! It IS relevant that these matrices are both 4×4 and real, because the real eigenspaces are going to have to be 2-dimensional.)

2. (10B1) Is there an infinite sequence of real numbers $a_1, a_2, a_3, \ldots$ such that

$$a_1^m + a_2^m + a_3^m + \cdots = m$$

for every positive integer m?

ANSWER: Apply the Cauchy-Schwarz Inequality to the vectors $v_k = (a_1^k, a_2^k, \ldots)$ and $v_n = (a_1^n, a_2^n, \ldots)$ to conclude $(\sum a_i^{k+n})^2 \leq (\sum a_i^{2k})(\sum a_i^{2n})$, which from the assumed equations would mean $(k+n)^2 \leq 4kn$, i.e. $(k-n) \leq 0$. So we get a contradiction just from any two (distinct) equations in the list!

3. (95A5) Let $x_1, x_2, \ldots, x_n$ be differentiable (real-valued) functions of a single variable f which satisfy

$$\begin{aligned}
\frac{dx_1}{dt} &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\
\frac{dx_2}{dt} &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\
&\cdots \\
\frac{dx_n}{dt} &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n
\end{aligned}$$

for some constants $a_{ij} > 0$. Suppose that for all i, $x_i(t) \to 0$ as $t \to \infty$. Are the functions $x_1, x_2, \ldots, x_n$ necessarily linearly dependent?

4. (95A6) Suppose that each of n people writes down the numbers 1,2,3 in random order in one column of a $3 \times n$ matrix, with all orders equally likely and with the orders for different columns independent of each other. Let the row sums a, b, c of the resulting matrix be rearranged (if necessary) so that $a \leq b \leq c$. Show that for some $n \geq 1995$, it is at least four times as likely that both $b = a + 1$ and $c = a + 2$ as that $a = b = c$.

ANSWER: I don't recall an aswer but I have recently taught probability, so let me show how one might compute the probabilities of these events precisely. (This is surely NOT the way to answer the Putnam question!)

There are six possible columns, "123", "132", etc.; let C be the set of all six of them. Then the outcome of this experiment will be a sequence s of n elements of C, i.e. an element of C^n. My reading of the problem is that all the elements of this sample space are equally likely to occur, so the probability of any event E (that is, any subset of the sample space) is the cardinality $|E|$ of that subset divided by $|C^n| = 6^n$.

We can count how many of each column occur in a particular sequence s, giving a function $F : C^n \to \mathbf{N}^6$ (where $\mathbf{N}$ is the set of natural numbers). Then we can compute the row sums of s from $F(s)$: if s contains n_{123} occurences of the column 123, n_{132} occurences of 132, etc., then the top row-sum is

$$1 \cdot n_{123} + 1 \cdot n_{132} + 2 \cdot n_{213} + 2 \cdot n_{231} + 3 \cdot n_{312} + 3 \cdot n_{321}$$

and similarly for the middle and bottom row-sums. We can assemble the three row-sums into a vector, giving another function $G : \mathbf{N}^6 \to \mathbf{N}^3$. Note that the sum of the three row-sums will be $6(\sum n_{ijk}) = 6n$.

So what are the events whose probabilities we wish to compute? In the one case we want all the sequences with $G(F(s)) = (2n, 2n, 2n)$; in the other we want the sequences with $G(F(s))$ being one of the six permutations of $(2n - 1, 2n, 2n + 1)$.

Very well then, when is $G(F(s)) = (2n, 2n, 2n)$? We need the three row-sums to be equal; that will give two homogeneous equations in the six unknowns n_{ijk} whose solution set is then a four-dimensional subspace of $\mathbf{N}^6$. This subspace obviously includes some simple sums of basis elements of $\mathbf{N}^6$: $e_{123} + e_{321}$, $e_{213} + e_{231}$, $e_{132} + e_{312}$ as well as $e_{123} + e_{231} + e_{312}$ and $e_{321} + e_{132} + e_{213}$. The first three and the last pair have equals sums, so the set of five is linearly dependent, but any four are independent, and hence span the solution space. In other words: a matrix with three equal row-sums must consist of some linear combination (say) m_1 pairs of columns 123 and 321, plus m_2 pairs of a 213 and a 231, plus m_3 pairs 132 and 312, plus m_4 sets of a 123 and a 231 and a 312. Or, said differently, the 123, 132, 213, 231, 312, 321 components of $F(s)$ must respectively be $(m_1+m_4, m_3, m_2, m_2+m_4, m_3+m_4, m_1)$. Since each of these is a non-negative integer, the m_i are all integers, with $m_1, m_2, m_3 \geq 0$ and with $m_4 \geq -m_1, -m_2, -m_3$. The condition that the six components of $F(s)$ sum to n forces $2(m_1 + m_2 + m_3) + 3m_4 = n$, from which m_4 may be computed (and from which we find the constraint $m_1 + m_2 + m_3 \equiv -n \pmod 3$.) To summarize: we may characterize all sequences $s \in C^n$ whose three row-sums are equal as those for which the vector $F(s)$ is of the form

$$(n + m_1 - 2m_2 - 2m_3)/3, m_3, m_2, (n - 2m_1 + m_2 - 2m_3)/3, (n - 2m_1 - 2m_2 + m_3)/3, m_1)$$

for some non-negative integers m_i having $m_1+m_2+m_3 \equiv -n \pmod 3$ and $-m_1+2m_2+2m_3 \le n, 2m_1-m_2+2m_3 \le n, 2m_1+2m_2-m_3 \le n$. (The domain is a double-tetrahedron: the triangle spanned by the points with $m_i = n/2, m_{i+1} = m_{i-1} = 0$ is extended on the one side to the origin and on the other side to the point $m_1 = m_2 = m_3 = n/3$.)

To recap, here is how we compute the probability that the three row sums are equal. For each m_1, m_2, m_3 meeting the conditions above, compute (m_4 and then) the 6-tuple of n_{ijk}'s above. The number of sequences having exactly these counts of the six possible columns is the multinomial coefficient $\binom{6n}{n_{123}, n_{132}, \dots} = (6n)!/(n_{123}!n_{132}!\dots)$

Here are the calculations in the (small!) case $n = 12$. There are 50 eligible triples $(m_1, m_2, m_3) = (0,0,0), (3,0,0), (2,1,0), (1,1,1), \dots (3,3,3), (2,3,4), (1,4,4), (4,4,4)$. For each, we compute the six numbers n_{ijk}:

$$(4,0,0,4,4,0), (5,0,0,2,2,3), (4,0,1,3,2,2), (4,1,0,2,3,2), \dots, (0,4,4,0,0,4)$$

For each of these we count the 12-tuples with these characteristics, as multinomial coefficients: $34650, 166320, 831600, 831600, \dots 34650$. This gives a grand total 47977776 and a probability of 333179/15116544= 0.02204 .

I suppose a less-precise form of this analysis can be used to *estimate* the probabilities sufficiently well to answer the Putnam question.

5. Suppose $A \in M_n(\mathbf{C})$ has rank r, where $1 \le r \le n-1$ and $n > 1$. Show that there exist matrices $B \in M_{n,r}(\mathbf{C})$ and $C \in M_{r,n}(\mathbf{C})$ with $A = BC$.

ANSWER: Geometrically this simply asserts that a linear map from $\mathbf{C}^n$ to itself may be written as a composite of a projection and an injection. If r is the rank of A then its kernel $V = \ker(A)$ has dimension $n-r$, and then the orthogonal complement $V^\perp$ of V has dimension r. Pick a basis $v_1, \dots, v_r$ of $V^\perp$ and let C be the matrix that has these v_i as its rows. Then C has the same kernel as A. Since CC^T is positive definite, it is invertible, so let $B = AC^T(CC^T)^{-1}$. Then $BCC^T = AC^T$, i.e. $(BC-A)C^T = 0$. Thus $(BC-A)v$ vanishes for any v in the span of the columns of C^T (i.e. in the span of the rows of the C, which are the v_i). But this means $BC - A$ annihilates all of $V^\perp$, and as earlier noted it also annihilates all of V, so it annihilates all of $\mathbf{C}^n$, i.e. $A = BC$.

6. (Problem 2008-A-2). Alan and Barbara play a game in which they take turns filling entries of an initially empty 2008×2008 array. Alan plays first. At each turn, a player chooses a real number and places it in a vacant entry. The game ends when all the entries are filled. Alan wins if the determinant of the resulting matrix is nonzero; Barbara wins if its is zero. Which player has a winning strategy?

ANSWER: Barbara can win by duplicating Alan's move in an adjoining row, so as to make the 1st and 2nd rows identical, as well as the third and fourth, etc. (Barbara will end up creating a matrix of rank only 1004!) A similar analysis works whenever the number of rows is even. I don't know who has a winning strategy even when $n = 3$.

7. (1990-A-5). If A and B are square matrices of the same size such that $ABAB = 0$, does it follow that $BABA = 0$?

ANSWER: No. This is a little tricky because there are many theorems that say, "If (blah blah) and $(AB)^2 = 0$ then $(BA)^2 = 0$." Here the conditions might be: A is invertible; B is invertible; A and B commute; A or B has rank 1; $AB = 0$; $A, B \in M_2(\mathbf{C})$; A, B symmetric; etc. So we can construct an example but have to avoid these conditions. Also the import of the theorem is unchanged under change of basis and under scalar multiplication, so we can assume at least A is of a fairly simple form.

So let A be a 3×3 matrix $\begin{pmatrix} I_2 & 0 \\ 0 & 0 \end{pmatrix}$, and assume $B = \begin{pmatrix} M & v \\ w^t & c \end{pmatrix}$, where $v, w \in \mathbf{R}^2$. Then it's easy to compute $(AB)^2 = \begin{pmatrix} M^2 & Mv \\ 0 & 0 \end{pmatrix}$ and $(BA)^2 = \begin{pmatrix} M^2 & 0 \\ w^t M & 0 \end{pmatrix}$. So we need to choose M, v, w so that $M^2 = 0$ and v lies in its kernel but w' does not. A combination that works is $M = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $v = (1, 0)^t$, $w^t = (0, 1)$.

8. (1994-A-4). Let A and B be 2×2 matrices with integer entries such that A, $A + B$, $A + 2B$, $A + 3B$, and $A + 4B$ are all invertible matrices whose inverses have integer entries. Show that $A + 5B$ is invertible and that its inverse has integer entries.

ANSWER: If M and N are matrices with integer entries, their determinants are integers. If $MN = I$ then $\det(M) \cdot \det(N) = 1$ so $\det(M)$ must be ± 1. So each of $\det(A + kB)$ (for $k = 0, 1, 2, 3, 4$) is either $+1$ or -1. By the Pigeonhole Principle, at least three of these values must agree. On the other hand, $\det(A + xB)$ is a quadratic polynomial in x, and hence can only achieve any particular value at most twice as x varies, unless it's actually constant, so in our case $\det(A + 5B)$ will again be ± 1, making the matrix invertible with integer inverse.

Now flip over for some additional practice with Axiomatic Mathematics!

BONUS ROUND! A vector space may be defined as a set V on which two binary operations called $+$ and $\cdot$ are defined (respectively as functions $V \times V \to V$ and $\mathbf{R} \times V \to V$) subject to a set of axioms. We may express these axioms in the following way:

$\texttt{VS}_1$. For all $u, v, w \in V$ we have $u + (v + w) = (u + v) + w$

$\texttt{VS}_2$. For all $u, v \in V$ we have $u + v = v + u$

$\texttt{VS}_3$. There is a vector $u \in V$ so that for all $v \in V$ we have $v + u = v$

$\texttt{VS}_4$. For all $u, v, w \in V$, if $u + w = v + w$ then $u = v$; likewise if $w + u = w + v$ then $u = v$.

$\texttt{VS}_5$. For all $u, v \in V$ and all $a \in \mathbf{R}$ we have $a \cdot (u + v) = a \cdot u + a \cdot v$

$\texttt{VS}_6$. For all $u \in V$ and all $a, b \in \mathbf{R}$ we have $(a + b) \cdot u = a \cdot u + b \cdot u$

$\texttt{VS}_7$. For all $u \in V$ and all $a, b \in \mathbf{R}$ we have $(ab) \cdot u = a \cdot (b \cdot u)$

$\texttt{VS}_8$. For all $u \in V$ we have $1 \cdot u = u$

The challenge I posed was this: *For each of these axioms, give an example of an object which satisfies all the axioms EXCEPT the given one, that is, a non-vector-space that satisfies the other seven axioms.*

I gave a warning: *This can be done for seven of the axioms, but one of these axioms is actually redundant — it automatically follows from the other seven axioms. Which of the eight axioms is redundant?*

Turns out that's not *quite* right, because I phrased the axioms just a little differently from the way I remembered. But here are some of the not-quite-vector-spaces I had in mind.

First an example of an impostor that meets all the axioms except $\texttt{VS}_1$. Let $V = \mathbf{R}^2$, on which we use ordinary scalar multiplication. But define the addition of two vectors u, v to be $u + v = |\cos(\theta)|(u + v)$, where θ is the angle between u and v; if either of u or v is zero, take θ to be zero.

Here's an example where only $\texttt{VS}_4$ fails. It's essentially "the real numbers with a second zero". Let $V = \mathbf{R} \cup \{z\}$ where z is anything outside $\mathbf{R}$. Define "+" so that its restriction to real numbers is ordinary addition, and then define $r + z = z + r = r$ for all real numbers r; only $z + z = z$ gives a sum not in $\mathbf{R}$. So this z (and not the number $0 \in \mathbf{R}$) is the "zero element" u that is called for by axiom $\texttt{VS}_3$. Next define the scalar multiplication to be ordinary multiplication on $\mathbf{R}$, and define $a \cdot z = z$ for all reals a. I leave it to you to check all the other axioms hold, but $\texttt{VS}_4$ is not, because $0 + z = 0 + 0$ even though $z \neq 0$.

Skipping $\texttt{VS}_5$ is tricky. It's easier to violate this axiom in the complex case: Let $V = \mathbf{C} \times \mathbf{C}$ with usual vector addition but to define scalar multiplication, let σ be any nontrivial automorphism of the complex field (e.g. complex conjugation) and then let

$$a \cdot (x, y) = \begin{cases} (a\,x, a\,y) & \text{if } x \neq 0 \\ (0, \sigma(a)\,y) & \text{if } x = 0. \end{cases}$$

We can use the same trick over any other field, except that $\mathbf{R}$ has no nontrivial automorphisms! And if we just restrict the previous example to the reals inside $\mathbf{C}$ if σ is complex conjugation, then axiom $\texttt{VS}_5$ *will* hold. But we can use the fact that $\mathbf{C}$ has automorphisms

which do not preserve $\mathbf{R}$(not even setwise), and then use the above formulas to define a scalar multiplication on the real vector space $\mathbf{C} \times \mathbf{C}$ (or if you prefer, on $\mathbf{R} \times \mathbf{C} = \mathbf{R}^3$.)

Here's an example in which all axioms hold except $\mathtt{VS}_6$. In Group Theory language, this example is the group $\mathbf{R} \times \mathbf{Z}_2$: as a set, V is the collection of ordered pairs (x, y) where x is any real number but $y = \pm 1$ only. "Addition" is defined by $(x, y) + (z, w) = (x + z, yq)$ and "scalar multiplication" is defined by $a \cdot (x, y) = (ax, y)$. I will leave it to you to check that the other 7 axioms hold but as for $\mathtt{VS}_6$: $(a + b) \cdot (x, -1) = ((a + b)x, -1) = (ax + bx, -1)$ but since $a \cdot (x, -1) = (ax, -1)$ and $b \cdot (x, -1) = (bx, -1)$, their sum is $a \cdot (x, -1) + b \cdot (x, -1) = (ax + bx, +1)$. (A second example uses $V = \mathbf{R}$ with the usual addition but with scalar multiplication being $a \cdot v = a^2 v$.)

My example of a system lacking $\mathtt{VS}_7$ is a bit subtle. Let $V = \mathbf{R}$, with vector addition being the usual addition, but define scalar multiplication by: $a \cdot v = \phi(a)\, v$ where $\phi : \mathbf{R} \to \mathbf{Q}$ is any homomorphism of (additive) groups. You may find this unsatisfying because we know such a ϕ exists but we can't write one down!

I already gave an example lacking $\mathtt{VS}_8$ when I posed the problem: we take the set V to be the set of real numbers; define "vector addition" on V to be ordinary addition of real numbers; and define "scalar multiplication" by: $c \cdot v = 0$ for all scalars c and vectors v. Then axioms $\mathtt{VS}_1$ through $\mathtt{VS}_7$ are satisfied but axiom $\mathtt{VS}_8$ is not.

Now what about the other two axioms? As Dylan noted, $(1 + 1) \cdot (u + v)$ may be expanded in two ways using axioms 5 and 6 in different orders: on the one hand it's

$$(1 + 1) \cdot u + (1 + 1) \cdot v = (1 \cdot u + 1 \cdot u) + (1 \cdot v + 1 \cdot v)$$

and on the other hand it's

$$1 \cdot (u + v) + 1 \cdot (u + v) = (1 \cdot u + 1 \cdot v) + (1 \cdot u + 1 \cdot v)$$

Using axioms 1 and 8 to remove the parentheses and the 1's, we conclude $u + u + v + v = u + v + u + v$. Then using (both parts of) axiom 4 we may cancel a u from the left end of both sides and then a v from the right end, giving $u + v = v + u$. That means axiom 2 automatically holds in any situation where axioms 1,4,5,6, and 8 hold. (Axioms 3 and 7 were never used here.)

(In Group Theory, this proof is given to show that a group in which squaring is a homomorphism is necessarily abelian; in particular, a group in which every element has order 2 must be abelian.)

So there is no example in which all axioms hold except axiom 2.

There's also no example violating only axiom $\mathtt{VS}_3$; in fact you may have noticed while working in Linear Algebra that you needn't postulate the existence of a zero element because you can compute it, as $0 \cdot v$ for any vector v. The trick is proving that all these 0-scalar-multiples are the same element of V.

Well, for any $v, w \in V$, first use axioms $\mathtt{VS}_1$ and $\mathtt{VS}_6$ to show

$$0 \cdot w + (0 \cdot w + 0 \cdot v) = (0 \cdot w + 0 \cdot w) + 0 \cdot v = (0 + 0) \cdot w + 0 \cdot v = 0 \cdot w + 0 \cdot v$$

By symmetry (and $\texttt{VS}_2$) this is then equal also to $0 \cdot v + (0 \cdot w + 0 \cdot v)$. By $\texttt{VS}_4$ we conclude $0 \cdot w = 0 \cdot v$ as I suggested. So now let's verify that axiom $\texttt{VS}_3$ is satisfied; we will define our z by picking any $v \in V$ and defining $z = 0 \cdot v$, but we have now proven that this same z can also be written as $0 \cdot w$ for any other $w \in V$. So then for any $w \in V$ we have $w + z = w + 0 \cdot v = 1 \cdot w + 0 \cdot w = (1 + 0) \cdot w = w$. by $\texttt{VS}_8$. Thus this z really does satisfy axiom $\texttt{VS}_3$.

The reason I claimed only one axiom was dependent on the others is that one can streamline the axioms a bit. The definition of a vector space can also be given with axiom $\texttt{VS}_4$ replaced by

$$\texttt{VS}'_4\text{: For all } u, v, w \in V, \text{ if } u + w = v + w \text{ then } u = v$$

That is, we have only a right-cancellation property. In order to prove $\texttt{VS}_2$, Dylan used both cancellation properties implicit in $\texttt{VS}_4$. If instead we have available only this weaker axiom $\texttt{VS}'_4$, no such proof can exist, because here is a model of something satisfying all the axioms except $\texttt{VS}_2$ (including $\texttt{VS}'_4$ but not $\texttt{VS}_4$): Let $V = \mathbf{R}$, and define the vector addition by: $u + w = u$ for any two vectors u and v, and (likewise) define $a \cdot u = u$ for any vector u and real number a. Then axiom $\texttt{VS}_2$ fails pretty spectacularly but you can check that the other 7 axioms hold. (In particular, $u + w = v + w$ does indeed imply $u = v$, although from $w + u = w + v$ we can conclude nothing.)

If you find any examples that are simpler than these (or more enlightening!) then I would be happy to learn of them. In the mean time I recommend that whenever you find yourself teaching linear algebra, you omit axioms 3 and 4 altogether and simply add a

$$\texttt{VS}_9\text{: For all } u, v \in V,\ 0 \cdot u = 0 \cdot v$$

Then this common element of V becomes the zero vector ($\texttt{VS}_3$), $(-1) \cdot v$ becomes the inverse of v, and with inverses for each element we get cancellation properties ($\texttt{VS}_4$).