Department of Applied Mathematics and Statistics
The Johns Hopkins University

INTRODUCTORY EXAMINATION–SPRING SESSION

Tuesday, January 24, 2006

Instructions: Read carefully!

1. This **closed-book** examination consists of 20 problems (sorry, no choices), each worth 5 points. The passing grade has been set at $66\frac{2}{3}\%$. Partial credit will be given as appropriate; each part of a problem will be given the same weight. If you are unable to prove a result asserted in one part of a problem, you may still use that result to help in answering a later part.

2. You have been provided with a syllabus indicating the scope of the exam. Our purpose is to test not only your knowledge, but also your ability to apply that knowledge, and to provide mathematical arguments presented in **clear, logically justified steps**. The grading will reflect that broader purpose.

3. The problems have not been grouped by topic, but there are roughly equally many mainly motivated by each of the four areas identified in the syllabus (linear algebra; real analysis; probability; discrete mathematics and operations research/optimization). Nor have the problems been arranged systematically by difficulty. If a problem directs you to use a particular method of analysis, you *must* use it in order to receive substantial credit.

4. Start your answer to each problem on a NEW sheet of paper. Write only on ONE SIDE of each sheet, and please do not write very near the margins on any sheet. Arrange the sheets in order, and write your NAME and the PROBLEM NUMBER on each sheet.

5. The examination will begin at 8:30 AM; lunch and refreshments will be provided. The exam will end just before 5:00 PM. You may leave before then, but in that case you may not return.

6. Paper will be provided, but you should bring and use writing instruments that yield marks dark enough to be read easily.

7. **No calculators of any sort are needed or permitted.**

1. Let $S = \{v_1, v_2, v_3\}$ be a basis for $\mathbb{R}^3$, where

$$v_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 3 \\ 3 \\ 4 \end{bmatrix}, \quad \text{and} \quad v_3 = \begin{bmatrix} 2 \\ 9 \\ 0 \end{bmatrix}.$$

(a) Find the coordinate vector of $v = \left[\begin{smallmatrix} 1 \\ 1 \\ 1 \end{smallmatrix}\right]$ with respect to S.

(b) Find the vector v in $\mathbb{R}^3$ whose coordinate vector with respect to S is $[v]_S = \left[\begin{smallmatrix} 1 \\ 1 \\ 1 \end{smallmatrix}\right]$.

Solution: (a) We must find scalars c_1, c_2, c_3 such that $v = c_1v_1 + c_2v_2 + c_3v_3$. We do this by row reducing the matrix

$$\begin{bmatrix} 1 & 3 & 2 & 1 \\ 2 & 3 & 9 & 1 \\ 1 & 4 & 0 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 0 & -7 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

so $[v]_S = \left[\begin{smallmatrix} -7 \\ 2 \\ 1 \end{smallmatrix}\right]$. (b) $v = v_1 + v_2 + v_3 = [6, 14, 5]^T$.

---

2. Estimate the probability that in 60 independent tosses of a pair of fair dice the sum is never equal to four. Use the Poisson approximation to the Binomial.

Solution: Let $X_n = 1$ if the sum obtained on the n'th toss is four and let $X_n = 0$ otherwise, $n = 1, 2, \ldots, 60$. Then $X_1, \ldots, X_{60}$ are independent Bernoulli random variables with $P(X_n = 1) = \frac{3}{36} = \frac{1}{12}$ for each n and therefore $S \doteq \sum_{n=1}^{60} X_n$ is Binomial with sample size 60 and $p = \frac{1}{12}$. By the Poisson approximation, the distribution of S is approximately Poisson with mean $\lambda = \frac{60}{12} = 5$. Consequently,

$$P(S = 0) \approx e^{-5}.$$

---

3. Let $f, g : \mathbb{R} \to \mathbb{R}$ be continuously differentiable and let $a < b$ be reals. Suppose that $fg' - gf'$ never vanishes on the interval $[a, b]$ and that $f(a) = f(b) = 0$. Prove that g has a root in the interval $[a, b]$.

Solution: Suppose, for contradiction, that $g \neq 0$ on $[a, b]$. The hypothesis implies that $h = f/g$ is a well-defined function on $[a, b]$ with $h(a) = h(b) = 0$. By Rolle's theorem (or the mean value theorem) there is a $c \in [a, b]$ such that $h'(c) = 0$. But $h'(c) = [f'(c)g(c) - f(c)g'(c)]/[g(c)]^2$ which contradicts the fact that $f'(c)g(c) - f(c)g'(c) \neq 0$.

4. Let J be a 100×100 matrix all of whose entries are equal to 1. Find (with justification) the rank, nullity (dimension of the null space), determinant, and eigenvalues (with multiplicities) of J.

Solution: The columns are identical and nonzero, so the rank is 1. By the Rank-Nullity Theorem, the nullity of this matrix is $100-1=99$. Since the matrix does not have full rank, its determinant is 0.

Since the nullity is 99, 0 is an eigenvalue with multiplicity 99. The all ones vector is clearly an eigenvector with corresponding eigenvalue 100.

---

5. Let H be an $n \times n$ matrix and let I be the $n \times n$ identity matrix. Define $f : \mathbb{R} \to \mathbb{R}$ by

$$f(t) = \log|\det(I+tH)|$$

Show that (a) f is differentiable at 0 and that (b) $f'(0) = \operatorname{trace}(H)$.

Solution: Let h_i denote the i'th column of H and e_i the i'th column of I. Since the determinant is multilinear,

$$P(t) = \det(I+tH) = \det[e_1+th_1, \ldots, e_n+th_n]$$

is a polynomial of degree n in t and so is infinitely differentiable. Note that at $t=0$, $P(t)=1$, so $|P(t)| = P(t)$ near $t=0$, so $\log|P(t)|$ is differentiable at $t=0$.

By the chain rule and that $|x|'=1$ near $x=1$ we have $f'(0) = \frac{P(0)}{P'(0)}$.

Now

$$\begin{aligned} P(t) &= \det[e_1+th_1, \ldots, e_n+th_n] \\ &= \det(I) + t\Big(\det[h_1, e_2, \ldots, e_n] + \det[e_1, h_2, e_3, \ldots, e_n] + \cdots + \det[e_1, e_2, \ldots, h_n]\Big) + O(t^2) \end{aligned}$$

so

$$P'(0) = \det[h_1, e_2, \ldots, e_n] + \det[e_1, h_2, e_3, \ldots, e_n] + \cdots + \det[e_1, e_2, \ldots, h_n]$$

and it is easy to see that $\det[e_1, e_2, \ldots, h_i, \ldots, e_n] = H_{ii}$ and so $P'(0)/P(0) = \operatorname{trace}(H)/1$.

---

6. Let c, b be real vectors and let A be a real matrix. Let $P(c,A,b)$ denote the linear program

$$\text{find } x \text{ to} \quad \max c^T x \quad \text{subject to} \quad Ax \le b,\ x \ge 0.$$

Let $\mathscr{P}$ denote the set of all linear programs that can be expressed in the form $P(\_,\_,\_)$.

Let $L \in \mathscr{P}$. Although one typically considers the dual $D(L)$ as a minimization problem show that, in fact, $D(L) \in \mathscr{P}$ and that $D(D(L)) = L$.

Solution: If $L = P(c,A,b)$ then the dual (in minimization form) is

$$\text{find } y \text{ to} \quad \min b^T y \quad \text{subject to} \quad A^T y \ge c,\ y \ge 0$$

but this can be rewritten as

$$\text{find } y \text{ to} \quad \max -b^T y \quad \text{subject to} \quad -A^T y \le -c,\ y \ge 0$$

which is the program $P(-b,-A^T,-c)$. Thus D is the transformation

$$D : L(c,A,b) \mapsto L(-b,-A^T,-c).$$

and so

$$D(D(L(c,A,b)) = D(L(-b,-A^T,-c)) = L(-(-c),-(-A^T)^T,-(-b)) = L(c,A,b).$$

---

7. Let $A = [a_{ij}]$ be an $n \times n$ matrix. Define c_k (where $1 \le k \le n$) to be the sum of the determinants of all the $k \times k$ principal submatrices of A (a principal submatrix is a submatrix formed by taking any k rows and the same k columns). In this problem we ask you to prove that

$$|c_k| \le \binom{n}{k} k^{k/2} \alpha^k$$

where $\alpha = \max_{ij} |a_{ij}|$.

To this end, you may use (without proof) Hadamard's inequality that

$$|\det B| = \prod_{i=1}^{m} \left(\sum_{j=1}^{m} |b_{ij}|^2 \right)^{1/2}$$

where $B = [b_{ij}]$ is any $m \times m$ matrix.

Solution: The number of principal submatrices of A is $\binom{n}{k}$. By Hadamard's inequality, they are bounded in absolute value by

$$\prod_{i=1}^{k} \left(\sum_{j=1}^{k} \alpha^2 \right)^{1/2} = \prod \left(k\alpha^2\right)^{1/2} = k^{k/2} \alpha^k$$

and so c_k is bounded by $\binom{n}{k}$ times this expression.

8. Let $f : \mathbb{R} \to \mathbb{R}$ be the function defined by $f(x) := x|x|$. (In answering the following questions about f, you may find it notationally convenient to refer to the "signum" function sgn, which has value $1, -1, 0$ according as its argument is positive, negative, or zero.)

   (a) Show that f is differentiable and strictly increasing.

   (b) By part (a), f must have a continuous and strictly increasing inverse function g. Obtain and justify a formula for $g(y)$.

Solution:

(a) On the positive x-axis, $f(x) = x^2$ so that $f'(x) = 2x > 0$; thus f is differentiable and (strictly) increasing there, and also positive. Similarly, on the negative x-axis, $f(x) = -x^2$ so that $f'(x) = -2x > 0$, implying that f is differentiable and increasing, as well as negative, there. Because $f(0) = 0$, while $\text{sgn}[f(x)] = \text{sgn}(x)$ for $x \neq 0$, it follows that f is increasing throughout $\mathbf{R}$.

It remains only to show that f is differentiable at 0. Indeed, we will show that $f'(0) = 0$. As $x \to 0$ we have

$$\frac{f(x) - f(0)}{x - 0} = \frac{x|x|}{x} = |x| \to 0,$$

giving the desired derivative.

(b) Reversing the axes in the graph of f, and staring at the result, leads to the claim that $g(y) = \text{sgn}(y)\sqrt{|y|}$. Indeed, with this choice,

- for $x = 0$ we have $g(f(0)) = g(0) = 0 = x$,
- for $x > 0$ we have $g(f(x)) = \text{sgn}(x^2)\sqrt{|x^2|} = 1 \times x = x$, and finally
- for $x < 0$ we have

$$g(f(x)) = \text{sgn}(-x^2)\sqrt{|-x^2|} = (-1)\sqrt{x^2} = (-1)(-x) = x.$$

So $g(f(x)) = x$ for all x, showing that g is the inverse of f.

---

9. Let G be a (finite, simple) graph, $\Delta(G)$ the maximum degree in G, and $\chi(G)$ the chromatic number of G. Prove that $\chi(G) \leq \Delta(G) + 1$.

Solution: The proof is by induction on the number of vertices in G. In the basis case $n = 1$, $\chi(G) = 1$ and $\Delta(G) = 0$, so the result holds.

Suppose the result has been proved for graphs with fewer than n vertices, and let G be a graph with n vertices. Choose a vertex v and note (using induction) that $\chi(G-v) \leq \Delta(G-v)+1 \leq \Delta(G)+1$. Properly color $G-v$ with a pallete of $\Delta(G)+1$ colors. Since v is adjacent to at most $\Delta(G)$ vertices of $G-v$, there remains a color we can apply to v giving a proper color of G with $\Delta(G)+1$ colors.

---

10. Let N be a random variable whose distribution is Poisson with parameter λ, so that

$$P[N=n] = e^{-\lambda}\lambda^n/n!,\ n = 0,1,2,\ldots,$$

and suppose that $X_1, X_2, \ldots,$ is a sequence of iid positive-valued random variables independent of N and whose expected value is μ. Show that

$$E\left[\prod_{i=1}^{N} X_i\right] = e^{\lambda(\mu-1)}.$$

Solution: By conditioning on N we see that

$$E\left[\prod_{i=1}^{N} X_i\right] = \sum_{n=0}^{\infty} P[N=n] E\left[\prod_{i=1}^{N} X_i | N=n\right]$$

$$= \sum_{n=0}^{\infty} \left\{e^{-\lambda}\lambda^n/n!\right\} E\left[\prod_{i=1}^{n} X_i | N=n\right]$$

$$= \sum_{n=0}^{\infty} \left\{e^{-\lambda}\lambda^n/n!\right\} \mu^n$$

$$= e^{-\lambda} \sum_{n=0}^{\infty} (\mu\lambda)^n/n!$$

$$= e^{-\lambda} e^{\mu\lambda} = e^{\lambda(\mu-1)}.$$

---

11. Define the function $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \begin{cases} x^2, & \text{if } x \text{ is rational,} \\ 0, & \text{if } x \text{ is irrational.} \end{cases}$$

(a) f is continuous at exactly one point x_0. What is the value of x_0 and prove that f is continuous there. (You do not need to show that f is not continuous elsewhere.)

(b) Is f differentiable at x_0? (Prove your answer.)

Solution: (a) We show that f is continuous at 0 where $f(0) = 0$. Let ε be given with $\varepsilon > 0$, and choose $\delta = \min\{1, \varepsilon\}$. We show that if $|x| < \delta$, then $|f(x) - 0| < \varepsilon$. Note that

$$|f(x)| \leq |x^2| < \delta^2 \leq \delta \leq \varepsilon$$

(since $\delta \leq 1$).

(b) Yes. Note that

$$g(h) = \frac{f(h) - f(0)}{h - 0} = \begin{cases} h^2/h & \text{if } h \text{ is rational} \\ 0/h & \text{if } h \text{ is irrational} \end{cases}$$

and so $|g(h)| \leq h$ for all h. Therefore $\lim_{h \to 0} g(h) = 0$ and so $f'(0) = 0$.

---

12. What is the expected number of times a fair die is rolled until all 6 sides appear at least once?

Solution: One side appears on the first roll. The number of additional rolls until the second side appears is a geometric random variable with parameter 5/6. Given that k sides have appeared, the number of additional rolls until the $k+1$st side appears is a geometric random variable with parameter $(6-k)/6$. Summing the initial one and the expectation of five geometric random variables, we obtain

$$1 + \frac{6}{5} + \frac{6}{4} + \frac{6}{3} + \frac{6}{2} + 6 = 14.7$$

---

13. Prove that if K is a real skew-symmetric matrix and $I + K$ is nonsingular, then $(I - K)(I + K)^{-1}$ is orthogonal.

Note: K is skew-symmetric means that $K^T = -K$.

Solution: First note that $I - K$ is also nonsingular, since $I - K = (I + K)^T$. Now we have

$$\left[(I-K)(I+K)^{-1}\right]^T \left[(I-K)(I+K)^{-1}\right] = \left[(I-K)^{-1}(I+K)\right] \left[(I-K)(I+K)^{-1}\right],$$

which clearly equals the identity because $I + K$ and $I - K$ commute.

14. Two random variables X and Y have the joint density function

$$f_{(X,Y)}(x,y) = \frac{e^{-\frac{x^2}{2}+x}}{\sqrt{2\pi}} e^{-y} \text{ for } x \le y \text{ and 0 otherwise.}$$

(a) Compute and identify the marginal density function of X.

(b) What is the conditional density function of Y given $X = x$?

Solution: (a) We have

$$\begin{aligned}
f_X(x) &= \int_{\mathbb{R}} f_{(X,Y)}(x,y)dy \\
&= \int_x^{+\infty} \frac{e^{-\frac{x^2}{2}+x}}{\sqrt{2\pi}} e^{-y} dy \\
&= \frac{e^{-\frac{x^2}{2}+x}}{\sqrt{2\pi}} [-e^{-y}]_x^{+\infty} \\
&= \frac{e^{-\frac{x^2}{2}+x}}{\sqrt{2\pi}} [-e^{-y}]_x^{+\infty} \\
&= \frac{e^{-\frac{x^2}{2}+x}}{\sqrt{2\pi}} e^{-x} \\
&= \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}
\end{aligned}$$

This is the density of a standard Gaussian distribution.

(b) The conditional density of Y given X is

$$g_Y(y|x) = \frac{f_{(X,Y)}(x,y)}{f_X(x)}$$

It is 0 for $y < x$ and for $y \ge x$, it is (cancelling terms which appear both in the numerator and the denominator)

$$g_Y(y|x) = e^{-y+x}$$

The expectation is

$$E(Y|X=x) = \int_x^{\infty} y e^{-y+x} dy = \int_0^{\infty} (z+x)e^{-z} dz = x+1.$$

15. Let $y, x_1, \ldots, x_n$ be (scalar) variables, and consider the single constraint

$$ny \geq x_1 + x_2 + \cdots + x_n \tag{1}$$

and the family of constraints

$$y \geq x_j \text{ for all } j. \tag{2}$$

(a) Show that if all the variables are binary (0 and 1 are the only values), then (1) and (2) are equivalent.

(b) In contrast, show that if all the variables are continuous with values in $[0,1]$, and also $n > 1$, then one of (1) and (2) (which one?) is strictly more restrictive than the other.

Solution:

(a) Suppose all the variables are binary. If $y = 0$, then since all $x_j \geq 0$, constraint (1) forces all $x_j = 0$, while (2) does exactly the same. And if $y = 1$, then since all $x_j \leq 1$, constraint (1) permits the x's to take any combination of values from $[0,1]$, while (2) does exactly the same. So (1) and (2) permit exactly the same sets of variable-values, and thus are equvalent.

(b) Now suppose all the variables are $[0,1]$-valued. Summing the constraints (2) gives (1), so (2) $\Longrightarrow$ (1), i.e., (2) is no less restrictive than (1). To show it is strictly more restrictive, i.e., that (1) does not imply (2), it suffices to take (e.g., with $y = 2/3$)

$$1/2 \leq y < 1, \qquad x_1 = 1, \qquad x_2 = 2y - 1, \qquad x_j = y \text{ for all } j > 2.$$

---

16. A man claims to have extrasensory perception (ESP): a supernatural ability to predict the future. As a test, before tossing a fair coin 10 times, he is asked to predict the outcomes in advance. The experiment is performed, and he gets 7 of 10 correct. What is the probability that he would have done at least this well if he does not have ESP?

Solution: If he does not have ESP, he guesses for each toss, so has probability 1/2 of being correct on each toss. Assuming the tosses are independent, the number of correct guesses is a Binomial(10, 1/2) random variable. Thus, the probability that it takes a value of 7, 8, 9, or 10 is

$$2^{-10}\left(\binom{10}{7} + \binom{10}{8} + \binom{10}{9} + \binom{10}{10}\right) = \frac{120 + 45 + 10 + 1}{1024} = \frac{176}{1024} = \frac{11}{64}.$$

---

17. Suppose $T : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation with the property that its matrix is the same relative to any basis. Show that this matrix must be a scalar multiple of the identity matrix.

Hint: If there exists x for which $T(x)$ is not a scalar multiple of x, consider what the matrix of T looks like relative to a basis containing both x and $T(x)$.

Solution: Suppose there exists a nonzero x_1 such that $T(x_1)$ is not a scalar multiple of x_1. Set $x_2 = T(x_1)$. Then x_1, x_2 are linearly independent, and we can expand to an ordered basis $x_1, x_2, \ldots, x_n$. The expression of T in this basis has a column of the form $[0, 1, 0, \ldots, 0]$.

If $n = 2$ then we can take as our ordered basis x_2, x_1 and since the matrix of T is of the same form we must have $T(x_2) = x_1$ and the form of the matrix of T relative to this basis is

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

But when we take our basis to be $y_1 = x_1 + x_2$ and $y_2 = x_1 - x_2$ we get $T(y_1) = y_1$ and $T(y_2) = -y_2$ so relative to this basis, the matrix of T is a diagonal matrix, a contradiction.

If $n > 2$, by assumption, we get the same matrix if we express T using a basis $x_1, x_i, \ldots$, where x_i and x_2 have been permutated. But this would give $T(x_1) = x_i = x_2$ which contradicts the linear independence of the basis elements.

In either case, we conclude $T(x)$ is a scalar multiple of x for every nonzero $x \in \mathbb{R}^n$. This tells us that the matrix of T is diagonal. Finally, if we pick a basis $x_1, \ldots, x_n$ and write $T(x_i) = a_{ii} x_i$, then when we swap x_i and x_j in the basis, we should get the same matrix, so $a_{ii} = a_{jj}$ for all choices of i and j.

---

18. Suppose that $x = (3, 1)$ and $w = (0, 2/3, 1/3)$ are optimal solutions to the following problem and its dual, respectively.

$$\begin{array}{lrcl} \max & \alpha x_1 - x_2 & & \\ \text{s.t.} & -x_1 + x_2 & \leq & 2 \\ & x_1 + \beta x_2 & \leq & 1 \\ & \gamma x_1 + x_2 & \leq & 4 \\ & x_1, \quad x_2 & \geq & 0 \end{array}$$

Find the values of α, β and γ.

Solution: The dual problem is:

$$\begin{array}{ll} \min & 2w_1 + w_2 + 4w_3 \\ s.t. & -w_1 + w_2 + \gamma w_3 \quad \geq \alpha \\ & w_1 + \beta w_2 + w_3 \quad \geq -1 \\ & w_i \geq 0 \end{array}$$

By the Complementary Slackness Principle, we have:

$w_2 = \frac{2}{3} > 0$ implies $x_1 + \beta x_2 = 1$. Therefore, $\beta = -2$ because $x_1 = 3$ and $x_2 = 1$.

$w_3 = \frac{1}{3} > 0$ implies $\gamma x_1 + x_2 = 4$. Therefore, $\gamma = 1$

$x_1 = 3 > 0$ implies $-w_1 + w_2 + \gamma w_3 = \alpha$. It follows from $w = (0, 2/3, 1/3)$ and $\gamma = 1$ that $\alpha = 1$.

---

19. Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function with the property that

$$f(x+y) = f(x) + f(y)$$

for all values of $x, y \in \mathbb{R}$.

Prove that $f(cx) = cf(x)$ for all $c, x \in \mathbb{R}$.

Solution: From $f(x+y) = f(x) + f(y)$ it follows (e.g., by induction) that $f(nx) = nf(x)$ for any positive integer n. From this it follows that $f(x)/n = f(x/n)$. Therefore

$$f(mx/n) = mf(x/n) = \frac{m}{n} f(x) \qquad (*)$$

for any positive integers m, n.

$f(0+0) = f(0) + f(0)$ gives $f(0) = 0$ and so $f(x-x) = f(0) = f(x) + f(-x)$, so $f(-x) = -f(x)$. Thus we can extend $(*)$ to any rational multiplier, i.e.,

$$f(qx) = qf(x)$$

for any $q \in \mathbb{Q}$.

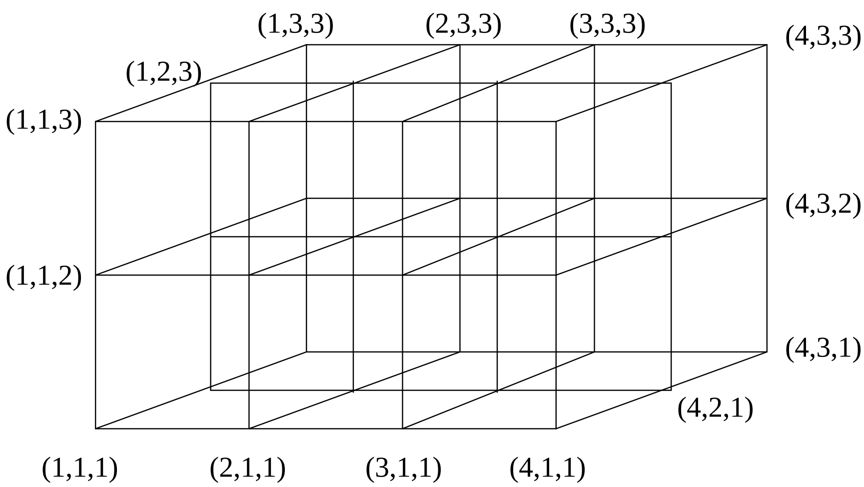
Choose rationals $q_1, q_2, \ldots$ that converge to c, and so by continuity $f(q_k x) \to f(cx)$ but $f(q_k x) = q_k f(x) \to cf(x)$ and so $f(cx) = cf(x)$.

---

20. For given positive integers p, q and r, the $p \times q \times r$ wire mesh in $\mathbb{R}^3$ involves wires connecting pairs of points in the set

$$\mathscr{P}_{p,q,r} = \{(i,j,k) \,:\, 1 \le i \le p, 1 \le j \le q, 1 \le k \le r\},$$

and wire connections between every pair of points in $\mathscr{P}_{p,q,r}$ of the form $((i,j,k),(i+1,j,k))$, $((i,j,k),(i,j+1,k))$ or $((i,j,k),(i,j,k+1))$. For example, the figure below shows a $4 \times 3 \times 3$ mesh. Suppose an ant starts at the point $(1,1,1)$ and moves to the point (p,q,r) in the



$p \times q \times r$ wire mesh, using exactly $p+q+r-3$ steps, in such a way that

- at the end of every step the ant is still at a point with integer coordinates,
- each step involves moving a Euclidean distance of one unit.

How many distinct paths are there that the ant can follow?

Solution: There are $p+q+r-3$ steps to make, and 3 different types: $p-1$ of the first type (along x axis) , $q-1$ of the second type (y axis), and $r-1$ of the third type (z axis). The number of paths is the number of ways of ordering the types, so the answer is

$$\binom{p+q+r-3}{p-1,q-1,r-1}.$$