Department of Applied Mathematics and Statistics
The Johns Hopkins University

# INTRODUCTORY EXAMINATION–WINTER SESSION

Wednesday, January 17, 2007

**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 $n$ be a positive integer. Prove that

$$(2n)(2n-1)(2n-2)\cdots(n+1)$$

is divisible by $n!$

*Solution:* We know that $\binom{2n}{n}$ is the number of $n$-element subsets of a $2n$-element set, and is therefore an integer. Note that

$$\binom{2n}{n} = \frac{(2n)!}{n!n!} = \frac{(2n)(2n-1)\cdots(n+1)}{n!}$$

and the result follows.

---

2. Let $(x_i, y_i),\ i = 1,2,3$ be three points in the plane. Define

$$A = \begin{bmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix}$$

and let $S$ be defined as the set of points $(x,y)$ for which

$$\det \begin{bmatrix} 1 & 1 & 1 & 1 \\ x & x_1 & x_2 & x_3 \\ y & y_1 & y_2 & y_3 \\ x^2+y^2 & x_1^2+y_1^2 & x_2^2+y_2^2 & x_3^2+y_3^2 \end{bmatrix} = 0$$

Show that if $\det A \neq 0$ then $S$ is a circle (with positive radius) that passes through all of the three points $(x_i, y_i)$.

*Solution:* Observe that $(x_i, y_i) \in S$ for $i = 1,2,3$ since substituting $(x_i, y_i)$ for $(x,y)$ in the $4 \times 4$ matrix leads to a pair of equal columns hence a vanishing determinant.

We can expand the determinant of the $4 \times 4$ matrix using the first column and the resulting equation for $S$ takes the form

$$a(x^2+y^2) + bx + cy + d = 0$$

with $a = \det A$. Since $a \neq 0$ we can complete the square and the defining equation for $S$ takes the form

$$(x-U)^2 + (y-V)^2 = W,$$

for some constants $U, V$ and $W$. We know that $S \neq \emptyset$ so it must be the case that $W \geq 0$. Furthermore, since $a \neq 0$ the points must be distinct, and we can conclude that $W > 0$ so $S$ forms a circle with radius $\sqrt{W} > 0$ that contains all of the $(x_i, y_i)$.

---

3. Let $X$ and $Y$ be independent Bernoulli random variables with parameter $p = 1/2$. Let $U = X+Y$ and $V = |X-Y|$.

   Compute the correlation between $U$ and $V$. Are $U$ and $V$ independent?

   *Solution:* Since $V = 0$ or 1, we have $V = V^2 = X+Y-2XY$. We have $E(U) = 1$ and $E(V) = 1-2/4 = 1/2$. We have $UV = (X+Y)^2 - 2XY(X+Y) = X+Y+2XY-4XY = X+Y-2XY = V$ so that $E(UV) = 1/2$ and $\mathrm{cov}(U,V) = 0$: the variables are uncorrelated.

   They cannot be independent since $U = 0 \Rightarrow V = 0$ so that $P(U = 0.V = 0) = P(U = 0) \neq P(U = 0)P(V = 0)$.

---

4. Let $I \subseteq \mathbb{R}$ be an interval and let $f : I \to \mathbb{R}$ be increasing on $I$. Suppose that $c \in I$ is not an end point of $I$. Show that $\lim_{x\to c^-} f = \sup\{f(x) : x \in I, x < c\}$.

   *Solution:* First note that if $x \in I$ and $x < c$, then $f(x) \leq f(c)$. Thus the set $\{f(x) : x \in I, x < c\}$, which is non-void because $c$ is not an end point of $I$, is bounded above by $f(c)$. Hence the indicated supremum exists; we denote it by $L$. If $\varepsilon > 0$ is given, then $L-\varepsilon$ is not an upper bound of this set. Hence there exists $y_\varepsilon \in I$, $y_\varepsilon < c$ such that $L-\varepsilon < f(y_\varepsilon) \leq L$. Since $f$ is increasing, we deduce that if $\delta(\varepsilon) := c - y_\varepsilon$, and if $0 < c-y < \delta(\varepsilon)$, then $y_\varepsilon < y < c$ so that $L-\varepsilon < f(y_\varepsilon) \leq f(y) \leq L$. Therefore $|f(y)-L| < \varepsilon$ when $0 < c-y < \delta(\varepsilon)$. Since $\varepsilon > 0$ is arbitrary we infer the desired result.

---

5. Let $A = \{a_1, a_2, \ldots, a_m\}$ be distinct "demand points" in $\mathbb{R}^n$, and let $W = \{w_1, w_2, \ldots, w_m\}$ be a corresponding set of positive numerical "weights". The Weber Plant Location Problem with these data, $P(A,W)$, calls for determining a point $x$ in $\mathbb{R}^n$ with a minimum sum of weighted Euclidean distances from the demand points, i.e., calls for

   $$\text{Minimize } f(x) := w_1|x-a_1| + w_2|x-a_2| + \cdots + w_m|x-a_m|.$$

   Show that $P(A,W)$ is a convex program, and that it has at least one optimal solution.

   *Solution:* We first show that the function $f$ is both continuous and convex. That's true because these properties of functions are preserved both under addition and under multiplication by a positive constant. Thus it suffices to use the properties of the Euclidean norm to show that functions of the form $|x-a|$ have the properties. Applying the Triangle Inequality to $\{x,y,a\}$, gives the continuity-exhibiting

$$\big||x-a|-|y-a|\big| \le |x-y|.$$

And for all $t \in [0,1]$ we have

$$\begin{aligned}|((1-t)x+ty)-a| &= |(1-t)(x-a)+t(y-a)|\\ &\le |(1-t)(x-a)|+|t(y-a)|\\ &= |1-t||x-a|+|t||y-a|\\ &= (1-t)|x-a|+t|y-a|,\end{aligned}$$

exhibiting convexity.

The problem is a convex program, since it involves minimizing a convex function over the convex set $\mathbb{R}^n$. To assure existence of an optimal solution, it suffices to show that the minimization of the continuous function $f$ can be confined to a closed bounded subset of $\mathbb{R}^n$.

To obtain such a subset $K$, choose any point $p \in \mathbb{R}^n$, and define

$$K := \{x \in \mathbb{R}^n : f(x) \le f(p)\}.$$

Then the minimization can be confined to $K$ since any point not in $K$ is "worse" than member $p$ of $K$. The subset is closed as a "lower level set" of the continuous function $f$, and is bounded because

$$f(x) \to \infty \text{ as } |x| \to \infty,$$

this "coercive" property holding because $f(x) \ge w_1|x-a_1|$ and because, by the Triangle Inequality applied to the origin and $(x, a_1)$,

$$|x-a_1| \to \infty \text{ as } |x| \to \infty.$$

---

6. Recall that the Frobenius norm of real matrix $M = [m_{ij}]$ is given by

$$||M||_F^2 = \sum_{ij} m_{ij}^2.$$

A standard decomposition of a real square ($n \times n$) matrix $M$ into symmetric and antisymmetric parts is given by $M = S + A$, where $S = (M+M^T)/2$ and $A = (M-M^T)/2$.

Prove that $||M||_F^2 = ||S||_F^2 + ||A||_F^2$.

*Solution:* For $M$ a $1\times 1$ matrix (or for the diagonal, in general), we see that $m_{ii}^2 = m_{ii}^2 + 0^2$. For the off-diagonals, in general, we see that $m_{ij}^2 + m_{ji}^2 = 2((m_{ij}+m_{ji})/2)^2 + 2((m_{ij}-m_{ji})/2)^2$.

---

7. The probability of being dealt a full house in a single hand of poker is approximately .0014. Find an approximation for the probability that in 1000 independent hands of poker you will be dealt at least 2 full houses.

*Solution:* Since the poker hands are assumed to be Bernoulli trials, the random number of hands which are full houses has a Binomial(1000,.0014) distribution. Since $n$ is large and $p$ is small, this Binomial distribution is well-approximated by a Poisson distribution with mean $np = 1.4$. The event of interest is the complement of the event that there are 0 or 1 full houses, so the probability is approximately

$$1 - e^{-1.4} - (1.4)e^{-1.4} = 1 - (2.4)e^{-1.4}.$$

---

8. Assume that $a, b, c, d$ are real numbers in $[-1, 1]$ such that $|a+c| \leq 1$ and $|b+d| \leq 1$. Show that $|ad - bc| \leq 1$.

*Solution:* Assume that $ad > 1 + bc$. Then, $a$ and $d$ must have same signs, and we can assume that they are positive (otherwise, change the sign of the 4 numbers together). Also, $b$ and $c$ must have opposite signs, and we can assume that $b \geq 0$ since we can swap the pairs $(a, c)$ and $(d, b)$ to reduce to this situation. Thus, $0 \leq b + d \leq 1$ and $d \leq 1 - b$ which yields

$$1 + bc < a - ab$$

or $1 + b(c+a) < a$. Adding $c$ to both terms yields $1 + c < (1-b)(c+a)$. This implies that $c + a \geq 0$ (otherwise $c < -1$), but if $c + a \geq 0$, then $a > 1 + b(c+a) \geq 1$ which is impossible. Thus $ad - bc \leq 1$. The fact that $ad - bc \geq -1$ comes from this result, by changing the sign of $a$ and $c$.

---

9. Show that the simplex method will terminate in a finite number of iterations if degeneracy never occurs.

*Solution:* To solve a linear program $\min\{\, c^T x \,: Ax = b,\ x \geq 0\}$, the simplex method searches along the vertices of the polyhedron $\{\, x \,:\, Ax = b, x \geq 0\}$. If degeneracy does not occur, the new vertex has a strictly smaller objective function value than the previous vertex; therefore, no vertex will be repeated in the search. Because there are only a finite number of vertices, the method will terminate either at an optimal solution or detect that the objective function is unbounded below in the feasible region.

10. Let $A = \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix}$. Does there exist a matrix $P$ such that $PAP^{-1}$ is a diagonal matrix? If not, prove that no such matrix exists and if so, give such a matrix.

*Solution:* Observe that if such a $P$ were to exist, its columns would be eigenvectors of $A$ with eigenvalues given by the diagonal entries of $PAP^{-1}$. $A$ has characteristic polynomial $(\lambda-2)^2$ which has a single root $\lambda = 2$. We can solve for the corresponding eigenspace of $A$

$$A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

gives $x_2 = 0$, so the eigenspace is the span of $[1,0]$. It follows that there do not exist a pair of linearly independent eigenvectors so a diagonalizing $P$ does not exist.

---

11. In a probability space, suppose that $\{E_n, n \geq 1\}$ and $\{F_n, n \geq 1\}$ are increasing sequences of events having limits $E$ and $F$ respectively. Show that if $E_n$ is independent of $F_n$ for all $n \geq 1$ then $E$ is independent of $F$.

*Solution:* By independence, for each $n$, $P[E_n \cap F_n] = P[E_n]P[F_n]$.

Since $\{E_n, n \geq 1\}$ and $\{F_n, n \geq 1\}$ are increasing sequences of events, and the intersection of two events is an event, $\{E_n \cap F_n, n \geq 1\}$ is also an increasing sequence of events, which has a limit of $E \cap F$.

Taking limits, using the continuity property of a probability measure for each of the expressions in the independence equation, we obtain $P[E \cap F] = P[E]P[F]$.

---

12. Let $A$ be a $n \times n$ symmetric, nonsingular real matrix, $b \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$. Show that if $\hat{x}$ is a stationary (critical) point of the function $f(x) = x^T Ax + 2b^T x + \alpha$ then

$$f(\hat{x}) = \frac{\det\left(\begin{bmatrix} A & b \\ b^T & \alpha \end{bmatrix}\right)}{\det(A)}.$$

*Solution:* Since $\hat{x}$ is a critical point, we have $A\hat{x} + b = 0$. From the first determinant below, we multiply the i-th column by $\hat{x}_i$ and add it to the last column to get the equalities:

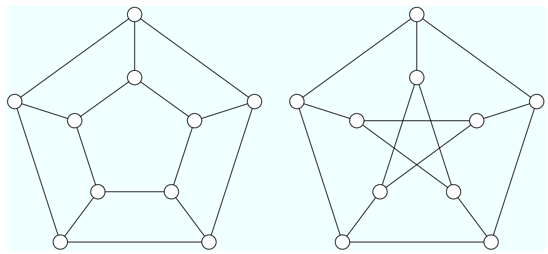
$$\det\left(\begin{bmatrix} A & b \\ b^T & \alpha \end{bmatrix}\right) = \det\left(\begin{bmatrix} A & A\hat{x}+b \\ b^T & b^T\hat{x}+\alpha \end{bmatrix}\right) = \det(A)(b^T\hat{x}+\alpha).$$

Therefore, the result follows from

$$f(\hat{x}) = \hat{x}^T(A\hat{x}+b) + b^T\hat{x} + \alpha = b^T\hat{x} + \alpha.$$

---

13. Consider the two graphs in the figure. Prove that they are not isomorphic.



*Solution:* The graph on the right is Petersen's graph which is known to be nonplanar. [*Proof*: Show that it contains a subgraph homeomorphic to $K_{3,3}$ or contract the five "radial" edges to show that it has a $K_5$ minor.]

The graph on the left is clearly planar.

Therefore, they are not isomorphic.

---

14. (a) Prove that a real symmetric matrix $A$ is positive definite if and only if $A = B^T B$, where $B$ is a real, non-singular matrix.

(b) If a real symmetric matrix $A$ is positive definite, show that $\det A > 0$.

*Solution:* (a) Assume that $A = B^T B$. Then, $x^T A x = \|Bx\|_2^2 \geq 0$ and $= 0$ iff $Bx = 0$. Since $B$ is nonsingular, equality holds only if $x = 0$.

For the converse, note that all the eigenvalues $\lambda_i$ of a real, symmetric matrix are real and the eigenvectors $e_i$ may be chosen to be real and orthogonal. By the spectral decomposition

$$A = \sum_{i=1}^{n} \lambda_i e_i e_i^T.$$

If $A$ is positive definite, then all $\lambda_i > 0$. Defining

$$B = \sum_{i=1}^{n} \sqrt{\lambda_i} e_i e_i^T$$

gives $A = B^T B$ with $B$ real and non-singular (and symmetric).

(b) Using (a), note that $\det A = \det(B^T B) = (\det B)^2 > 0$.

15. Let $X$ be a positive random variable with density $f(x)$. Show that

$$E(X) = \int_0^\infty P(X \geq t)dt.$$

*Solution:*

$$\begin{aligned} E(X) &= \int_0^\infty x f(x)dx \\ &= \int_0^\infty \left(\int_0^x dt\right) f(x)dx \\ &= \int_0^\infty \left(\int_t^\infty f(x)dx\right) dt \\ &= \int_0^\infty P(X \geq t)dt. \end{aligned}$$

---

16. Consider the space $C[0,1]$ of real-valued continuous functions on the domain $[0,1]$. For $f \in C[0,1]$, define the function $Jf$ by indefinite integration:

$$(Jf)(x) := \int_0^x f(t)\,dt, \quad x \in [0,1].$$

(a) Does $J$ map $C[0,1]$ into $C[0,1]$? Is it one-to-one? Is it onto $C[0,1]$?

(b) For which functions $f$ does the sequence $(J^n f)$ converge uniformly to a limit function belonging to $C[0,1]$? Here $J^n$ denotes the $n$th iterate of $J$.

*Solution:*

(a) By the fundamental theorem of integral calculus, $J$ is a one-to-one map of $C[0,1]$ onto the space of continuously differentiable functions on $[0,1]$ that vanish at the origin. So the answers are yes, yes, and no.

(b) We claim that $J^n f$ converges uniformly to the zero function for *every* $f \in C[0,1]$. To see this, first observe that

$$|(Jf)(x)| \leq \|f\|x, \quad x \in [0,1];$$

here $\|f\| < \infty$ denotes the supremum norm of $f$. By induction,

$$|(J^n f)(x)| \leq \|f\|\frac{x^n}{n!} \leq \frac{\|f\|}{n!}, \quad x \in [0,1],$$

and the claim follows.

17. Let $T$ be a tree and let $u$ and $v$ be distinct vertices of $T$ that are not adjacent. Prove that $T$ contains an edge $e$ such that $(T-e)+uv$ is also a tree.

Note: The notation $(T-e)+uv$ means the graph formed from $T$ by deleting the edge $e$ and then adding in the edge $uv$.

*Solution:* Since $T$ is a tree there is a unique path from $u$ to $v$. Since $u$ and $v$ are not already adjacent, adding the edge $uv$ to $T$ creates a cycle $C$; indeed, this is the only cycle of $T+uv$. Let $e$ be any edge of $T$ on this cycle. Note that $T+uv$ is connected since $T$ is connected.

Because $e$ is on a cycle to $T+uv$, $e$ is not a cut edge, and so $(T+uv)-e$ is connected. Also $(T+uv)-e$ has the same number of edges as $T$, and therefore is a tree. Note, the graph $(T-e)+uv$ is exactly the same as $(T+uv)-e$.

---

18. Suppose that A is an $n\times n$ matrix with distinct eigenvalues $\lambda_i$, $i=1,..,n$. If $e(t)$ is the $n$-vector with components $e_i(t)=e^{\lambda_i t}$ and if $V$ is the Vandermonde matrix with entries $V_{ij}=\lambda_i^{j-1}$, then prove that

$$e^{At}=\sum_{j=1}^{n} d_j(t)A^{j-1}$$

where the coefficients are the components of the $n$-vector

$$d(t)=V^{-1}e(t).$$

Hint: Use the Cayley-Hamilton theorem.

*Solution:* By the Cayley-Hamilton theorem,

$$A^n+c_nA^{n-1}+\cdots+c_{n-1}A+c_nI=O,$$

where $\det(zI-A)=z^n+c_1z^{n-1}+\cdots+c_{n-1}z+c_n$ is the characteristic polynomial of $A$. It is then an easy proof by induction that

$$\sum_{k=0}^{N}\frac{1}{k!}t^kA^k=\sum_{j=1}^{n}d_j^{(N)}(t)A^{j-1} \qquad (*)$$

for scalar coefficients $d_j^{(N)}(t),\ j=1,\ldots,n$. Applying this matrix equation to the eigenvector $u_i$ for eigenvalue $\lambda_i$ gives

$$\sum_{k=0}^{N}\frac{1}{k!}(\lambda_i t)^k=\sum_{j=1}^{n}\lambda_i^{j-1}d_j^{(N)}(t)$$

for $i = 1, \ldots, n$, or

$$e^{(N)}(t) = Vd^{(N)}(t).$$

where $e_i^{(N)}(t) = \sum_{k=0}^{N} \frac{1}{k!}(\lambda_i t)^k$. Since all the eigenvalues are distinct, the Vandermonde matrix is invertible and

$$d^{(N)}(t) = V^{-1}e^{(N)}(t).$$

Because $\lim_{N\to\infty} e^{(N)}(t) = e(t)$, we get

$$\lim_{N\to\infty} d^{(N)}(t) = V^{-1}e(t) \equiv d(t),$$

and then the desired result follows by passing to the limit $N \to \infty$ in (*).

---

19. Show that if $X_1$ and $X_2$ are independent Poisson random variables with parameters $\theta_1$ and $\theta_2$, respectively, then $X_1 + X_2$ has a Poisson distribution with parameter $\theta_1 + \theta_2$.

*Solution:*

$$\begin{aligned}
P(X_1 + X_2 = k) &= \sum_{j=0}^{k} P(X_1 + X_2 = k, X_1 = j) \\
&= \sum_{j=0}^{k} P(X_2 = k - j, X_1 = j) \\
&= \sum_{j=0}^{k} \frac{e^{-\theta_2}\theta_2^{k-j}}{(k-j)!} \frac{e^{-\theta_1}\theta_1^{j}}{j!} \\
&= \frac{e^{-(\theta_1+\theta_2)}}{k!} \sum_{j=0}^{k} \frac{k!}{(k-j)!\,j!} \theta_1^j \theta_2^{(k-j)} \\
&= \frac{e^{-(\theta_1+\theta_2)}}{k!} (\theta_1 + \theta_2)^k.
\end{aligned}$$

---

20. Let $f : \mathbb{R}^p \to \mathbb{R}$ be differentiable at $c$.

    (a) Express the directional derivative of $f$ at $c$ in the direction of a given unit vector $w = (w_1, \ldots, w_p)$. Do this in terms of the partial derivatives of $f$.

    (b) Show that there is a direction in which the directional derivative is maximum and that this direction is uniquely determined if at least one of the partial derivatives is nonzero at $c$.

*Solution:*

(a) The directional derivative of $f$ at $c$ in the direction of $w$ is given by

$$f_w(c) = f_1(c)w_1 + \cdots + f_p(c)w_p,$$

where, for $j = 1, \ldots, p$, the notation $f_j(c)$ denotes the evaluation at $c$ of the partial derivative of $f$ with respect to its $j$th argument.

(b) By the Cauchy–Schwarz inequality,

$$|f_w(c)| \leq \left[ \sum_{j=1}^{p} \left| f_j(c) \right|^2 \right]^{1/2},$$

with equality if (and when at least one partial derivative is nonzero at $c$, only if) $w$ is a multiple of the vector $(f_1(c), \ldots, f_p(c))$.

---