Solutions to Review Questions for 14.102 Midterm
10/14/05

Note: For true/false questions you should either prove the statement or provide a counterexample.

1 Real Analysis

1. Give an example of a relation R that is transitive, but not symmetric.

   Solution: $>$ is one such relation; $x > y$ & $y > z \Rightarrow x > z$, but $x > y \not\Rightarrow y > x$.

2. Suppose S is an ordered set, $E \subset S$, and E is bounded above. Define the supremum of E.

   Solution: Suppose there exists an $\alpha \in S$ with the following properties:

   - α is an upper bound of E.
   - If $\gamma < \alpha$ then γ is not an upper bound of E.

   Then α is called the supremum of E, and we write $\alpha = \sup E$.

   Comment: It is very easy to leave off $\alpha \in S$, which is an important part of the definition, or to imply that $\alpha \in E$, which is not necessarily the case.

3. List the properties of a distance function. Is the following statement true: 'if $d(\cdot,\cdot)$ is a distance, then $d'(x,y) = (d(x,y))^2$ is a distance'?

   Solution:

   - $d(x,y) \geq 0$ (we do not want negative distance),
   - $d(x,y) = 0 \iff x = y$ (moreover, we want strictly positive distance between distinct points),
   - $d(x,y) = d(y,x)$ (symmetry),
   - $d(x,y) + d(y,z) \geq d(x,z)$ (triangle inequality).

   The statement is not true - for example, in $\mathbb{R}^1$, the standard Euclidean metric $d(x,y) = \sqrt{(x-y)^2}$ is a metric, whereas we showed in Problem Set 1 that $d'(x,y) = (x-y)^2$ is not.

4. State the Separating Hyperplane Theorem. Is it true that for any two disjoint *closed* convex sets C_1 and C_2 there exists a hyperplane $H(p, a)$ such that $p \cdot x < a$ for all $x \in C_1$ and $p \cdot y > a$ for any $y \in C_2$?

   Solution: The theorem states: let C_1 and C_2 be two disjoint (i.e., $C_1 \cap C_2 = \emptyset$) convex sets in $\mathbb{R}^n$. Then there exists a hyperplane $H(p, a)$ that separates $C_1$and C_2, i.e., such that $\forall x \in C_1 \ p \cdot x \leq a$ and $\forall y \in C_2 \ \ p \cdot y \geq a$.

   If C_1 and C_2 are both closed, then we can indeed replace the weak inequalities with strict inequalities (for now no point of either set can be a limit point of the other; this is possible when the sets are not closed, and necessitates the weak inequalities).

5. Show that the intersection of (even infinitely many) closed sets is closed. Give an example of an infinite family of closed sets whose union is not closed.

   Solution: By definition, a set is closed if its complement is open. We have to prove that the complement of an intersection of closed sets is open, which is the same as the union of complements of each of them (why?), which are all open by definition. Therefore, we have to prove that the union of (any number of) open sets is open. The latter statement is straightforward: if a point belongs to the union, it belongs to one of the sets which, being open, contains a small enough open ball centered in our point and this ball must, therefore, be contained in the union, proving that the union is open.

   If $A_n = \{\frac{1}{n}\}$, then each A_n is closed (it is a singleton), but their union is not (limit point 0 is missing).

6. Prove carefully that the sum of two convergent sequences is convergent and its limit is the sum of the limits.

   Solution: Let $\{a_n\} \to a$ and $\{b_n\} \to b$. We want to show that $\{a_n + b_n\} \to a + b$. By definition of convergence, this means that $\forall \varepsilon > 0 \ \exists N : \ \forall n > N \ \ |a_n + b_n - (a + b)| < \varepsilon$. Indeed, for a given $\varepsilon > 0$ consider $\varepsilon_1 = \frac{\varepsilon}{2}$. By definition of convergence, there must exist N_a such that $\forall n > N_a \ \ |a_n - a| < \varepsilon_1$; likewise, there must exist N_b such that $\forall n > N_b \ \ |b_n - b| < \varepsilon_1$. Set $N = \max\{N_a, N_b\}$. Now for any $n > N$ we have $|a_n + b_n - (a + b)| = |a_n - a + b_n - b| \leq |a_n - a| + |b_n - b| < \varepsilon_1 + \varepsilon_1 = \varepsilon$. This completes the proof.

7. Define a limit point of a sequence. Is it true that if A is a limit point of a sequence $\{a_n\}$ and B is a limit point of a sequence $\{b_n\}$

then $A+B$ is a limit point of sequence $\{a_n+b_n\}$?

Solution: A (finite) number B is called a limit point of $\{a_n\}$ if $\forall\varepsilon>0\ \ \forall N\exists n>N:\ |a_n-B|<\varepsilon$.

The statement is not true. Consider the sequences $\{a_n\}=(-1)^n$ and $\{b_n\}=(-1)^{n+1}$. Then 1 and -1 are limit points for both sequences, but neither 2 nor -2 is a limit point for the sequence $\{a_n+b_n\}=\{0,0,0,...\}$

8. Let $A=[-1;0)$ and $B=(0,1]$. Examine whether each of the following statements is true or false:

   (a) $A\cup B$ is compact.

   Answer: False. Point 0 is missing making the set not closed.

   (b) $A+B=\{x+y|x\in A, y\in B\}$ is compact.

   Answer: False. $A+B=(-1;1)$ which is not closed.

   (c) $A\cap B$ is compact.

   Answer: True. $A\cap B=\emptyset$ which is closed and bounded.

9. Define

   $$f(x)=\begin{cases}1 \text{ if } 0\le x\le 1\\ 0\ \ \text{otherwise}\end{cases}$$

   Find an open set O such that $f^{-1}(O)$ is not open and a closed set C such that $f^{-1}(C)$ is not closed.

   Examples: $O=(\frac{1}{2},\frac{3}{2})$ and $C=\{0\}$.

10. Find all the limit points of the following sequence: 1,1,2,1,2,3,1,2,3,4,...

   Solution: A finite number B is called a limit point of $\{a_n\}$ if $\forall\epsilon>0,\forall N:\exists n>N$ such that $|a_n-B|<\epsilon$. For this sequence, any positive integer is a limit point.

11. Show that any finite union of finite sets (say, n of them) is itself finite.

   Solution: A finite set is defined as a set with a 1-1 correspondence to the set J_m, whose elements are the positive integers 1,2,3,...,m, for some positive integer m. The empty set is also defined to be finite. The proof of the statement follows by induction. It holds trivially for $n=1$; the union of a finite set with itself is simply itself, and is finite. Now suppose it holds for n, and let us show that it holds for $n+1$. Denote the finite union of $n+1$ sets by

$\bigcup_{i=1}^{n+1} A_i$, where each A_i is a finite set. Note that we can rewrite this as $(\bigcup_{i=1}^{n} A_i) \bigcup A_{n+1}$, the union of the first n sets with the $n+1^{st}$. By the induction hypothesis, both of these sets are finite. Thus, it suffices to show that the union of two finite sets is finite. To do so, let B be a finite set with a 1-1 correspondence to J_b, and let C be another finite set with a 1-1 correspondence to J_c, where b and c are both positive integers. Now simply take the first element of B, and let it correspond to the integer 1; take the second and let it correspond to the integer 2, and so on, up to the last element of B, which corresponds to the integer b. Take the first element of C; if it appears already in B, we can drop it, and if not we let it correspond to the integer $b+1$. Do the same for the second element of C, and so on. If the intersection of of B and C is empty, we will arrive at the last element of C and assign to it the integer $b+c$; if not, the last element will be assigned some lower number. Either way, the set consisting of all elements in B and all elements in C (that is, $\{x|x \in B \text{ or } x \in C\}$, which is nothing more than the union of B and C) will have a 1-1 correspondence with J_p, where $p \leq b+c$, p a positive integer. The only case we have not considered is if B or C (or both) is the empty set; but the union of a finite set with the empty set is simply the same finite set, and the union of the empty set with itself is the empty set. Thus, the union of two finite sets is a finite set, and this completes the proof by induction.

2 Linear Algebra

In what follows, unless otherwise noted, let A be an $m \times n$ real matrix.

1. Define the nullspace of A.

   Solution: The nullspace of A is defined as $N(A) \equiv \{x \in \mathbb{R}^n | Ax = 0\}$.

2. State the rank-nullity theorem.

   Solution: For the given matrix A of dimension $m \times n$, $rk(A) + null(A) = n$.

3. Consider an $m \times n$ matrix A.

   (a) Let $B = A'A$. Give an upper bound for $rank(A)$. Give an upper bound for $rank(B)$.

Solution: $rank(A) \leq \min(m, n)$. $rank(A'A) \leq \min(m, n)$ as well; in general, the rank of the product of two matrices can be no greater than the minimum of the ranks of each of the two matrices. The easiest way to see this is to note that the columns of AB consist of linear combinations of the columns of A, weighted by the columns of B; clearly, then, the rank of AB cannot exceed the rank of A. To see that it cannot exceed the rank of B, note that $rank(AB) = rank(B'A')$, that $rank(B'A')$ cannot exceed $rank(B')$, and that $rank(B') = rank(B)$.

Comment: Writing $rank(B) \leq n$ was by far the most common error on the exam from which this question was taken. Make sure you understand why this is wrong.

(b) Provide a definition for the eigenvalues and eigenvectors of B.

Solution: We say that λ is an eigenvalue of B with corresponding eigenvector v if $Bv = \lambda v$ for some $v \neq 0$. Alternatively, λ is an eigenvalue of B if $|B - \lambda I| = 0$.

(c) Are the eigenvectors of B necessarily orthogonal to each other?

Solution: No, not necessarily, even though B is symmetric. If the eigenvalues are distinct, then they are, but if we have repeated eigenvalues then it is possible to find eigenvectors which are not orthogonal to each other (consider $A = I$, $B = I$). The theorem we covered in class says only that we can always find a set of orthogonal eigenvectors for a symmetric matrix, not that every set of eigenvectors will be orthogonal.

(d) Are the eigenvectors of B necessarily orthonormal? If not, show how you would find a pair of orthonormal eigenvectors.

Solution: No, they are not necessarily orthonormal (especially since they aren't necessarily orthogonal). Assuming that we have found a set of orthogonal eigenvectors, which we always can for symmetric B, we must still normalize them to have unit length. To do so, we find $\|v_i\|$ for each eigenvector v_i; this is a scalar, so $\frac{v_i}{\|v_i\|}$ will still be an eigenvector, and will have unit length.

4. Define a symmetric matrix. Is it true that the product of two symmetric matrices is a symmetric matrix?

Solution: A matrix A is symmetric if $A' = A$, or equivalently $a_{ij} = a_{ji}\ \forall i, j$. This imposes that A is square. If AB is defined and A and B are both symmetric, then we have $(AB)' = B'A' = BA$, which is not necessarily the same as AB.

5. True or false: for a matrix to be diagonalizable it is both necessary and sufficient that it have n distinct eigenvalues.

   Solution: This is false. A matrix must admit n linearly independent eigenvectors to be diagonalizable. Having n distinct eigenvalues is sufficient for this to hold, but not necessary; consider the $n \times n$ identity matrix.

6. Consider the $m \times n$ system of equations $Ax = b$. Under what conditions does there exist **no** solution to the system?

   Solution: If $rank[A, b] > rank(A)$, which necessarily implies $b \neq 0$, $m > rank(A)$, and $b \notin S(A)$, then and only then the system has no solution, $X^* = \emptyset$.

   Comment: It would also be correct to write that b must not be expressible as a linear combination of the columns of A; it is good to realize that this is what $b \notin S(A)$ means.

7. Give the eigenvalues for the diagonal square matrix

$$\begin{bmatrix} a_1 & 0 & \ldots & 0 \\ 0 & a_2 & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & a_n \end{bmatrix}$$

   where $a_i = 2^i$, $i = 1, 2, \ldots, n$.

   Solution: The eigenvalues of any diagonal matrix are simply the diagonal elements of that matrix - in this case, 2^i, $i = 1, 2, \ldots, n$.

8. Show that $S(A)$ and $N(A')$ are orthogonal subspaces, in the sense that $z \in S(A), u \in N(A') \Rightarrow z'u = 0$. Show further that $S(A) + N(A') = \mathbb{R}^m$, in the sense that for every $y \in \mathbb{R}^m$ there are vectors $z \in S(A)$ and $u \in N(A')$ such that $y = z + u$.

   Solution: For the first part, we just use the definitions of $S(A)$ and $N(A')$. $z \in S(A)$ means that $z = Ax$ for some $x \in \mathbb{R}^n$, and $u \in N(A')$ means that $A'u = 0$. So we have

$$z'u = x'A'u = 0.$$

   For the second part, I will first simply show that such z and u exist - then show how we get them. It turns out that if we choose any $x \in S(A)$, then $z = x\frac{(x'y)}{(x'x)}$ and $u = y - z$ will satisfy $y = z + u$. It is clear that $z \in S(A)$; $\frac{(x'y)}{(x'x)}$ is simply a scalar. It is also clear that $y = z + u$, since $z + u = z + y - z = y$. The only thing to check is that $z'u = 0$, since we know that $u \in N(A') \Leftrightarrow u \perp S(A)$. Checking this, we have

$$\begin{aligned} z'u &= (x\frac{(x'y)}{(x'x)})'(y - x\frac{(x'y)}{(x'x)}) \\ &= x'y(\frac{(x'y)}{(x'x)}) - x'x(\frac{(x'y)}{(x'x)})^2 \\ &= x'y(\frac{(x'y)}{(x'x)}) - x'y(\frac{(x'y)}{(x'x)}) = 0 \end{aligned}$$

Now - how would we have found this z? You might have some idea if you notice that $x\frac{(x'y)}{(x'x)}$ is commonly called the **orthogonal projection of** y **on** x. The name comes from the picture that goes with it - suppose for intution that y and x are vectors in $\mathbb{R}^2$, and that they're linearly independent. Now extend x, drawing in its full span (which is a line). If we drop a line from y to the span of x such that the line is perpendicular to x, the vector which ends at the intersection of this line and the span of x is the orthogonal projection of y on x. Notice that precisely because the line was perpendicular to the span of x, we have found the vector in the span of x whose head is closest to the head of y. This is how ordinary least squares works in econometrics. You have a LHS variable y, and you're trying to describe it as a function of RHS variables X. In this example, there's just one RHS variable, x. Your best prediction of y is going to be the vector in x that is as close as possible to y. The difference is simply $y - x$, which is our u here, commonly called the residual. You want to choose a vector in the span of x so as to minimize the norm of the residual, which is a sum of squares - hence the name ordinary least squares. Going through this minimization reveals where the formula for the orthogonal projection comes from:

$$\begin{aligned} &\min_\beta u'u \text{ s.t. } u = y - \beta x \\ &\min_\beta (y - \beta x)'(y - \beta x) \\ &\min_\beta y'y - 2\beta x'y + \beta^2 x'x \quad \text{(note that } x'y = y'x) \\ -2x\,'y + 2\beta x'x = 0& \\ &\beta = \frac{x'y}{x'x} \end{aligned}$$

So $z = x\frac{(x'y)}{(x'x)}$ is precisely the vector in the span of x whose head is closest to the head of y, which means that the segment connecting the two heads, which is the residual, must be perpendicular to it. Moreover, since this segment is simply $u = y - x\frac{(x'y)}{(x'x)}$, we know that $u + z = y$.

3 Optimization in $\mathbb{R}^n$

1. State the Weierstraß theorem. Is it true that any function that is differentiable on a compact set is bounded on that set?

   Solution: The theorem states that if $X \subset \mathbb{R}^n$ is a compact set and $f : X \to \mathbb{R}$ is a continuous function, then f attains a maximum on X, that is, there exists a point $x^* \in X$ such that $\forall x \in X$ $f(x) \leq f(x^*)$. The theorem implies that a function which is differentiable on a compact set is indeed bounded on that set, for any function which is differentiable is continuous, and any function which achieves a maximum and a minimum over a set is bounded on that set.

2. Give an example of a set $X \subset \mathbb{R}^n$, and a function $f : X \to \mathbb{R}$, such that the conditions of the Weierstraß theorem do **not** hold, but such that f nevertheless attains a maximum and a minimum on X.

   Solution: If $X \subset \mathbb{R}^n$ is a compact set and $f : X \to \mathbb{R}$ is a continuous function, then f attains a maximum on X, that is, there exists a point $x^* \in X$ such that $\forall x \in X \quad f(x) \leq f(x^*)$. A simple example is $X = \mathbb{R}$ and $f(x) = \sin x$, which has many global maxima and minima. Remember that the conditions of the theorem are sufficient but not necessary for optima to exist.

3. State the implicit function theorem. Find all points on the curve $x^4 - 2x^2y^2 + y^4 = 0$ around which either y is not expressible as a function of x or x is not expressible as a function of y. Compute $y'(x)$ along the curve at point $(1, -2)$.

   Solution: The theorem states: Let $F(x_1, ..., x_n)$ be a C^1 function around the point $(x_1^*, ..., x_n^*, y^*)$ such that $\frac{\partial F}{\partial y}(x_1^*, ..., x_n^*, y^*) \neq 0$. Denote $c = F(x_1^*, ..., x_n^*, y^*)$. Then there exists a C^1 function $y = y(x_1, ..., x_n)$ defined around $(x_1^*, ..., x_n^*)$ such that:

   - $F((x_1, ..., x_n, y(x_1, ..., x_n)) = c$
   - $y^* = y(x_1^*, ..., x_n^*)$
   - $\frac{\partial y}{\partial x_i}(x_1^*, ..., x_n^*) = -\frac{\frac{\partial F}{\partial x_i}(x_1^*, ..., x_n^*, y^*)}{\frac{\partial F}{\partial y}(x_1^*, ..., x_n^*, y^*)}$.

For $F(x,y) = x^4 - 2x^2y^2 + y^4$, y is not expressible as a function of x when $\frac{\partial F}{\partial y} = -4x^2y + 4y^3 = 0$; i.e. when $y = 0$ or $y = x$ or $y = -x$.

x is not expressible as a function of y when $\frac{\partial F}{\partial x} = 4x^3 - 4xy^2 = 0$; i.e. when $x = 0$ or $x = y$ or $x = -y$. Notice that the function is symmetric in its two arguments - indeed, that it is equal to $(x^2 - y^2)^2 = [(x+y)(x-y)]^2$ Rewriting the function in this way makes it much easier to visualize the set of points described by the level set $F(x,y) = 0$.

$y'(x)$ at $(1,-2)$ is $-\frac{\frac{\partial F}{\partial x}(1,-2)}{\frac{\partial F}{\partial y}(1,-2)} = \left.\frac{4x^3-4xy^2}{4x^2y-4y^3}\right|_{(1,-2)} = \frac{4-16}{-8+32} = -\frac{1}{2}$.

4. Consider the problem of maximizing $f(x)$ subject to $h(x) \leq 0$. The Lagrangian is

$$L(x_1, ..., x_n, \lambda_1, ..., \lambda_k) = f(x_1, ..., x_n) - \sum_{i=1}^{k} \lambda_i h_i(x_1, ..., x_n) \quad (1)$$

Explain in words why it must be the case that $\lambda_i \geq 0$ for $i = 1, ..., k$ at a local optimum (x^*, λ^*).

Solution: Recall that the intution for the Lagrange theorem is that at an optimum, we have $\nabla f = \lambda \nabla h$. In an optimum where the i^{th} constraint is not binding the problem locally looks like an unconstrained problem with respect to this constraint, and the first order condition will be $\nabla f = 0$, i.e. $\lambda_i = 0$. For a constraint that is binding, $\lambda_i \neq 0$. But we can say more: if λ_i were negative, we could move slightly from the prospective maximum in the direction of ∇f, and that will not violate the constraint (we would be moving in the direction opposite to ∇h_i, so h_i would decrease and hence still remain nonpositive). Therefore, at any local optimum it must be the case that $\lambda \geq 0$.

Comment: Note that we often say that λ is the shadow value of the constraint, and thus must be zero when the constraint does not bind and positive when it does. This is a very good economic interpretation of the previous paragraph.

5. Alden and Nicole recently moved into an apartment in need of a few repairs. For instance, the shower, water heater, and gas lines all needed some work before they could be used. Let us denote repairs to the shower by x, repairs to the water heater by y, and repairs to the gas lines by z, where $x, y, z \in \mathbb{R}$, and let us denote

Alden and Nicole's utility from repairs to the three by

$$U(x, y, z) = xyz$$

Alden and Nicole are endowed with one dollar (they're both grad students), which they can spend on repairs to the shower, water heater, and gas lines (assume they get utility only from the consumption of these three goods). Assume that $p_x = p_y = p_z = 1$.

(a) Write down the maximization problem faced by Alden and Nicole, including their constraints (assume that we do not allow negative consumption of x, y, or z[1]).

Solution: The problem is

$$\begin{aligned} \max_{x,y,z} & \, xyz \\ s.t. x + y + z & \leq 1 \\ x & \geq 0 \\ y & \geq 0 \\ z & \geq 0 \end{aligned}$$

(b) Solve the maximization problem by the Kuhn-Tucker (i.e., Lagrangian) method, writing down all first order conditions, including complementary slackness conditions.

Solution: Usually with Kuhn-Tucker maximization, we like to write all our inequality constraints the same way (as something less than a constant, in this case zero); thus, we rewrite the last three constraints as

$$\begin{aligned} -x & \leq 0 \\ -y & \leq 0 \\ -z & \leq 0 \end{aligned}$$

We now form the Lagrangian

$$L(x, y, z, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = xyz - \lambda_1(x+y+z-1) + \lambda_2 x + \lambda_3 y + \lambda_4 z$$

Note that because the last three constraints are on $-x, -y, -z$, the constraints enter the Lagrangian positively rather than negatively (and indeed, we can skip the step of rewriting the

[1] In real life, Alden could, say, try to work on the gas lines himself, making matters worse, but we'll just ignore that.

constraints if we remember to enter nonnegativity constraints in this way).

The first order conditions are:

$$\begin{aligned}
\frac{\partial L}{\partial x} &= yz - \lambda_1 + \lambda_2 = 0\\
\frac{\partial L}{\partial y} &= xz - \lambda_1 + \lambda_3 = 0\\
\frac{\partial L}{\partial z} &= xy - \lambda_1 + \lambda_4 = 0\\
\lambda_1(x+y+z-1) &= 0, \lambda_2 x = 0, \lambda_3 y = 0, \lambda_4 z = 0\\
\lambda_1 &\geq 0, \lambda_2 \geq 0, \lambda_3 \geq 0, \lambda_4 \geq 0\\
x+y+z &\leq 1\\
x &\geq 0, y \geq 0, z \geq 0
\end{aligned}$$

The fourth row contains the complementary-slackness conditions, while the last two rows give the first order constraints with respect to the Lagrange multipliers, and simply restate the intitial constraints.

We can rewrite the first three constraints as

$$\lambda_1 = yz + \lambda_2 = xz + \lambda_3 = xy + \lambda_4 \qquad (2)$$

Now consider two cases: $\lambda_1 = 0$ and $\lambda_1 > 0$.

First, if $\lambda_1 = 0$, then the fact that every variable in 2 is nonnegative implies that:

$$xy = yz = xz = 0 \text{ and } \lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0 \qquad (3)$$

3 leads to the infinite set of candidate solutions in which two of the variables are zero and the third is any number in $[0,1]$. Utility equals zero for all x, y, z satisfying 3. Can we do better?

Aside: we'd better hope so; Alden and Nicole would really like a hot shower one of these days.

If $\lambda_1 > 0$, it must be that $x + y + z = 1$; that is, the budget constraint binds. Moreover, this means that at least one of $x, y,$ and z must be nonzero. Suppose for a moment that $x = 0$. Then, using 2 and the assumption that $\lambda_1 > 0$, we see that $\lambda_1 = \lambda_3 = \lambda_4 > 0$. But if this is the case, the complementary slackness conditions tell us that $y = z = 0$, a contradiction of the fact that at least one of $x, y,$ and z

must be nonzero. Thus, it must be that $x > 0$, and similar arguments tell us that $y > 0$ and $z > 0$ as well. Now the complementary slackness conditions tell us that $\lambda_2 = \lambda_3 = \lambda_4 = 0$ (which implies that none of the other three constraints bind, something which we just established as being true). Now we have

$$xy = xz = yz$$

and from this and the binding budget constraint it follows that $x = y = z = \frac{1}{3}$. These values for x, y, and z give utility of $\frac{1}{27}$, which is indeed the maximal value of the utility function given the constraints. $\frac{1}{27}$ may not sound like much, but it's a whole heck of a lot better than zero. Take it from me - er, Alden.

(c) Extra Credit: What (in words) is the interpretation of the functional form of Alden and Nicole's utility function?

Solution: The functional form simply implies that utility is zero if any of the three arguments are zero. We can also see that x, y, and z are *complements* - the higher is one argument, the greater is the rate at which utility rises with an increase in any of the other components (you can see this mathematically by observing that the cross-partial derivatives $\frac{\partial^2 U}{\partial x \partial y}, \frac{\partial^2 U}{\partial x \partial z}, \frac{\partial^2 U}{\partial y \partial z}$ are all nonnegative). This makes sense for the story we told. A shower's not much good without gas or hot water, and similarly for the other two goods (remember that we're assuming that utility is derived only from repairs to these three, so we're ignoring the fact that gas and hot water have other uses). Finally, note that this is a standard Cobb-Douglas utility function with elasticity the same on each good; that, combined with the price being the same on each good, implies an even distribution of consumption among the three.

6. (Sundaram 6.12, page 171) A firm produces a single output y using three inputs x_1, x_2, x_3 in nonnegative quantities through the relationship $y = x_1(x_2 + x_3)$.The unit price of y is $p_y > 0$ while that of the input x_i is $w_i > 0, i = 1, 2, 3$.

   (a) Describe the firm's profit-maximization problem and derive the equations that define the critical points of the Lagrangian L in this problem.

   Solution: The firm want to maximize $p_y y - w_1 x_1 - w_2 x_2 - w_3 x_3$ subject to $y = x_1(x_2 + x_3)$ and $x_1, x_2, x_3 \geq 0$. We may immediately plug the first constraint to the objective to make life easier.

The Lagrangian is $L(x_1, x_2, x_3, \lambda_1, \lambda_2, \lambda_3) = p_y x_1(x_2 + x_3) - w_1 x_1 - w_2 x_2 - w_3 x_3 + \lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3$. Critical points are then given by:

$$\begin{cases} p_y(x_2 + x_3) = w_1 - \lambda_1 \\ p_y x_1 = w_2 - \lambda_2 \\ p_y x_1 = w_3 - \lambda_3 \\ \lambda_1 x_1 = 0 \\ \lambda_2 x_2 = 0 \\ \lambda_3 x_3 = 0 \end{cases}$$

(b) Show that the Lagrangian L has multiple critical points for any choice of $(p_y, w_1, w_2, w_3) \in \mathbb{R}^4_{++}$.

Solution: $x_1 = x_2 = x_3 = 0$, $\lambda_i = w_i$, $i = 1, 2, 3$ is always a critical point. So is $\lambda_1 = \lambda_2 = 0 \neq \lambda_3$ for $w_3 > w_2$ and $\lambda_1 = \lambda_3 = 0 \neq \lambda_2$ for $w_2 > w_3$. For $w_2 = w_3$ any x_2 and x_3 such that $x_2 + x_3 = \frac{w_1}{p_y}$ is a critical point.

(c) Show that none of these critical points identifies a solution of the profit-maximization problem. Can you explain why this is the case?

Solution: This is the case simply because there is no solution to the profit-maximization problem at all. That is, the profits can be potentially made infinite. To see this, consider moving along the line $x_1 = x_2 = a$, $x_3 = 0$. Profits then are $p_y a^2 - (w_1 + w_2)a$, which grows to infinity as $a \to \infty$.

7. (Sundaram 8.25, page 201) An agent who consumes three commodities has a utility function given by $u(x_1, x_2, x_3) = \sqrt[3]{x_1} + \min\{x_2, x_3\}$. Given an income of I and prices p_1, p_2, p_3, write down the consumer's utility-maximization problem. Can the Weierstraß and/or Kuhn-Tucker theorems be used to obtain and characterize a solution? Why or why not?

Solution: The problem is to maximize $u(x_1, x_2, x_3) = \sqrt[3]{x_1} + \min\{x_2, x_3\}$ subject to $p_1 x_1 + p_2 x_2 + p_3 x_3 \leq I$. The Weierstraß theorem is applicable, as long as prices are strictly positive, but Kuhn-Tucker theorem is not applicable, since u is not a C^1 function (it is not differentiable at points where $x_2 = x_3$). The way to solve it will be to notice that it is never optimal to have $x_2 \neq x_3$ (if, say, $x_2 < x_3$, then cutting down on x_3 and buying some more of x_2 will improve utility). Therefore, we may denote by x the composite good composed of equal quantities of x_2 and x_3, which is going at price

$p = p_2 + p_3$. The problem then becomes to maximize $u(x_1, x) = \sqrt[3]{x_1} + x$ subject to $p_1 x_1 + px \leq I$; now Kuhn-Tucker theorem is applicable.

(a) Consider the Euclidean distance from the origin to the point (x, y) in $\mathbb{R}^2$: $d(x, y) = \sqrt{x^2 + y^2}$. Suppose $d(x, y)$ reaches its global maximum on a compact set X at the point (x^*, y^*), and suppose that $z \to h(z)$ is a monotonically increasing transformation. Where does $F(x, y) \equiv (h \circ d)(x, y)$ attain its global maximum on X?

Solution: h monotonically increasing implies that $d(x^*, y^*) \geq d(x, y) \forall (x, y) \in X \implies F(x^*, y^*) \geq F(x, y) \forall (x, y) \in X$; that is, F attains its global maximum at x^*, y^* as well.

(b) Consider the function $G(x, y) = x^2 + 2y^2 - 6x - 7$. Find the maximum and minimum of G on $\mathbb{R}^2$, if any. Does the Weierstraß theorem apply?

Solution: Weierstraß does not apply, because $\mathbb{R}^2$ is unbounded. Nevertheless, we can find a minimum.

Observe first that $G(x, y)$ can be split into $f(x) = x^2 - 6x$ and $g(y) = 2y^2 - 7$. Both of these functions describe convex parabolae (parabolas? paraboleese? whatever), so it should be clear that the function will have no global max, that it will have a global min, and that G as a whole describes a paraboloid in $\mathbb{R}^3$. Taking first order conditions, we find that

$$2x - 6 = 0 \tag{4}$$
$$4y = 0 \tag{5}$$

so the only critical point is $(3, 0)$. Since both second order conditions are positive, we again see that the function is convex, and that this is therefore a global minimum.

(c) Consider now the curve described by $G(x, y) = 0$. Where does this curve **not** implicitly define y as a function of x? Where does the curve **not** implicitly define x as a function of y? Find the slope of the curve when $x = 2$.

Solution: This curve is the intersection of the xy plane with the paraboloid, which is an ellipse. To see this, note that we can write $G(x, y) = (x - 3)^2 + 2y^2 - 16 = 0$, which clearly defines an ellipse with center $(3, 0)$, and major axis on the x-axis, with length 8.

According to the implicit function theorem, the curve does not define $y(x)$ where $\frac{\partial G}{\partial y} = 4y = 0$, or at $y = 0$. Similarly,

$x(y)$ is not defined where $\frac{\partial G}{\partial x} = 2x - 6 = 0$, or $x = 3$. Notice that these points are the 'top', 'bottom', and 'sides' of our ellipse, where the curve goes just vertical and horizontal.

The implicit function theorem also tells us that, so long as $y(x)$ is defined, $y'(x) = \frac{\partial y}{\partial x_i}(x_1^*, ..., x_n^*) = -\frac{\frac{\partial G}{\partial x_i}(x_1^*,...,x_n^*,y^*)}{\frac{\partial G}{\partial y}(x_1^*,...,x_n^*,y^*)}$. Here, we have $y'(x) = -\frac{2x-6}{4y} = \frac{1}{2y}$ when $x = 2$. Plugging $x = 2$ into $G(x, y) = 0$, we find that $y = \pm\sqrt{\frac{15}{2}}$, so $y'(x) = \pm\frac{\sqrt{30}}{30}$.

(d) Where is the curve $G(x, y) = 0$ closest to the origin? Does Weierstraß apply?

Solution: Weierstraß does apply, because the ellipse defined by the constraint is a compact set and the distance function is continuous.

One way to solve this problem is simply to look at the picture. The ellipse described by $G(x, y) = 0$ clearly comes closest to the origin at the point $(-1, 0)$ and is furthest from it at $(7, 0)$. But we can also get the same answer using the standard Lagrangian method, noting that the objective function is $d(x, y) =.\sqrt{x^2 + y^2}$ (the formula for distance from the origin).

$$L(x, y, \lambda) = \sqrt{x^2 + y^2} - \lambda(x^2 + 2y^2 - 6x - 7) \tag{6}$$

$$\frac{\partial L}{\partial x} = x(x^2 + y^2)^{-\frac{1}{2}} - \lambda(2x - 6) = 0 \tag{7}$$

$$\frac{\partial L}{\partial y} = y(x^2 + y^2)^{-\frac{1}{2}} - 4\lambda y = 0 \tag{8}$$

$$\frac{\partial L}{\partial \lambda} = x^2 + 2y^2 - 6x - 7 = 0 \tag{9}$$

We note that $y = 0$ satsfies the second FOC. Plugging this into the constraint, we find that two critical points are $(-1, 0)$ and $(7, 0)$. Technically, we've treated x as a function of y, and should therefore be concerned about the points where this is not defined (i.e., at $x = 3$). This would give us two more critical points: $(3, 2\sqrt{2})$ and $(3, -2\sqrt{2})$. If we have drawn the ellipse that $G(x, y)$ represents we know that we need not check these points, however - and indeed, we find that

$$F(-1, 0) = 1 \tag{10}$$

$$F(7, 0) = 49 \tag{11}$$

$$F(3, 2\sqrt{2}) = F(3, -2\sqrt{2}) = \sqrt{17} \tag{12}$$

So that $(-1,0)$ is our minimizer and $(7,0)$ is our maximizer.

Comment: We can solve the FOCs to get

$$\lambda = \frac{x}{2x-6} = \frac{y}{4y} \tag{13}$$

or some variant thereof (it depends whether you use $d(x,y)$ as the objective function or x^2+y^2 instead, a perfectly good thing to do). This is fine, so long as you notice that this equation assumes that $y \neq 0$ and $x \neq 3$ - which turn out to be the only candidate solutions! If you take the problem in cases - first, assume that $y \neq 0$ and $x \neq 3$, so that you can use the equation above, and then come back and check what you get when $y = 0$ or $x = 3$ - then this will work. Forgetting to check $y = 0$ and $x = 3$ causes some problems.

(e) Where is the curve $G(x,y) = 0$ farthest from the origin? Does Weierstraß apply?

Solution: See above, where I found both the minimum and the maximum.

(f) Maximize and minimize $d(x,y) = \sqrt{x^2+y^2}$ subject to $G(x,y) \leq 0$. Does Weierstraß apply?

Solution: This is nearly the same problem as above, except that now we have to check all points in the ellipse, in addition to its boundary. When the constraint binds, we have the same ciritical points as above; when it doesn't, we have the critical point $(0,0)$ - the origin itself. Thus, the maximum occurs at $(7,0)$, and the minimum at $(0,0)$. Weierstraß applies because the boundary and interior of the ellipse together are a compact set, and again, the objective function is continuous.

(g) Maximize and minimize $F(x,y) = 2x^2 + 2y^2 + 8$ subject to $G(x,y) \leq 0$.

Solution: This problem is the same as part (f). To see this, note that $F(x,y) = 2[d(x,y)]^2 + 8$, a monotonic transformation of $d(x,y)$. Thus, by the result in part (a), the extremal points must be the same.

Comment: It would seem interesting to investigate why it was that the candidate solutions in parts (d) and (e) were precisely those points where $y(x)$ and $x(y)$ aren't defined, and how this relates to the connection between the implicit function theorem and Lagrange, although it's hard to think of a way to ask this question in a way sufficiently narrow to be coherent and answerable. But this is nevertheless an important point: this is always going to be the case in this kind of problem when the ellipse represented by the constraint set has its center such that at least one of the coordinates of the center is the same as the coordinates of the point you're measuring distance from (for example, in this case we're measuring distance from the origin, and the ellipse is centered at $(3, 0)$. The intuitive way to think of this is simply to note that when things are lined up like this, the circle (or sphere) representing the distance from the origin (or whatever point it is) will always be such that the slope of its level sets will be vertical expanding out to the sides, and horizontal expanding up and down (one former student had a good way of picturing this - 'a balloon within a ballon', if the origin is contained within your constraint set). And an ellipse also has it slope vertical just on the ends of one axis, and horizontal just on the ends of its other axis. And the sphere and the ellipse will touch - and hence be tangent, a necessary condition for an optimum with a binding constraint - right along one of these axes, and thus right where the slopes are either vertical or horizontal. If the center of the constraint set isn't lined up with the origin (or other point) like this, we no longer have this condition, and we will no longer find our optima right where $x(y)$ or $y(x)$ isn't defined.