Math 1c: The Midterm Review

1 What You Have (Hopefully) Learned Thus Far

Thus far, in the first four weeks of Math 1c, we've covered the following topics:

- The topology of $\mathbb{R}^n$: specifically, the definitions of open, closed, and compact sets, the concept of a limit in $\mathbb{R}^n$, and the idea of continuity in $\mathbb{R}^n$. Also, we discussed how the idea of continuity was linked to this idea of open sets, in that a function f is continuous iff f^{-1} sends any open set to an open set.
- The concept of a derivative of a function $f : \mathbb{R}^n \to \mathbb{R}^m$: specifically, the three distinct notions of a partial derivative, a directional derivative, and a total derivative. As well, we discussed how these three definitions are connected to each other.
- Various tools we have for calculating the derivative: specifically, the chain and product rules for functions $f : \mathbb{R}^n \to \mathbb{R}^m$.
- Applications of the derivative: how to use the gradient to find critical points, how to use the Hessian matrix to classify some critical points as minima or maxima, and how to use the method of Lagrange multipliers to find constrained minima and maxima.

These topics are what your midterm will be on! Between my notes, Alden's notes, the class notes, and Apostol, you've got all of the key definitions and theorems reprinted approximately a thousand times; so, if you feel like you're still shaky with them, go back through your notes (or contact me/your TA, and we'll be happy to review them with you!) In lieu of this, I want to instead present some examples of these ideas *in action.* Each of the following eight examples illustrates ideas you may run into on your midterm; I've tried to make these all a bit trickier than anything you'll see on your test, so that if you're comfortable with how to solve these eight problems you should have no difficulties on the midterm. Enjoy!

2 Worked Examples

Problem 1 *Define* $f : \mathbb{R}^n \to \mathbb{R}$ *as the function*

$$f(x_1, \ldots x_n) = \begin{cases} (q_1 \cdot \ldots q_n)^{-1}, & x_i = \frac{p_i}{q_i} \in \mathbb{Q}, \forall i \\ 0, & \textit{otherwise.} \end{cases}.$$

1. *For any fixed constant $a > 0$, what kind of set (i.e open or closed) is*

$$A = \left\{ (x_1, \ldots x_n) : f(x_1, \ldots x_n) \leq \frac{1}{a} \right\}?$$

2. *How about the set*

$$B = \{(x_1, \ldots x_n) : f(x_1, \ldots x_n) \neq 0\}?$$

3. *Using only your answers to the above two questions as justifications, say whether or not f is a continuous function.*

Solution.

1. We claim that the set A above is open. To prove this, we will construct an open ball around any point $\mathbf{x}$ in A such that this open ball lies within A; as this is the definition of open, it will prove that A is an open set.

   To do this, take any point $\mathbf{x}$ in A, and look at the open ball $B_{\mathbf{x}}(1)$ of radius 1 around $\mathbf{x}$. We seek to find a smaller ball $B_{\mathbf{x}}(\epsilon)$ around $\mathbf{x}$ within this ball $B_{\mathbf{x}}(1)$, such that this smaller ball is contained entirely within A.

   How can we do this? Well: there are only finitely many rational numbers $\frac{p}{q}$ such that $q < a$ and $|x_i - \frac{p}{q}| < 1$; therefore, there are only finitely many rational points $\left(\frac{p_1}{q_1}, \dots \frac{p_n}{q_n}\right)$ in our ball $B_{\mathbf{x}}(1)$ such that $q_1 \cdot \ldots \cdot q_n < a$ in this open ball. These points are the only points for which it's even possible that $f\left(\frac{p_1}{q_1}, \dots \frac{p_n}{q_n}\right) > \frac{1}{a}$; therefore, there are at most finitely many points $\mathbf{y}$ in the ball $B_{\mathbf{x}}(1)$ on which $f(\mathbf{y}) > \frac{1}{a}$.

   List these points as $\mathbf{y}_1, \dots, \mathbf{y}_k$, and let $\epsilon = \min\{||\mathbf{x} - \mathbf{y_i}|| : 1 \leq i \leq n\}$. Because there are only finitely many $\mathbf{y}_i$'s and none of them are equal to x, this minimum exists and is > 0. So, draw a ball of radius ϵ around $\mathbf{x}$: by construction, it contains no points $\mathbf{y}$ such that $f(\mathbf{y}) > \frac{1}{a}$. So this ball $B_{\mathbf{x}}(\epsilon)$ is contained entirely within A: thus A is open, as claimed.

2. First, notice that because $f(\mathbf{x}) \neq 0$ if and only if $\mathbf{x} \in \mathbb{Q}^n$, we have that $B = \mathbb{Q}^n$. Therefore, B is neither closed nor open, as the rationals $\mathbb{Q}^n$ and irrationals $(\mathbb{Q}^n)^c$ are both dense in $\mathbb{R}^n$, as proven on your first homework set.

3. So, notice that we can write $B = f^{-1}\left((-\infty, 0) \cup (0, \infty)\right)$: i.e. we can write B as the preimage of an open set. So, because B is not open, we know that f cannot be continuous, as a function f is continuous if and only if f^1 takes open sets to open sets.

Problem 2 *For*

$$f(x_1, \dots x_n) = \sin\left(x_1^2 + \dots x_n^2\right),$$

find the sets

$$S_1 = \{\mathbf{x} : f(\mathbf{x}) > 0\},$$
$$S_2 = \{\mathbf{x} : f(\mathbf{x}) < 0\}, \text{ and}$$
$$S_3 = \{\mathbf{x} : f(\mathbf{x}) = 0\}.$$

1. *What kinds of sets (i.e. open/closed, bounded or unbounded, compact or noncompact) are these? Geometrically, what do they look like?*

2. *Show that $(0, \dots 0)$ is a critical point of f. Using the Hessian, determine what kind of critical point it is. Geometrically, why is this obvious?*

Solution.

1. So: notice first that f is a continuous function, and that

$$\begin{aligned} S_1 &= f^{-1}\left((0,\infty)\right), \\ S_2 &= f^{-1}\left((-\infty,0)\right), \text{ and} \\ S_3 &= f^{-1}\left(\{0\}\right). \end{aligned}$$

   Therefore, because the inverse f^{-1} of any continuous function f sends open sets to open sets, we know that because $(-\infty, 0)$ and $(0, \infty)$ are both open, their images S_1 and S_2 under f^{-1} must also both be open. As well, because S_3 is just all of the points in $\mathbb{R}^n$ not in either S_1 or S_2, we can write $S_3 = (S_1 \cup S_2)^c$, and thus see that S_3 is closed.

   More explicitly, we can see by solving for $\mathbf{x}$ that

$$\begin{aligned} S_1 &= \{\mathbf{x} : \sum_{i=1}^n x_i^2 \in (2k\pi, (2k+1)\pi), \text{ for some } k\} \\ S_2 &= \{\mathbf{x} : \sum_{i=1}^n x_i^2 \in ((2k+1)\pi, (2k+2)\pi), \text{ for some } k\}, \text{ and} \\ S_3 &= \{\mathbf{x} : \sum_{i=1}^n x_i^2 = k\pi, \text{ for some } k\}. \end{aligned}$$

   Therefore, as there are points in each of these sets of arbitrarily large size, none of these sets are bounded. As a consequence, because a set is compact in $\mathbb{R}^n$ iff it is closed and bounded, none of these sets are compact.

2. To see that $(0, \ldots 0)$ is a critical point of f, we merely need to look at f's gradient, which is

$$\nabla(f) = \left(2x_1 \cdot \cos\left(\sum_{k=1}^n x_i^2\right), 2x_1 \cdot \cos\left(\sum_{k=1}^n x_i^2\right), \ldots \right).$$

   When $\mathbf{x} = (0, \ldots 0)$, this gradient is 0; therefore, $(0, \ldots 0)$ is a critical point.

   To determine what kind of critical point this is, we can use the Hessian, which is

$$\begin{bmatrix} 2\cos\left(\sum x_i^2\right) - 4x_1^2 \sin\left(\sum x_i^2\right) & -4x_1x_2 \sin\left(\sum x_i^2\right) & -4x_1x_3 \sin\left(\sum x_i^2\right) & \ldots \\ -4x_2x_1 \sin\left(\sum x_i^2\right) & 2\cos\left(\sum x_i^2\right) - 4x_2^2 \sin\left(\sum x_i^2\right) & -4x_2x_3 \sin\left(\sum x_i^2\right) & \ldots \\ -4x_3x_1 \sin\left(\sum x_i^2\right) & -4x_3x_2 \sin\left(\sum x_i^2\right) & 2\cos\left(\sum x_i^2\right) - 4x_3^2 \sin\left(\sum x_i^2\right) & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

   At $(0, \ldots, 0)$, this simplifies considerably to

$$\begin{bmatrix} 2 & 0 & \ldots & 0 \\ 0 & 2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 2 \end{bmatrix},$$

   which has 2 as an eigenvalue n times. As all of the eigenvalues of the Hessian are positive, and there are n of them, we can conclude that $(0, \ldots, 0)$ is a local minimum.

   If we didn't use the Hessian, we could also trivially note that because our function $f(x_1, \ldots x_n) = \sin\left(x_1^2 + \ldots x_n^2\right)$,, for any value of $\mathbf{x}$ in the ball $B_0(\sqrt{\pi})$, we have $0 \leq \sum x_i^2 < \pi$ and therefore $f(\mathbf{x}) = \sin(\sum x_i^2) \geq 0$. Therefore, $(0, \ldots, 0)$ is a relative minimum within this ball.

Problem 3 *Let $f : \mathbb{C}^2 \to \mathbb{C}$ be defined as*

$$f(z_1, z_2) = z_1 \cdot z_2.$$

1. *Thinking of $\mathbb{C}$ as $\mathbb{R}^2$ via the map $x + iy \mapsto (x, y)$, interpret this as a function from $\mathbb{R}^4$ to $\mathbb{R}^2$.*
2. *Thinking of $Re(f), Im(f)$ as functions $\mathbb{R}^4 \to \mathbb{R}$, find the directional derivatives of $Re(f)$ and $Im(f)$ at the point $(0, 1, 2, 3)$ in the direction $(4, 5, 6, 38)$.*

Solution.

1. So: if we write $z_1 = x_1 + iy_1, z_2 = x_2 + iy_2$, we have

$$\begin{aligned} f(z_1, z_2) &= z_1 \cdot z_2 \\ &= (x_1 + iy_1) \cdot (x_2 + iy_2) \\ &= x_1x_2 - y_1y_2 + i(x_1y_2 + x_2y_1). \end{aligned}$$

   Therefore, if you regard a point (x_1+iy_1, x_2+iy_2) in $\mathbb{C}^2$ as equivalent to the point $(x_1, y_1, x_2, y_2) \in \mathbb{R}^4$, we can think of our function as the map

$$f(x_1, y_1, x_2, y_2) = (x_1x_2 - y_1y_2, x_1y_2 + x_2y_1)$$

   from $\mathbb{R}^4$ to $\mathbb{R}^2$.

2. So, notice that if we're still thinking of f as a function $\mathbb{R}^4 \to \mathbb{R}^2$, we have

$$\begin{aligned} Re(f)(x_1, y_1, x_2, y_2) &= x_1x_2 - y_1y_2, \text{ and} \\ Im(f)(x_1, y_1, x_2, y_2) &= x_1y_2 + x_2y_1. \end{aligned}$$

   Therefore, we have that

$$\begin{aligned} \nabla(Re(f))(x_1, y_1, x_2, y_2) &= (x_2, -y_2, x_1, -y_1), \text{ and} \\ \nabla(Im(f))(x_1, y_1, x_2, y_2) &= (y_2, x_2, y_1, x_1). \end{aligned}$$

   Consequently, we can write the directional derivative of $Re(f)$, $Im(f)$ at the point $(0, 1, 2, 3)$ in the direction $(4, 5, 6, 38)$ as simply the dot product of the normalized vector $(4, 5, 6, 38) \cdot \frac{1}{||(4,5,6,38)||} = (4, 5, 6, 38) \cdot \frac{1}{39}$, and the gradients of these functions at $(0, 1, 2, 3)$:

$$\begin{aligned} \nabla(Re(f))(0, 1, 2, 3) \cdot \frac{(4, 5, 6, 38)}{39} &= (2, -3, 0, -1) \cdot \frac{(4, 5, 6, 38)}{39} \\ &= \frac{8 - 15 - 38}{39} \\ &= -\frac{45}{39}, \text{ and} \\ \nabla(Im(f))(0, 1, 2, 3) \cdot \frac{(4, 5, 6, 38)}{39} &= (3, 2, 1, 0) \cdot \frac{(4, 5, 6, 38)}{39} \\ &= \frac{12 + 10 + 6}{39} \\ &= \frac{28}{39}. \end{aligned}$$

Problem 4 *Define $f : \mathbb{R}^n \to \mathbb{R}$ as the function*

$$f(\mathbf{x}) = \mathbf{x}^T \cdot A \cdot \mathbf{x},$$

where A is the matrix

$$A = \begin{bmatrix} 1 & 2 & 2 & \dots & 2 \\ 2 & 1 & 2 & \dots & 2 \\ 2 & 2 & 1 & \dots & 2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 2 & 2 & 2 & \dots & 1 \end{bmatrix},$$

and n is even.

1. *Find the directional derivative of f at $(1, \dots 1)$ in the direction $(1, -1, 1, -1, \dots 1, -1)$.*
2. *Find all of the critical points of f, and classify them.*

Solution.

1. So: notice that if we explicitly write out what f does to a vector $(x_1, \dots x_n)$, we can see that

$$\begin{aligned} f(x_1, \dots x_n) &= \begin{bmatrix} x_1 & x_2 & x_3 & \dots & x_n \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 2 & \dots & 2 \\ 2 & 1 & 2 & \dots & 2 \\ 2 & 2 & 1 & \dots & 2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 2 & 2 & 2 & \dots & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} \\ &= \begin{bmatrix} x_1 & x_2 & x_3 & \dots & x_n \end{bmatrix} \cdot \begin{bmatrix} x_1 + \sum_{i \neq 1} 2x_i \\ x_2 + \sum_{i \neq 2} 2x_i \\ x_3 + \sum_{i \neq 3} 2x_i \\ \vdots \\ x_n + \sum_{i \neq n} 2x_i \end{bmatrix} \\ &= \left(\sum_{i=1}^{n} x_i^2 \right) + \left(\sum \sum_{(i,j): i \neq j} 4 \cdot x_i x_j \right). \end{aligned}$$

   This tells us that the gradient of f is given by the vector

$$\nabla(f) = \left(\left(2x_1 + \sum_{i \neq 1} 4x_i \right), \left(2x_2 + \sum_{i \neq 2} 4x_i \right), \dots \right);$$

   so, at $(1, \dots 1)$, this is just $(4n-2, 4n-2, \dots 4n-2)$. Therefore, the directional derivative of f in the direction $(1, -1, 1, -1, \dots 1, -1)$ is just

$$\nabla(f)\Big|_{(1,\dots1)} \cdot \frac{(1, -1, 1, -1, \dots 1, -1)}{||(1, -1, 1, -1, \dots 1, -1)||} = \frac{(4n-2) - (4n-2) + \dots + (4n-2) - (4n-2)}{\sqrt[n]{n}} = 0.$$

2. So: to find the critical points of f, we return to the gradient. We know that $\nabla(f) = (0, \dots 0)$ iff $2x_i + \sum_{j \neq i} 4x_j$ is zero, for every i: i.e. whenever there is a solution to the system of linear

equations

$$\begin{bmatrix} 2 & 4 & 4 & \dots & 4 \\ 4 & 2 & 4 & \dots & 4 \\ 4 & 4 & 2 & \dots & 4 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 4 & 4 & 4 & \dots & 2 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = 0.$$

This matrix is clearly rank n: therefore, the only combination of its rows that equals the zero vector is the trivial one $(x_1, \dots x_n) = (0, \dots, 0)$.

So the only critical point of f is at the origin. What kind of critical point is the origin?

Well: to identify this point, we turn to the Hessian of f, which (by taking derivatives) we can see is just the matrix

$$\begin{bmatrix} 2 & 4 & 4 & \dots & 4 \\ 4 & 2 & 4 & \dots & 4 \\ 4 & 4 & 2 & \dots & 4 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 4 & 4 & 4 & \dots & 2 \end{bmatrix}$$

What are the eigenvalues of this matrix? In other words, for what values of λ does

$$\begin{bmatrix} 2-\lambda & 4 & 4 & \dots & 4 \\ 4 & 2-\lambda & 4 & \dots & 4 \\ 4 & 4 & 2-\lambda & \dots & 4 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 4 & 4 & 4 & \dots & 2-\lambda \end{bmatrix}$$

not have full rank? Well: if $\lambda = -2$, this matrix is simply the matrix of all 4's, and thus has rank 1; therefore, -2 is an eigenvalue of this matrix of multiplicity $n-1$. As well, if $\lambda = 4n-2$, we have that our matrix is of the form

$$\begin{bmatrix} 4n-4 & 4 & 4 & \dots & 4 \\ 4 & 4n-4 & 4 & \dots & 4 \\ 4 & 4 & 4n-4 & \dots & 4 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 4 & 4 & 4 & \dots & 4n-4 \end{bmatrix}.$$

If we add all of the rows of this matrix together we get the 0-vector: therefore, this matrix does not have full rank, and thus $4n-2$ must also be an eigenvector of our matrix. Therefore, $H(f)$ has both positive and negative eigenvectors at (0,...0), and thus $(0, \dots 0)$ must be a saddle point.

Problem 5 *Let* $g : \mathbb{R}^4 \to \mathbb{R}$ *be defined as*

$$g(w, x, y, z) = wz - xy,$$

and $f : \mathbb{R}^2 \to \mathbb{R}^4$ *be defined as*

$$f(a, b) = (a, b, \lambda a, \lambda b),$$

for some constant λ.

1. *Use the chain rule to find the total derivative of $g \circ f$.*

2. *In one sentence, explain why the above result is obvious. (You may need to use some commas and/or a semicolon.)*

Solution. 1. So, we know that both g and f are continuous functions on all of their domains; therefore, we know that their composition is continuous everywhere. Therefore, we know that the total derivative of $g \circ f$ is just given by the partial derivatives of $g \circ f$: i.e. $T(g \circ f) = D(g \circ f)$. So, we can use the chain rule:

$$\begin{aligned}
D(g \circ f)\Big|_{\mathbf{a}} &= D(g)\Big|_{f(a,b)} \cdot D(f)\Big|_{(a,b)} \\
&= \begin{bmatrix} d & -c & b & -a \end{bmatrix}\Big|_{f(a,b)} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \lambda & 0 \\ 0 & \lambda \end{bmatrix} \\
&= \begin{bmatrix} \lambda b & -\lambda a & b & -a \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \lambda & 0 \\ 0 & \lambda \end{bmatrix} \\
&= \lambda b - \lambda b + \lambda a - \lambda a \\
&= 0.
\end{aligned}$$

2. This is trivially true because the function g is just the determinant of the matrix $\begin{pmatrix} w & x \\ y & z \end{pmatrix}$, the function f outputs a matrix of rank 1, and the determinant of a 2×2 matrix of rank 1 is always the constant 0, (which has derivative 0.)

Problem 6 *Let $g : \mathbb{R}^3 \to \mathbb{R}^2$ be defined as*

$$g(a, b, c) = \left(\frac{b}{\sqrt{c}}, \frac{b}{\sqrt{a}} \right),$$

and $f : \mathbb{R}^2 \to \mathbb{R}^3$ be defined as

$$f(x, y, z) = \left(x^2, xy, y^2 \right).$$

1. *Use the chain rule to find the total derivative of $g \circ f$, at any point (x, y) where $x, y > 0$.*

2. *In one sentence, explain why the above result is exactly what you'd expect.*

Solution. 1. So, we know that both g and f are continuous functions on all of their domains; therefore, we know that their composition is continuous everywhere. Therefore, we know that the total derivative of $g \circ f$ is just given by the partial derivatives of $g \circ f$: i.e. $T(g \circ f) =$

$D(g \circ f)$. So, we can use the chain rule:

$$\begin{aligned}
D(g \circ f)\Big|_{\mathbf{a}} &= D(g)\Big|_{f(x,y)} \cdot D(f)\Big|_{(\mathbf{x},\mathbf{y})} \\
&= \begin{bmatrix} 0 & \frac{1}{\sqrt{c}} & -\frac{b}{2c^{3/2}} \\ -\frac{b}{2a^{3/2}} & \frac{1}{\sqrt{a}} & 0 \end{bmatrix}\Bigg|_{f(x,y)} \cdot \begin{bmatrix} 2x & 0 \\ y & x \\ 0 & 2y \end{bmatrix} \\
&= \begin{bmatrix} 0 & \frac{1}{y} & -\frac{x}{2y^2} \\ -\frac{y}{2x^2} & \frac{1}{x} & 0 \end{bmatrix}\Bigg|_{f(x,y)} \cdot \begin{bmatrix} 2x & 0 \\ y & x \\ 0 & 2y \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.
\end{aligned}$$

2. We expect this because whenever $x, y > 0$, $g \circ f$ is just the identity function $\mathbb{R}^2 \to \mathbb{R}^2$.

Problem 7 *Using the method of Lagrange multipliers, prove the harmonic mean-geometric mean inequality. In other words, show that for any $x_1, \ldots x_n > 0$, we have*

$$\frac{n}{\frac{1}{x_1} + \ldots + \frac{1}{x_n}} \leq (x_1 \cdot \ldots x_n)^{1/n}.$$

Solution. This problem illustrates one of the more nonobvious uses of the method of Lagrange multipliers: proving inequalities! (See my lecture notes from week 4 for another example of how to do this.) Specifically, to prove an inequality like the above using Lagrange multipliers, we can simply do the following:

1. Pick one of the functions above –say, $(x_1 \cdot \ldots x_n)^{1/n}$ – and choose it to be the "constraining" function g: i.e. define $g(\mathbf{x}) = (x_1 \cdot \ldots x_n)^{1/n}$.

2. Now, choose any constant c, and look at all of the points $\mathbf{x}$ such that $g(\mathbf{x}) = c$. Our goal is now to prove that amongst the set $S = \{\mathbf{x} : g(\mathbf{x} = c)\}$ of all points constrained by $g(\mathbf{x}) = c$, we always have $\frac{n}{\frac{1}{x_1}+\ldots+\frac{1}{x_n}} \leq g(\mathbf{x}) = c$: in other words, that the maximum value of $\frac{n}{\frac{1}{x_1}+\ldots+\frac{1}{x_n}}$ on this constrained set is $\leq c$. Define $f(\mathbf{x})$ as this expression $\frac{n}{\frac{1}{x_1}+\ldots+\frac{1}{x_n}}$.

3. So, we've reduced our problem to one we *can* solve with Lagrange multipliers! In specific, we have

   - $g(\mathbf{x}) = (x_1 \cdot \ldots x_n)^{1/n} = c$, for any positive number c,
   - $f(\mathbf{x}) = \frac{n}{\frac{1}{x_1}+\ldots+\frac{1}{x_n}}$,
   - and we want to maximize $f(\mathbf{x})$ on the set $S = \{\mathbf{x} : g(\mathbf{x} = c)\}$, and then finally show that that maximum is $\leq c$.

So, this is totally doable! To do this via the method of Lagrange multipliers, we merely need to check that the ∇g's are never a linearly dependent set. As there is only one constraint g, this is just equivalent to checking that ∇g is never identically 0; so, because

$$\begin{aligned}
\nabla(f) &= \left(\frac{\partial}{\partial x_1}\left((x_1\cdot\ldots\cdot x_n)^{1/n}\right), \frac{\partial}{\partial x_2}\left((x_1\cdot\ldots\cdot x_n)^{1/n}\right), \ldots\right) \\
&= \left(\frac{\partial}{\partial x_1}\left((x_1)^{1/n}\cdot(x_2\cdot\ldots\cdot x_n)^{1/n}\right), \frac{\partial}{\partial x_2}\left((x_2)^{1/n}\cdot(x_1\cdot x_3\cdot\ldots\cdot x_n)^{1/n}\right), \ldots\right) \\
&= \left(\frac{1}{n}(x_1)^{1/n-1}\cdot(x_2\cdot\ldots\cdot x_n)^{1/n}, \frac{1}{n}(x_2)^{1/n-1}\cdot(x_1\cdot x_3\cdot\ldots\cdot x_n)^{1/n}, \ldots\right) \\
&= \left(\frac{1}{n\cdot x_1}(x_1\cdot\ldots\cdot x_n)^{1/n}, \frac{1}{n\cdot x_2}(x_1\cdot\ldots\cdot x_n)^{1/n}, \ldots\right),
\end{aligned}$$

we can see that this is never 0 at any point $(x_1, \ldots x_n)$ where all of the x_1's are > 0.

So, amongst the space of all points with positive coördinates, the method of Lagrange multipliers says that any critical point of $f(\mathbf{x})$ must occur at a point where

$$\nabla(f)(\mathbf{x}) = \lambda\cdot\nabla(g)(\mathbf{x}),$$

for some constant λ.

Because

$$\begin{aligned}
\nabla(f) &= \left(\frac{\partial}{\partial x_1}\left(\frac{n}{\frac{1}{x_1}+\ldots+\frac{1}{x_n}}\right), \frac{\partial}{\partial x_2}\left(\frac{n}{\frac{1}{x_1}+\ldots+\frac{1}{x_n}}\right), \ldots\right) \\
&= \left(\frac{1}{x_1^2}\cdot\frac{n}{\left(\frac{1}{x_1}+\ldots+\frac{1}{x_n}\right)^2}, \frac{1}{x_2^2}\cdot\frac{n}{\left(\frac{1}{x_1}+\ldots+\frac{1}{x_n}\right)^2}, \ldots\right),
\end{aligned}$$

we will have $\nabla(f)(\mathbf{x}) = \lambda\cdot\nabla(g)(\mathbf{x})$ whenever

$$\frac{1}{x_i^2}\cdot\frac{n}{\left(\frac{1}{x_1}+\ldots+\frac{1}{x_n}\right)^2} = \lambda\cdot\frac{1}{n\cdot x_i}(x_1\cdot\ldots\cdot x_n)^{1/n},$$

for every i.

Multiplying both sides by $\lambda^{-1}\cdot n\cdot x_i^2\cdot(x_1\cdot\ldots\cdot x_n)^{-1/n}$, we can see that this will hold whenever

$$\frac{n^2}{\lambda(x_1\cdot\ldots\cdot x_n)^{1/n}\cdot\left(\frac{1}{x_1}+\ldots+\frac{1}{x_n}\right)^2} = x_i;$$

in other words, when all of the coördinates are equal to each other! Because our constraint is that $g(\mathbf{x}) = (x_1\cdot\ldots\cdot x_n)^{1/n} = c$, we can see that this will uniquely happen at the point $(c, c, \ldots, c)$.

What kind of point is $(c, c, \ldots, c)$ – a maxima or a minima, or neither? How can we determine this?

In recitation/office hours/when talking to many of you, people often said "the Hessian" as an answer here! **NO. DO NOT USE THE HESSIAN WHEN USING LAGRANGE MULTIPLIERS. BAD STUDENTS.** Mostly kidding, but seriously, don't use the Hessian; it only applies when you're trying to find *unconstrained* maxima and minima! (For a quick example that shows how the Hessian will lie to you when you're using Lagrange multipliers: consider the function $f(x, y, z) = x^2 + y^2 - z^2$ and the constraint $g(x, y, z) = x - z = 0$, i.e. the plane $x = z$. on this plane, the method of Lagrange multipliers will tell you that the points $(x, 0, x)$ are your critical points, and inspection of the graph will tell you that these are minima on our constrained set: increasing y will only increase f, and increasing the x/z coördinate will do nothing, as $x = z$ is our constraint. However, the Hessian of f is the matrix $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$, which corresponds to a saddle point because it has both positive and negative eigenvalues! DOOM. So, don't do this!)

So, if we can't use the Hessian, what can we do? Well, there are usually two situations we'll find ourselves in:

- Often, the set S of constrained points will be a closed and bounded set. This means it will be compact! In this case, if our function f is continuous, it must attain its absolute minima and maxima on this set, and these points will be critical points of our function – and thus picked up by the method of Lagrange multipliers! Therefore, in this case, **after noting that the set S is compact in your proof**, you can just evaluate f on all of your critical points, and conclude that the largest value is the absolute maxima and the smallest one is the absolute minima.

- Sometimes, however, your set S of constrained points will not be a bounded set. That, in fact, is what we have going on in this situation – the set S is made up out of points $\mathbf{x}$ with geometric mean c, which includes points of the form $(x_1, \frac{c}{x_1}, c, \ldots c)$ for arbitrarily large x_1. What do we do here? Well, what you can do is just arbitrarily "cut off" S to some really big but closed and bounded subset S': in this case, we can consider the closed and bounded set $S' = \{\mathbf{x} : g(\mathbf{x}) = c, \text{ and } x_i \leq \kappa^n, \text{ for all } i\}$, say, for some really big number κ. Then, to tell if our point $(c, \ldots c)$ is an absolute minima or maxima, we just need to compare it to all of the critical points on this set – i.e all of the **boundary points**!

So: at any point on the boundary of the above set $S' = \{\mathbf{x} : g(\mathbf{x}) = c, \text{ and } x_i \leq \kappa^n, \text{ for all } i\}$, we have $x_i = \kappa^n$ for some coördinate x_i. Therefore, because $(x_1 \cdot \ldots x_n)^{1/n} = c$, we know that, by dividing both sides by x_i, we must have

$$(x_2 \cdot \ldots x_n)^{1/n} = \frac{c}{\kappa},$$

and thus that at least one coördinate x_j must be such that $x_j < \left(\frac{c}{\kappa}\right)^{n/(n-1)}$. In this case, we have

$$f(\mathbf{x}) = \frac{n}{\frac{1}{x_1} + \ldots + \frac{1}{x_n}} \geq \frac{n}{\frac{1}{x_j}} = \frac{n}{(\kappa/c)^{n/(n-1)}} = \frac{n \cdot c^{n/(n-1)}}{\kappa^{n/(n-1)}}.$$

For really really big values of κ, this goes to 0 and therefore is $< c$; therefore, all of the values on the boundary of this set S are less than the value attained by $f(c, \ldots c) = \frac{n}{\frac{1}{c} + \ldots + \frac{1}{c}} = c$. Thus, we've proven that $(c, \ldots c)$ is an absolute maxima on our specially constrained set S': therefore, because it's the *only* critical point on all of S, we know that it's an absolute maxima on all of S! Therefore, we've proven that amongst all of the points $\mathbf{x}$ with geometric mean c, the function $f(\mathbf{x})$ is always $\leq$ its maxima, c!

Therefore, we've proven that the harmonic mean is always less than the geometric mean.

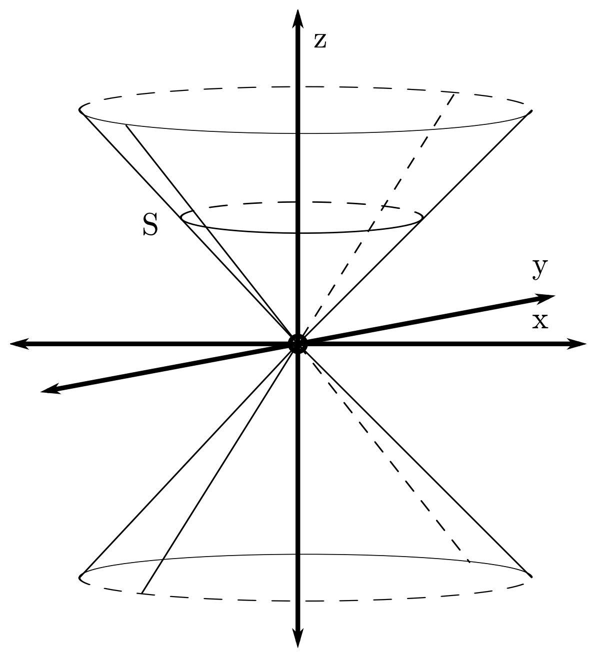
Problem 8 *Find the closest point to* $(1, 1, 1)$ *in the set*

$$A = \{(x, y, z) : x^2 + y^2 = 4, z = 2\}.$$

Solution. This example, hopefully, should be much easier than our earlier problem. So: what is our setup?

1. Well, it's pretty clear that the function we want to minimize is the function $f'(x, y, z)$ that outputs distance from the point $(1, 1, 1)$: i.e. $f'(x, y, z) = \sqrt{(x-1)^2 + (y-1)^2 + (z-1)^2}$. If you want to be clever, though, you can notice that (because distances are always nonnegative numbers,) the distance function will achieve its minima and maxima at the same places that the square of the distance function does – i.e. that we can simply try to minimize the function $f(x, y, z) = (x-1)^2 + (y-1)^2 + (z-1)^2$. There's not too much difference here, but avoiding square roots will make taking derivatives easier on you.

2. The constraint functions are perhaps harder to find. Specifically: what constraints will yield the set $A = \{(x, y, z) : x^2 + y^2 = 4, z = 2\}$? Well: graphically, this set is a circle of radius 2 around the z-axis, in the plane $z = 2$: in other words, it's the following **conic section**[1]:



Therefore, we can write S as the intersections of the level sets

$$g_1(x, y, z) = x^2 + y^2 - z^2 = 0,$$
$$g_2(x, y, z) = z = 2.$$

So, we've phrased our problem in the language of Lagrange multipliers: we have a function f we want to minimize with respect to the two constraints $g_1(x, y, z) = 0$, $g_2(x, y, z) = 2$. To do this, we first check that $\nabla(g_1)$ and $\nabla(g_2)$ will always be linearly independent, for any (x, y, z):

$$\nabla(g_1) = (2x, 2y, -2z),$$
$$\nabla(g_2) = (0, 0, 1).$$

Because the points x, y are constrained such that $x^2 + y^2 = 4$, we know that we cannot have both x and y equal to zero at the same time: therefore, there are no points in our set S at which these

[1] A **conic section** is simply one of the curves you can get by intersecting a cone (the shape sketched by the graph of $x^2 + y^2 = z^2$, up to various constants) with a plane. Conic sections can be either parabolas, hyperbolas, circles, ellipses, or just a single point.

curves are linearly dependent, and we can thus use the method of Lagrange multipliers to identify all possible critical points of f on S! To do this, we simply note that we're looking for points (x, y, z) where there are constants λ_1, λ_2 such that

$$
\begin{aligned}
&\nabla(f) = \lambda_1 \cdot \nabla(g_1) + \lambda_2 \cdot \nabla(g_2) \\
\Rightarrow &(2x-2, 2y-2, 2z-2) = \lambda_1 \cdot (2x, 2y, -2z) + \lambda_2 \cdot (0, 0, 1) \\
\Rightarrow &2x - 2 = 2\lambda_1 x, \\
&2y - 2 = 2\lambda_1 y, \text{ and} \\
&2z - 2 = -2\lambda_1 z + 2\lambda_2.
\end{aligned}
$$

So, if we multiply the equations equations $2x - 2 = 2\lambda_1 x$ and $2y - 2 = 2\lambda_1 y$ by y and x, respectively, we get that $2\lambda_1 xy = 2xy - 2x = 2xy - 2y$: i.e. that $x = y$. Because we know that $z = 2$ by default according to our set S, and because $x^2 + y^2 = 4$, we then know that the only critical points of our set occur at $(-\sqrt{2}, -\sqrt{2}, 2)$ and $(\sqrt{2}, \sqrt{2}, 2)$.

So, it suffices to classify these points. Because our set of constrained points S is a circle, it's a closed and bounded set; therefore, f must attain its absolute minima and maxima on this set, and these points must be amongst our critical points! Therefore, to find our minima, we just need to plug in both $(-\sqrt{2}, -\sqrt{2}, 2)$ and $(\sqrt{2}, \sqrt{2}, 2)$ into f. As $(\sqrt{2}, \sqrt{2}, 2)$ is clearly the closer of the two points to $(1, 1, 1)$, we can safely conclude that this is the closest point in S to (1,1,1).