MA 1C RECITATION 04/23/15

1. Lagrange Multipliers

Lagrange multipliers are a method to optimize a function $f(\overline{x})$ under some constraint $g(\overline{x}) = c$ for some constant c. It is a useful technique that is almost always faster doing things the "long way," that is, by finding the critical points, seeing which ones fit the constraint (or at least are close to it), and seeing which one is the maximum or minimum.

The reason why Lagrange multipliers work is a pretty clever application of the gradient and it's easy and satisfying to see how *geometric* these results are.

I'll state a commonly used version of the Lagrange multiplier theorem. More general versions are available, but I find that this relatively simple version helps illustrate most of the essential features.

Theorem 1.1. *Suppose that the constraint equation $g(\overline{x}) = c$ is **nonsingular**, that is, the function g is differentiable and $\nabla g \neq 0$ at all points in the set $\{g = c\}$. If $\overline{x}$ is a maximizing (or minimizing) input, then $\overline{x}$ also satisfies the equation*

$$(\nabla f)(\overline{x}) = \lambda(\nabla g)(\overline{x})$$

for some scalar λ.

In other words, for, say, the three-dimensional case, we can solve the system of equations

$$\begin{aligned} g(x, y, z) &= c \\ \frac{\partial f}{\partial x} &= \lambda \frac{\partial g}{\partial x} \\ \frac{\partial f}{\partial y} &= \lambda \frac{\partial g}{\partial y} \\ \frac{\partial f}{\partial z} &= \lambda \frac{\partial g}{\partial z} \end{aligned}$$

to find a list of f-inputs to check ("checkpoints"). Then we can plug these into f to get a set of f-outputs, and one of these checkpoints will be the maximum (or minimum) if it exists.

What's the geometric intuition here? Suppose that that you at some point in the domain of f. To increase the function, you want to follow the gradient. However, if you are constrained in some region (that must satisfy $g = c$), then we can imagine ourselves being restricted to this region, but pulled along by gradient (think of it as like a force vector) in its direction, at least, as much as we can while staying in the region. Even though the gradient might direct you outside constrained region, you will be pulled in the component of the gradient that keeps you inside the constrained region, and you will continue to move (increase) until the gradient points in a direct that is perpendicular to the constrained region, at which point you are at a local maximum.

Date: April 23, 2015.

1.1. Algebra Tips for Lagrange Multipliers.

- **Solving for λ and equating the results** is usually the fastest way to start, but sometimes a clever trick will work faster.
- **Be careful when you divide by 0!** If you solve this (or any system of equations), where you have to divide by an expression, it is *critical* that you consider a *separate* case where that expression is zero. For instance, the equation $AB = AC$ implies that $B = C$ *or* $A = 0$ (or both). You will miss check points if you don't do this!
- **Finding "too many" checkpoints** is not a problem, as long as they all satisfy the constraint $g = c$. When we check the outputs, we can quickly rule out the points that are not extrema. However, you must not miss any potential checkpoints!

1.2. Presentation Tips.

- **Make a list of (or draw a box around) all of your "checkpoints"** once you complete your problem. In other words, make it absolutely clear all the f-inputs that you checked, along with their outputs. Lagrange multiplier calculation can get messy quickly, so don't lose points unnecessarily!
- **Clearly indicate the cases considered** (e.g. by underlining them) to make sure that the grader knows which cases you considered.

Example 1. What is maximum and minimum of $f(x,y) = x^2 + y^3$ given the constraint $g(x,y) = x^4 + y^6 = 2$.

Solution. Check that $g(x,y) = 2$ is a nonsingular constraint. By applying the method of Lagrange multipliers, we have the system of equations

$$\begin{aligned} x^4 + y^6 &= 2 \qquad (*) \\ 2x &= \lambda 4x^3 \\ 3y^2 &= \lambda 6y^5 \end{aligned}$$

We want to divide by x and by y to solve for λ, so we need to consider the cases where x or y is zero.

Case $x = 0$: In this case, $(*)$ tells us that $y^6 = 2$ and $y = \pm\sqrt[6]{2}$, so we have the checkpoints $(0, \pm\sqrt[6]{2})$.

Case $y = 0$: Here, $(*)$ says $x^4 = 2$, so $x = \pm\sqrt[4]{2}$ and so we have checkpoints $(\pm\sqrt[4]{2}, 0)$.

Case $x \neq 0$ and $y \neq 0$: Since x and y are both nonzero, we can divide by x and y to see that $\lambda = 1/(2x^2) = 1/(2y^3)$, so $x^2 = y^3$. Then $(*)$ says that $x^4 + x^4 = 2$, so $x = \pm 1$ and $y^3 = x^2 = 1$, so we have checkpoints $(\pm 1, 1)$.

Now that we have all the checkpoints, let's check their outputs under f:

$$\begin{aligned} f(0, \sqrt[6]{2}) &= \sqrt{2} \\ f(0, -\sqrt[6]{2}) &= -\sqrt{2} \qquad \text{(min)} \\ f(\pm\sqrt[4]{2}, 0) &= \sqrt{2} \\ f(\pm 1, 1) &= 2 \qquad \text{(max)}. \end{aligned}$$

□

Remark 1.2. Observe that if we didn't consider the cases of $x = 0$ and $y = 0$ separately, we would not have found the minimum.

We could also have solved the third case differently. When $x^2 = 1$, we could have found y by solving $x^4 + y^6 = 2$, so we'd get "extra" checkpoints $(\pm 1, -1)$. These don't satisfy the full Lagrange system of equations because $x^2 \neq y^3$, but since they do satisfy the constraint equation, the theorem says that the checking process will show that they are not extrema. *Moral:* Having extra checkpoints is OK, as long as they satisfy the original constraint and that you check them at the end.

We could also have solved the example differently. For instance, if we divided by $x - 1$, then we need to check the case where $x = 1$. This also applies to more general situations. For example, if we divide by $y - z^3$, then we need to separately check the case when $y = z^3$.

Example 2. (Failure of the Lagrange Multiplier Method) The method of Lagrange multipliers doesn't always work, in particular, when the hypotheses of the theorem fail, you may not find the extrema.

A simple example is given by trying to find extrema of $f(x, y) = x^2 + y^2$ subject to the constraing $g(x, y) = x^2 - (y - 1)^3 = 0$. There is a local minimum at (0,1) (this even turns out to be the global minimum), but we have $\nabla g(0, 1) = 0$, so there does not exist a λ such that $\nabla f(0, 1) = (0, 2) = \lambda \cdot 0$.

Remark 1.3. A three-dimensional example of where the Lagrange multiplier method fails is given in Apostol, p. 317.

2. Multiple Integrals

Just like in the one-variable case, integrals are defined using step functions, but now the step functions are constant on a finite number of rectangles, instead of intervals. More precisely, we say that f **is integrable with integral** I—or $\int f = I$ for short—if there is some *unique* constant I if for every pair of step functions s and t such that $s(\mathbf{x}) < f(\mathbf{x}) < t(\mathbf{x})$ for all $\mathbf{x}$ in the domain of integration R, we have $\int_R s \leq I \leq \int_R t$.

This is a formal definition and you don't usually integrate things like this in practice, but it is nice to have this definition handy when dealing with tricky functions with many discontinuities.

An important result from your book about integration is the following theorem.

Theorem 2.1 (Apostol 11.5)**.** *If f is integrable in a region $Q = [a, b] \times [c, d]$ and $A(y) := \int_a^b f(x, y)\, dx$ exists for all y, then if $\int_c^d A(y)\, dy$ exists, it is equal to $\int\int_Q f$, that is*

$$\iint_Q f(x, y)\, dx\, dy = \int_c^d \int_a^b f(x, y)\, dx\, dy.$$

More results are in the chapter (11.12) about regions and when functions are integrable over those regions.

Let's go through a simple example to indicate what we expect on your homework solutions.

Example 3. Assuming that $f(x, y) = ye^x$ is integrable over $R = [0, 1] \times [0, 1]$, what is $\int\int_R f(x, y)\, dx\, dy$?

Solution. Since f is integrable, we want to apply Theorem 11.5. Since $A(y) := \int_0^1 ye^x\, dx = y(e - 1)$ we see that the integral $A(y)$ exists for all y. Since $A(y)$ ins also integrable on $[0,1]$, we have

$$\int_0^1 A(y)\,dy = (e-1)\int_0^1 y\,dy = \frac{1}{2}(e-1).$$

Therefore, by Theorem 11.5, we have

$$\iint_R f(x,y)\,dx\,dy = \int_0^1\int_0^1 ye^x\,dx\,dy = \frac{1}{2}(e-1).$$

□

Thus, in certain cases, such as when the functions are continuous over the domain of integration, Theorem 11.5 provides everything you need. However, for some tricky cases, we need to use the formal definition of integration to get an answer.

Example 4. If a function f defined on $R = [0,1]\times[0,1]$ is 1 at a finite number of points $x_1, x_2, \ldots, x_n$ and 0 elsewhere, then f is integrable and $\int_R f = 0$.

Proof. We note that f is bounded below by 0 and above by 1. We need to show that if s and t are step functions such that $s < f < t$, we must have $\int_R s \le 0 \le \int_R t$ and show that 0 is the only value that works in the inequality.

Now, suppose that we are given such step functions s and t. Given any rectangle in R, there must be some point $\mathbf{a}$ in the rectangle on which $f(\mathbf{a}) = 0$, so we must have $s \le 0$ in that rectangle. Since this holds for all rectangles in R, we must have $s \le 0$ on R and so $\int s \le 0$ by Theorem 11.3 (Comparison Theorem). Since $0 \le t$, we must have $0 \le \int_R t$ by Theorem 11.3 again. Thus, our desired inequality holds for 0. It remains to show that 0 is the unique value I such that $\int_R s \le I \le \int_R t$.

To do this, we argue as follows. Since $s \equiv 0 \le f$, if any other I satisfies the inequality on integrals, it must be positive. We will show that no positive number works. Let $\epsilon > 0$. For each $x_i = (x_{i_1}, x_{i_2})$, consider the rectangle

$$R \cap [x_{i_1} - \sqrt{\epsilon/8n}, x_{i_1} + \sqrt{\epsilon/8n}] \times [x_{i_2} - \sqrt{\epsilon/8n}, x_{i_2} + \sqrt{\epsilon/8n}] \subset R$$

and let t be the step function that is 1 on the rectangles above and 0 elsewhere. Let s be the step function that is zero everywhere. Therefore, we have

$$\int t \le n(2\sqrt{\epsilon/8n})^2 = \frac{\epsilon}{2}$$

This implies that $s \le f \le t$ but $\int t < \epsilon$. Therefore, we don't have the inequality for integrals for any $\epsilon > 0$. Hence, 0 is the only value such that $\int s \le 0 \le \int t$ for all step functions s and t, and so f is integrable with integral 0. □

2.1. **A motivating example.** Integration theory as defined above is quite powerful, and allows you to integrate functions that are quite pathological. Consider the following such case.

We want to find the integral of the following function, often called the *popcorn function* and is defined for $x \in [0,1]$ as follows:

$$f(x) = \begin{cases} \frac{1}{q}, & x = \frac{p}{q} \in \mathbf{Q} \\ 0, & x \notin \mathbf{Q}. \end{cases}$$

This function has some curious properties. For instance, it is *discontinuous* on $\mathbf{Q} \cap [0,1]$, because there is a sequence of irrational numbers approximating any rational number. However, it is also continuous on the irrationals. To see this, let

a be an irrational number and let $\epsilon > 0$. Then the set of all rational numbers with $1/q > \epsilon$ is a finite set S. Therefore, if we set δ to be the minimal distance between a and any of the points in the set S, then $|a - x| < \delta$ implies that $|f(x)| < \epsilon$.

It turns out that f is integrable and so its integral has a well-defined value. You might guess that it is zero, but how can we see this? To do this, it helps to first understand the concepts of measure zero and content zero. They allow us to give us an answer to the question: *When does the integral of a function exist?*

3. Measure Zero, Content Zero

Definition 3.1. A set $A \subset \mathbf{R}^n$ has **content zero** if for all $\epsilon > 0$ there exists a *finite* collection of rectangles whose union contains A and whose total area (in $\mathbf{R}^n$ sense, e.g. volume when $n = 3$) is less than ϵ.

Definition 3.2. A set $A \subset \mathbf{R}^n$ has **measure zero** if for all $\epsilon > 0$ there exists a (possibly infinite) collection of rectangles whose union contains A and whose total area is less than ϵ.

Obviously, a set being content zero implies that it is also measure zero. Let's start with the most basic example.

Example 5. Let $A = \{x_1, \ldots, x_k\}$ be a finite collection of points in $\mathbf{R}$. Then A has both measure and content zero, since given ϵ, we can take rectangles $[x_i - \epsilon/2k, x_i + \epsilon/2k]$ which contain A and whose area (i.e. length) is less than epsilon.

Example 6. An interval in $\mathbf{R}$ (or a rectangle in higher dimensions) has nonzero measure and content.

Example 7. The rational numbers $\mathbf{Q}$ are measure zero but the content is nonzero. For instance, as we saw in the homework assignment a couple of weeks ago, if we can to cover the rationals in $[0, 1]$ with a finite number of intervals, we must cover all of $[0, 1]$ and so we cannot have total area less than 1, so the content 0 criterion fails for $\epsilon < 1$ (at least).

The reason why the rationals are measure zero is because they are a countable set of points. For any countable set of points A, we can put it in bijection with the natural numbers and consider rectangles of radius $\frac{\epsilon}{2}\left(\frac{1}{2}\right)^i$ around every point $x_i \in A$. These rectangles cover all the points and the total area is

$$\frac{\epsilon}{2}\sum_{i=1}^{\infty}\left(\frac{1}{2}\right)^i = \frac{\epsilon}{2} \cdot 2 = \epsilon.$$

Since we can do this for any ϵ, we see that, for example, the set $\mathbf{Q} \cap [0, 1]$ has measure zero. (Indeed, $\mathbf{Q}$ itself is a measure zero set in $\mathbf{R}$).

Example 8. The Cantor set an example of measure zero set with uncountably many points. (In fact, it is even content zero!) Therefore, while countable sets have measure zero, being measure zero doesn't imply that your set is countable.

4. The Integral of the Popcorn Function

One of the central results in the theory of integration is following theorem.

Theorem 4.1 (Lebesgue)**.** *Suppose that f is a bounded function defined on some domain R. Then f is integrable if and only if the set of discontinuous points of f has measure zero.*

I'm not sure if you can use this result yet, since its proof uses methods outside the scope of this class. However, it's good result to know, just so you can determine whether a set is integrable or not before you try and work it out.

What you are definitely allowed to use is the following theorem from the book.

Theorem 4.2 (Apostol 11.7)**.** *Let f be a bounded function defined on R. If the set of discontinuous points of f is a set of content zero in R, then f is integrable over R.*

In particular, a function that has only finitely many discontinuities is integrable, as we saw in example 4.

Anyway, since the popcorn function is discontinuous at the rationals, a measure zero set, it is integrable, so we know that our proof won't just crash and burn when we try and construct the appropriate step functions.

So let's find the integral of the popcorn function.

Solution. We will show that $\int f = 0$. We begin with a couple of observations. Any step function s less than f must be zero everywhere, since every interval contains irrational numbers. Furthermore, any step function t greater than f must be positive everywhere and thus we must have $\int t > 0$. Thus, we want to show that given $\epsilon > 0$, we can find a step function t such that $f \leq t$ and $\int t < \epsilon$.

Now, there are a finite number of points $\{x_i\}_{i=1}^N$ such that $f(x_i) > \epsilon/2$. Let d be the minimum distance between any two points x_i and x_j, and define a quantity $D = \min(\frac{\epsilon}{2N}, \frac{d}{2})$. Define a step function t

$$t(x) = \begin{cases} 1, & x \in [x_i - D/2, x_i + D/2] \text{ for all } x_i \\ 0, & \text{otherwise.} \end{cases}$$

Then $f \leq t$ and we have $\int t \leq ND + \epsilon/2 \leq \epsilon$. Therefore, f is integrable and $\int f = 0$. □