MA1C, WEEK 4: LAGRANGE MULTIPLIERS

TA: PADRAIC BARTLETT

These notes, like all future notes, can be found on my website.

1. Last Week's HW

Average: was around 83%. There was an email sent out earlier about common errors; basically, people made computational errors at times, forgot to show that functions satisfied certain theorems before applying these theorems, and occasionally forgot the old-school definition of partial derivatives via limits.

2. Random Question

Question 2.1. *So: suppose that you've been teleported back to Rome, and (in keeping with all good Roman stereotypes) you have been promptly placed in a gladiatorial arena, which is in the shape of a very large equilateral triangle. You sit at the center of this triangle; at each vertex, there is a lion, which is free to run along the boundary of the triangle but cannot escape the boundary of the triangle due to a complicated system of chains.*

Suppose that both you and the lions move at 10m/s, can pivot and change direction instantly, are arbitrarily brilliant, and know no fear. Can you escape from the triangle?

What if it were a square? How about any regular polygon?

3. Lagrange Multipliers

So: we do a brief review of our methods for cataloging extrema. (We assume here that f is a C^2 function from an open set $U \subseteq \mathbb{R}^n$ to $\mathbb{R}$.)

- We say that x_0 is a critical point of f iff $Df(x_0)$ is positive-definite.
- If $Hf(x_0)$ is positive-definite and x_0 is a critical point, x_0 is a local minimum.
- If $Hf(x_0)$ is negative-definite and x_0 is a critical point, x_0 is a local maximum.
- If the matrix $Hf(x_0)$ has both positive and negative eigenvalues, and x_0 is a critical point, then x_0 is a saddle point.

So: sometimes we want to find the maximum points of a function given some set constraints; given a function $f(x, y)$, we might want to only find the maximum values of f along the line $y+x = 1$, say. To do this, we use the method of **Lagrange Multipliers**, which we outline below:

Proposition 3.1. *Suppose that*

- *$f, g : U \subseteq \mathbb{R}^n \to \mathbb{R}$ are C^1 functions,*
- *S is the level set $g(x) = c$, for some constant c, and*
- *x_0 is a point in S.*

Then, we claim that whenever $f|_S$, the function f restricted to the set S, has a critical point at x_0, there is some $\lambda \in \mathbb{R}$ such that

$$\nabla f(x_0) = \lambda \nabla f(x_0).$$

Proof. If we have a maximum at x_0, then along any path γ in S, we have that $f \circ \gamma : \mathbb{R} \to \mathbb{R}$ has a maximum; so its derivative is 0 at x_0, by our knowledge of how single-variable calculus works. Then, by the chain rule, this means that $\nabla f(x_0) \cdot \gamma'(x_0)$ is 0, for every curve γ – i.e. that $\nabla f(x_0)$ is a vector orthogonal to the entire tangent space of S at x_0.

Similarly, for any path γ in S, $g \circ \gamma$ is the constant c (because S is a level set of g); so, as the derivative of a constant is 0, we have again by the chain rule that $\nabla g(x_0) \cdot \gamma'(x_0)$ is 0, for every path γ, and thus that $\nabla g(x_0)$ is orthogonal to the entire tangent space of S at x_0.

But the tangent space to S at x_0 is a space of dimension $n-1$, because S is a level curve of g; so the dimension of the space orthogonal to it is 1. So $\nabla g(x_0)$ and $\nabla f(x_0)$ both lie in the same 1-dimensional space – i.e. they both lie on a line – i.e. they are scalar multiples of each other. $\square$

So: we work an example, to demonstrate how the theory is actually applied.

Example 3.2. Let $f(x, y) = x^3 - 3xy^2$, $g(x, y) = x^2 + y^2$, and $c = 1$. Find all of the extremal points of f restricted to $g(x, y) = c$.

Proof. So: we have that

$$\begin{aligned} \nabla f(x,y) = (3x^2 - 3y^2, -6xy) &= \lambda \cdot \nabla g(x,y) = (2\lambda x, 2\lambda y) \\ &\Rightarrow -6xy = 2\lambda y \\ &\Rightarrow \text{either } y = 0 \text{ or } x = \frac{-1}{3}\lambda. \end{aligned}$$

Then, since $3x^2 - 3y^2 = 2\lambda x$, either

$$\begin{aligned} y = 0 \Rightarrow & 3x^2 = 2\lambda x \\ & \Rightarrow \text{either } x = 0 \text{ or } x = \frac{2}{3}\lambda, \\ \text{or } x = \frac{-1}{3}\lambda \Rightarrow & 3\left(-\frac{1}{3}\lambda\right)^2 - 3y^2 = -\frac{2}{3}\lambda^2 \\ \Rightarrow & \lambda^2 = 3y^2 \\ \Rightarrow & y = \pm\frac{\lambda}{\sqrt{3}}. \end{aligned}$$

So: the only possible points are either $(0, 0)$, $(\frac{2}{3}\lambda, 0)$, $(-\frac{1}{3}\lambda, \pm\frac{\lambda}{\sqrt{3}})$, for any value of λ such that these points lie on the unit circle $g(x, y) = x^2 + y^2 = 1$; i.e the six points

- $(\pm 1, 0)$,
- $(\pm\frac{1}{2}, \pm\frac{3\sqrt{3}}{6})$.

$\square$

(Explicitly, evaluating gives that the points $(1, 0), (-\frac{1}{2}, \pm\frac{3\sqrt{3}}{6})$ all evaluate to 1, while the points $(-1, 0), (\frac{1}{2}, \pm\frac{3\sqrt{3}}{6})$ all evaluate to -1; these turn out to be the minima and maxima of our function, though we haven't actually shown this yet –

we've merely established that these points are critical points, not necessarily that they are the maxima or minima of this function.)

It bears noting that our method of Lagrange multipliers can be extended to finding the extremal points of $U \cup \partial U$, where U is an open set and ∂U is its boundary, which is furthermore given by some equation $g(x) = c$. In this situation, we simply classify all of the maxima and minima by splitting it up into two cases:

(1) First, find the extremal points of U using our normal methods for finding extrema – i.e. look for all points $x_0 \in U$ where $Df(x_0) = 0$.
(2) Second, use the method of Lagrange multipliers above to find all of the extremal points of $\partial U =$ the level curve given by $g(x) = c$, for some c.
(3) Combine your results to get the total set of extremal points for $U \cup \partial U$.

4. Arc Length

So: we switch gears here.

Definition 4.1. For $c(t) = (c_1(t) \dots c_n(t))$ a path $\mathbb{R} \to \mathbb{R}^n$, we say that the **arc length** of the path c from the points t_0 to t_1 is

$$\int_{t_0}^{t_1} \sqrt{(c_1')^2(t) + \dots + (c_n')^2(t)}dt.$$

We work one quick example here, to show that this is actually something that "makes sense:"

Example 4.2. Let $\gamma(t) = (\sin(t), \cos(t))$. Find the arc length of γ from 0 to 2π.

Proof. So:

$$\int_0^{2\pi} \sqrt{\cos^2(t) + (-\sin)^2(t)}dt = \int_0^{2\pi} 1dt = 2\pi.$$

It bears noting that this makes sense; the graph of $(\sin(t), \cos(t))$ is a unit circle, and our normal formula for calculating the perimeter of a circle is $2\pi \cdot r$. □

5. Vector Fields and Flow Lines

Here are a few definitions:

Definition 5.1. A **vector field** on a set U is a map $F : U \subset \mathbb{R}^n \to \mathbb{R}^n$.

Definition 5.2. A **flow line** for a given vector field F on U is a path $\gamma : \mathbb{R} \to U$ such that at any point in $\mathbb{R}$,

$$\gamma'(t) = F(\gamma(t)).$$

The idea is that flow lines are paths that are in some sense determined by the vector field – at any point, the immediate direction that the path "wants" to go is given by the vector field. Given a vector field, you can quickly draw several flow lines by simply starting at a given point and moving your pencil in the direction of the vector field's arrows at that point; this is why we call them "flow lines," because they are the paths along which things "flow" in this vector field.

We work one example to demonstrate the mathematics behind this.

Example 5.3. Find a flow line for $F(x, y) = (-xy, y^2)$. Show that this is a flow line.

Proof. So: we're looking for a path $\gamma(t) = (\gamma_1(t), \gamma_2(t))$= such that

$$(\gamma_1'(t), \gamma_2'(t)) = F(\gamma(t)) = (-\gamma_1(t)\gamma_2(t), \gamma_2^2(t)).$$

So: because $\gamma_2'(t) = \gamma_2^2(t)$, one immediate candidate that comes to mind is $\gamma_2(t) = -\frac{1}{t}$, as this satisfies the equation in question. Then, this means that we are looking for a function $\gamma_1(t)$ that satisfies $\gamma_1'(t) = -\gamma_1(t)\gamma_2(t)$; setting $\gamma_1(t) = t$ satisfies this equation, as can be easily checked. Then, we have that

$$(\gamma_1'(t), \gamma_2'(t)) = (\frac{d}{dt}(t), \frac{d}{dt}(-\frac{1}{t})) = (1, \frac{1}{t^2}) = (-\gamma_1(t)\gamma_2(t), \gamma_2^2(t)) = F(\gamma(t));$$

so this is a flow line for F. The reader is invited to draw the vector field for F and the graph of γ, and check that this is indeed a path that "follows" the vectors given by F. $\square$