MATH 204: Homework 4

Solutions

Problems are from Rudin 3rd edition.

Problem 1. Chapter 9 (p. 239): 21

Problem 2. Let $E \subset \mathbb{R}^n$ open and $f : E \to \mathbb{R}$ be C^1. In class we saw that $f'(x) \in L(\mathbb{R}^n, \mathbb{R})$ can be represented as $f'(x)h = \nabla f(x) \cdot h$ for a vector $\nabla f(x) \in \mathbb{R}^n$ which we called the gradient of f. This makes ∇f a mapping from E to $\mathbb{R}^n$. Suppose that ∇f is itself differentiable on E and call it's derivative $D^2 f : E \to L(\mathbb{R}^n)$.

(a) Suppose that f has a local maximum at a point $x \in E$. Show that $D^2 f(x) \leq 0$ in the sense that,

$$\xi \cdot D^2 f(x)\xi \leq 0 \quad \text{for every} \quad \xi \in \mathbb{R}^n.$$

A linear operator with such a property is called *non-positive definite* (or *negative definite* if there is strict inequality).

(b) You saw on the midterm that if f has an interior local maximum at a point $x \in E$ then $\nabla f(x) = 0$. Now let us suppose that the domain $E = \{x \in \mathbb{R}^n : g(x) < 0\}$ for a C^1 function g which satisfies $\nabla g \neq 0$ on ∂E. Here we will need that both f and g are actually C^1 on an open set containing the closure of E. Suppose that f attains its maximum over the set $\overline{E}$ at a point $x \in \partial E$. Show that,

$$\nabla f(x) \cdot \nu(x) \geq 0,$$

where $\nu(x) = \frac{\nabla g(x)}{|\nabla g(x)|}$ is the unit normal vector to the domain E at the point x.

(c) In the same setting as part (b) assume that ∇f is also differentiable at the maximum point $x \in \partial E$. Let T_x be the orthogonal complement of the vector $\nu(x)$ in $\mathbb{R}^n$, i.e. T_x is the subspace of directions tangential to ∂E at x. Show that,

$$\xi \cdot (D^2 f(x) - \lambda D^2 g(x))\xi \leq 0 \quad \text{for every} \quad \xi \in T_x,$$

where $\lambda \in \mathbb{R}$ is the same constant as appears in part (d).

(d) Now lets consider the values of f restricted to $\partial E = \{x \in \mathbb{R}^n : g(x) = 0\}$. Suppose that f attains $\max_{\partial E} f$ at a point $x \in \partial E$. Finding the value/location of such a maximum is referred to as a *constrained optimization problem*. Show that at such a point,

$$\nabla f(x) = \lambda \nabla g(x) \quad \text{for some} \quad \lambda \in \mathbb{R}.$$

The parameter λ is often referred to as a *Lagrange Multiplier*.

part (a). Let a path $\gamma(t) = x + \xi t$ and consider $k(t) = f(\gamma(t))$ which has a local maximum at $t = 0$. First one should verify that k is twice differentiable at x. Next we claim that $k''(t) \leq 0$, one way is to look directly at the difference quotients,

$$k''(0) = \lim_{h \to 0} \frac{k(h) + k(-h) - 2k(0)}{h^2} \leq 0.$$

Now we evaluate $k'(t)$ and $k''(t)$ using the chain-rule,

$$k'(t) = \nabla f(\gamma(t)) \cdot \gamma'(t) \quad \text{and} \quad k''(t) = \gamma'(t) \cdot D^2 f(\gamma(t))\gamma'(t) + \gamma''(t) \cdot \nabla f(\gamma(t)).$$

Plugging in $\gamma'(t) = \xi$ and $\gamma''(t) = 0$ we get,

$$0 \geq k''(0) = \xi \cdot D^2 f(x)\xi.$$

part (b). Let $\gamma(t) = x + t\nu(x)$ then,

$$g(x + t\nu(x)) = g(x) + t\nabla g(x) \cdot \nu(x) + r(t) \quad \text{with} \quad \lim_{|t| \to 0} \frac{r(t)}{|t|} = 0.$$

Noting that $\nabla g(x) \cdot \nu(x) = |\nabla g(x)| > 0$ we see that $g(\gamma(t)) > 0$ for $t > 0$ small enough and $g(\gamma(t)) < 0$ for $t < 0$ small enough. In particular $\gamma(t) \in E$ for $t < 0$ small enough and so $f(\gamma(t)) \leq f(x)$ for $t < 0$ small enough and,

$$\nabla f \cdot \nu(x) = \lim_{t \to 0} \frac{f(x + t\nu(x)) - f(x)}{t} = \lim_{t \to 0-} \frac{f(x + t\nu(x)) - f(x)}{t} \geq 0.$$

You should attempt to understand the picture geometrically before trying to write down the details in terms of calculus. Using part (c) we could see that actually $\nabla f(x) = |\nabla f(x)| \nu(x)$.

part (d). The first thing I want to do is construct a C^1 curve $\gamma(t)$ mapping some interval $I \ni 0$ in $\mathbb{R}$ into ∂E passing through x at $t = 0$ and with $\gamma'(0) = \xi$. This way we can argue similarly to part (a) and to the test problem.

Let's change coordinates so that $\nu(x) = e_n$, and we write $y \in \mathbb{R}^n$ as (y', y_n) with $y' \in \mathbb{R}^{n-1}$. This makes $T_x = \mathrm{span}(e_1, \ldots, e_{n-1})$. Then we can apply to implicit function theorem to find a C^1 function $\psi : U \to \mathbb{R}$ with U a neighborhood of x' in $\mathbb{R}^{n-1}$,

$$g(y', \psi(y')) = 0 \ \text{ with } \ \psi(x') = x \ \text{ and } \ \nabla \psi(x') = -D_n g(x)^{-1} \nabla_{y'} g(x) = 0,$$

since $\nabla g(x)$ is parallel to e_n. Now for $\xi \in T_x$, i.e. $\xi = (\xi_1, \ldots, \xi_{n-1}, 0)$, take for $\gamma(t)$ the path,

$$\gamma(t) = (x' + \xi' t, \psi(x' + \xi' t)) \ \text{ with } \gamma'(0) = (\xi', \nabla \psi(x') \cdot \xi') = (\xi', 0) = \xi.$$

The path γ is C^1 in a neighborhood of $t = 0$ since ψ given from the implicit function theorem is C^1 in a neighborhood of x'.

Now that we have a curve $\gamma \in \partial E$ passing through the point x with velocity $\xi \in T_x$ we can compute, using that $f(\gamma(t))$ has a local maximum at $t = 0$,

$$0 = \frac{d}{dt} f(\gamma(t)) \Big|_{t=0} = \nabla f(x) \cdot \xi.$$

In other words $\nabla f(x)$ is orthogonal to every vector ξ of T_x, but since T_x was defined as the orthogonal complement of span of $\nabla g(x)$ this means that $\nabla f(x)$ is parallel to $\nabla g(x)$ or,

$$\nabla f(x) = \lambda \nabla g(x) \ \text{ for some } \ \lambda \in \mathbb{R}.$$

part (c). So part (c) is not stated correctly, and it is really more natural to put it after part (d), which I have done here. First we claim that for any unit vector $\xi \in T_x$ there exists a C^2 path $\gamma : I \to \partial E$, where $I \ni 0$ is an open sub-interval of $\mathbb{R}$, with $\gamma(0) = x$, $\gamma'(0) = \xi$. This requires g to be C^2, and it is not so obvious, in particular it is convenient to use the implicit function theorem as in part (d) above. I will skip this part and take for granted the existence of such a curve $\gamma(t)$.

Then we can also compute, by taking the derivative on both sides of $g(\gamma(t)) = 0$ twice,

$$\nabla g(x) \cdot \xi = 0 \ \text{ and } \ \xi \cdot D^2 g(x) \xi + \gamma''(0) \cdot \nabla g(x) = 0$$

Using the previous equation and part (d) we get that,

$$\gamma''(0) \cdot \nabla f(x) = \lambda \gamma''(0) \cdot \nabla g(x) = -\lambda \left[\xi \cdot D^2 g(x) \xi \right],$$

for the same λ as in part (d). Now we compute the second derivative of $f(\gamma(t))$, since $\gamma(t) \in \partial E$ for all $t \in I$ this function has a local max at $t = 0$ and so,

$$0 \geq \frac{d^2}{dt^2} f(\gamma(t)) \Big|_{t=0} = \xi \cdot D^2 f(x) \xi + \gamma''(0) \cdot \nabla f(x) = \xi \cdot (D^2 f(x) - \lambda D^2 g(x)) \xi.$$