Math 25b Homework 9 Solutions Part 1

Ivan Corwin and Alison Miller.

1 Alison's problems

There were a bunch of problems this that everyone did well on, so these aren't complete solutions: if you have questions about the parts not explained here, please come to office hours or section and ask me!

(1) *Second Derivative Test*

Solution. The interesting and hard part of this problem is rigorously proving the criterion for the local maximum.

We first do the case where $D > 0$ and $T > 0$. From previous parts of the problem, we know that for some y on the line between x and $x + h$,

$$\begin{aligned} f(x+h) - f(x) &= \sum_{i=1}^{n} D_i f(x) h^i + \frac{1}{2!} \sum_{i=1}^{n} \sum_{j=1}^{n} D_{ij} f(y) h^i h^j + \sum_{k=1}^{n} D_{ijk} f(y) h^i h^j h^k \\ &= h^T M h + \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} D_{ijk} f(y) h^i h^j h^k. \end{aligned}$$

We know that the third derivatives $D_{ijk} f$ are continuous in a neighborhood of x, and there are only finitely many of them, so we can bound them all by finding a neighborhood U of x so that there's a constant M such that for any $y \in U$ and any indices i, j, k, $|D_{ijk} f(y)| < M$. Note that for each i, $|h_i| < \|h\|$, so any $|h_{ijk}| \leq \|h\|$. These facts, and the triangle inequality, give us the bound

$$\left| \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} D_{ijk} f(y) h^i h^j h^k \right| \leq \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} |D_{ijk} f(y)| |h^i h^j h^k| < n^3 M \|h\|^3.$$

This will be useful, because it means that the above term is relatively small in size, so that the $h^T M h$ term will generally be more important in determining the sign of $f(x+h) - f(x)$: unfortunately, this need not always be the case – it's possible that the stars might align so that the $h^T M h$ term will be even smaller in absolute value (in the case when $D < 0$, we can even find h with $h^T M h = 0$). A number of you lost some points for this.

In the case where $D > 0$, $T > 0$, we need to show two things: first, that the $h^T M h$ is positive, and second, that it will be large enough to override the cubic term provided $\|h\|$ is sufficiently small. The first part is just linear algebra: we know from a previous part of the problem that M is diagonalizable with two real eigenvalues, λ_1 and λ_2 – without loss of generality, $\lambda_1 \leq \lambda_2$: let v_1, v_2 be the corresponding basis of eigenvectors, which can be taken to be orthonormal (with respect to the standard inner product $\langle v, w \rangle v^T w$) by the Spectral Theorem. Then $D = \lambda_1 \lambda_2 > 0$, so λ_1 and λ_2 have the same sign: also $T = \lambda_1 + \lambda_2 > 0$, so they must both be positive. Now, write our vector h in terms of our eigenbasis: $h = c_1v_1 + c_2v_2$. Then $Mh = c_1\lambda_1v_1 + c_2\lambda_2v_2$, and $h^TMh = (c_1\lambda v_1 + c_2\lambda v_2)^T(c_1v_1 + c_2v_2)$: by orthonormality, this is $\lambda_1c_1^2 + \lambda_2c_2^2$. Orthonormality also means that $\|h^2\| = c_1^2 + c_2^2$. As $\lambda_1 \le \lambda_2$, we can conlude that $h^TMh \ge \lambda_1\|h\|^2$ (apparently, this is also the Rayleigh-Ritz theorem – look it up). This not only shows that h^TMh is positive, but that it is large enough: for $\|h\|$ sufficiently small, $\lambda_1\|h\|^2 > n^3M\|h\|^3$, and

$$f(x+h) - f(x) = h^TMh + \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} D_{ijk}f(y)h^ih^jh^k \ge \lambda_1\|h\|^2 - n^3M\|h\|^3 > 0,$$

showing that f does in fact have a local maximum at x.

The exact same argument with signs switched shows that for $D > 0$, $T < 0$, f has a local minimum at x.

In the final case, we have to show that f has neither a max nor a min: that is, we need to find arbitrarily small h for which $f(x+h)$ takes on both positive and negative values. We can't do the sort of bound on h^TMh in general that we had in the previous case: instead, we just look at eigenvectors. We know that $D = \lambda_1\lambda_2 < 0$, so the two eigenvalues have different signs: assume that $\lambda_1 > 0$, $\lambda_2 < 0$. As before, let v_1, v_2 be the corresponding eigenvectors. Then conside $h = c_1v_1$: h is also an eigenvector, so $Mh = \lambda_1h$ and $h^TMh = \lambda_1h^Th = \lambda_1\|h\|^2$. We then apply the argument from above: choose c_1 so that $\|h\|$ is small enough that $\lambda_1\|h\|^2 > n^3M\|h\|^3$. As before, we conclude that $f(x+h) - f(x) > 0$: we can make h arbitrarily small, so f cannot have a local maximum at x. But we can also do the same argument for $h = c_2v_2$, so f can't have a local minimum either. □

Pretty much everyone got problems 2 and 3, and they're mostly computation, so I won't write them up here. If you still have questions about how to do these sorts of computation, ask me (or Ivan, or Elizabeth).

(4) *Measure Zero*

Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function. Show that the graph of f, which is a subset of $\mathbb{R}^2$, has measure zero.

Solution. The first thing to notice is that it suffices to show that, for all $n \in \mathbb{N}$ the graph of f restricted to $[-n, n]$ (that is $\{(x, f(x)) \mid x \in [-n, n]\}$ has measure zero. This is because the graph of f is the countable union of the graphs of f on $[-n, n]$ for all n, and Spivak 3-4 says that a countable union of measure 0 sets is measure 0.

Now let's show that the graph of f restricted to $[-n, n]$, which we'll call $G_{[-n,n]}$, has measure 0. Choose any $\epsilon > 0$: we need to cover $G_{[-n,n]}$ with rectangles of total area $< \epsilon$. The important thing here is that $[-n, n]$ is a compact interval, so f is uniformly continuous on $[-n, n]$: choose $m \in \mathbb{N}$ so that for $x, y \in [-n, n]$, if $|x - y| < 1/m$, $|f(x) - f(y)| < \frac{\epsilon}{2n}$. Then divide the interval $[-n, n]$ up into $2nm$ subintervals of length $1/m$. On each subinterval, the difference between the maximum and minimum values attained by f can be at most $\frac{\epsilon}{2n}$, so the graph of f on each subinterval can be covered by a rectangle of area less than $\frac{\epsilon}{2nm}$. The $2nm$ such rectangles cover all of $G_{[-n,n]}$ with an area of less than ϵ, as desired.

(There is an alternate proof that uses the fact that f is Riemann integrable on $[-n, n]$ – geometrically, for any partition, the difference between the upper and lower sums is the area of a collection of rectangles covering the graph of f. So we can use Riemann integrability to make this difference arbitrarily small, which gives us a collection of rectangles covering f of arbitrarily small area. If you think about it, the argument we used above is pretty much the same argument used to show that f is Riemann integrable on a compact interval, so it comes down to the same thing.)

$\square$