Notes for 8 Mar (Wednesday)

1 Recap

1. Proved that derivatives satisfy the intermediate value property, i.e., derivatives have only discontinuities of the second kind.
2. Proved a general version of L'Hospital's rule and warned against its usage when the limit of the ratios of the derivatives does not exist.
3. Proved Taylor's theorem.

2 Vector-valued functions

The notions of derivatives can be defined for $\vec{f} : (a, b)\mathbb{R}^n$ in exactly the same manner as before. A vector-valued function is differentiable if and only if its components are. The usual rules go through except for the quotient and the product rules. The product also goes through if we consider the dot product of vectors. However, L' Hospital's rule and the MVT fail spectacularly for complex-valued functions. Indeed,

1. Define $f(x) - e^{ix} = \cos(x) + i\sin(x)$. Thus $f^{'}(x) = ie^{ix}$. (We haven't defined these yet, but let's pretend that we have and move on for the purposes of this example.) Note that $f(2\pi) = f(0) = 1$. But $|f^{'}(x)| = 1$.
2. On $(0, 1)$ define $f(x) = x$ and $g(x) = x + x^2e^{i/x^2}$. Note that $\lim_{x\to 0}\frac{f(x)}{g(x)} = 1$. However, L' Hospital would say that it is limit of $\frac{1}{1+(2x-\frac{2i}{x})e^{i/x^2})}$. Note that the denominator $g^{'}(x)$ satisfies $|g^{'}(x)| \leq |2x - \frac{2i}{x}| - 1 = \sqrt{4x^2 + \frac{4}{x^2}} - 1 \geq \frac{2}{x} - 1$. Therefore, the limit of the fraction goes to 0.

But something similar to MVT is true.

Theorem 1. *Suppose $\vec{f} : [a, b] \to \mathbb{R}^n$ is continuous and $\vec{f}$ is differentiable on (a, b). Then there is a $c \in (a, b)$ such that $|f(\vec{b}) - \vec{f}(a)| \leq (b - a)|\vec{f}^{'}(c)|$.*

Proof. Let $\phi(t) = (f(\vec{b}) - \vec{f}(a)).\vec{f}(t)$. By the MVT $\phi(b) - \phi(a) = \phi^{'}(c)(b - a)$ for some c. Thus $|f(\vec{b}) - f(\vec{a})|^2 = (f(\vec{b}) - \vec{f}(a)).\vec{f}^{'}(c)(b - a)$. The Cauchy-Schwartz inequality gives us the result. □

3 The Riemann-Stieltjes integral

A review of the Riemann integral : So the symbol $\int_a^b f(x)dx$ was defined in order to calculate the area under the graph of $y = f(x)$. So naively speaking, one divides $[a, b]$ into a large number of pieces $a, a + \frac{b-a}{n}, \ldots, b$, then chooses points x_i in each of those pieces, constructs rectangles with base the i^{th} piece and height $f(x_i)$, sums the areas $\sum_i f(x_i)\delta x_i$ and hopes that as the number of pieces became large, the approximation gets better. Indeed, let's make this a little more rigorous :

Definition : A partition P of an interval $[a, b]$ is a finite set of points $a = x_0 \leq x_1 \leq x_2 \ldots \leq x_n = b$. We write $\delta x_i = x_i - x_{i-1}$.

Suppose $f : [a, b] \to \mathbb{R}$ is a bounded function. Then $M_i = \sup_{[x_{i-1},x_i]} f(x)$ and $m_i = \inf_{[x_{i-1},x_i]} f(x)$ exist. Denote the upper Riemann-Darboux sum $U(P, f) = \sum_{i=1}^{n} M_i \delta x_i$ and the lower Riemann-Darboux sum $L(P, f) = \sum_{i=1}^{n} m_i \delta x_i$.

Note that $U(P, f)$ is bounded below and $L(P, f)$ is bounded above. Thus the upper and lower Riemann integrals $\int_a^{\bar{b}} f dx = \inf_P U(P, f)$ and $\int_{\bar{a}}^b f dx = \sup_P L(P, f)$ exist. If these two are equal, then we say that f is Riemann integrable on $[a, b]$ with Riemann integral $\int_a^b f dx$ equal to the common value of the upper and lower integrals.

We could proceed from this point, ask when functions are Riemann integrable, how one calculates Riemann integrals, and so on. However, with just a little bit of effort, one can handle a more general situation. So we might as well do that. The more general integral is called the Riemann-Stieltjes integral. It is written as $\int_a^b f(x)d\alpha(x)$ where f is a bounded function on $[a, b]$ and $\alpha(x)$ is a monotonically increasing function on $[a, b]$. For instance, if $\alpha(x) = x$ then we get back our good ol' Riemann integral.

Before proceeding to define the Riemann-Stieltjes integral rigorously, you can see that if α has continuous first derivatives, it is reasonable to expect $\int f d\alpha = \int f\alpha' dx$ therefore reducing a Riemann-Stieltjes integral to an ordinary Riemann integral. Therefore the utility of the RS integral should be apparent in cases where α is not well-behaved. Why should we care about such cases ? Why bother with the RS integral in the first place ?

The point is that if you ever deal with probability theory, there are two kinds of situations : Tossing coins (discrete random variables) and throwing darts (continuous random variables). In the former case, you sum over all possibilities whereas in the latter case you integrate over all possibilities. The Riemann-Stieltjes integral is an attempt to unify these two situations. Indeed, naively speaking, if α is a step function, then its derivative is a sum of Dirac-delta "functions". Thus the integral reduces to a discrete sum. Ultimately, the Lebesgue integral is the most general theory of integration. The

RS integral is a precursor to the Lebesgue integral.

Definition : We write $\Delta\alpha = \alpha(x_i) - \alpha(x_{i-1})$. Define the upper and lower sums as $U(P, f, \alpha) = \sum_{i=1}^{n} M_i \Delta\alpha_i$ and $L(P, f, \alpha) = \sum_{i=1}^{n} m_i \Delta\alpha_i$. As before we define the upper and lower RS integrals. We say that f is RS integrable w.r.t α if the upper and lower RS integrals are equal.

This definition is too hard to check. In the quest of a simpler criterion, we need to make another small, natural definition : A partition P^* is said to be a refinement of a partition P if $P \subset P^*$. Given two partitions P_1, P_2, P^* is said to be a common refinement if $P^* = P_1 \cup P_2$. Let's investigate what happens to the upper and lower sums when we go to a refinement.

Lemma 3.1. *If P^* is a refinement of P, then $L(P, f, \alpha) \leq L(P^*, f, \alpha)$ and $U(P^*, f, \alpha) \leq U(P, f, \alpha)$.*

Proof. Suppose P^* has exactly one more point x which is not in P and satisfies $x_{i-1} < x < x_i$. The general case can be done using induction.
Note that $U(P^*, f, \alpha) = \sum_{k=1}^{i-1} \tilde{M}_k \Delta\alpha_k + \tilde{M}_i(\alpha(x) - \alpha(x_{i-1})) + \tilde{M}_{i+1}(\alpha(x_i) - \alpha(x)) + \sum_{i+1}^{n} \tilde{M}_k \Delta\alpha_k$. The $\tilde{M}$s are all equal for P and P^* except for $i, i+1$. Indeed, $U(P^*) - U(P) = \tilde{M}_i(\alpha(x) - \alpha(x_{i-1})) + \tilde{M}_{i+1}(\alpha(x_i) - \alpha(x)) - M_i(\alpha(x_i) - \alpha(x_{i-1})) \leq 0$. Likewise for the lower sum. □

This may be used to prove that indeed the lower Riemann integral is less than or equal to the upper one :
Proof : Suppose P_1, P_2 are two partitions. Let P^* be their common refinement. Then $L(P_1) \leq L(P^*) \leq U(P^*) \leq U(P_2)$. Thus $\sup_{P_1} L(P_1) \leq U(P_2)$. Now take infimum to conclude the result.

Finally, we have a nice criterion for RS integrability.

Theorem 2. *f is RS integrable w.r.t α if and only if for every $\epsilon > 0$ there exists a partition P such that $U(P, f, \alpha) - L(P, f, \alpha) < \epsilon$.*

Proof. Suppose f is RS integrable w.r.t. α with integral L. Then by definition for every $\epsilon > 0$, there exist partitions P_1 and P_2 such that $L + \epsilon/2 > U(P_1) \geq L$ and $L - \epsilon/2 < L(P_2) \leq L$. Suppose P^* is a common refinement, then $L + \epsilon/2 > U(P^*) \geq L$ and $L - \epsilon/2 < L(P^*) \leq L$. Therefore $U(P^*) - L(P^*) < \epsilon$.
Conversely, suppose for every ϵ such a partition exists. Let $L = \sup L(P)$ and $U = \inf U(P)$. Choose a partition P such that $L - \epsilon/2 < L(P) \leq L$ and $U + \epsilon/2 > U(P) \geq U$. Now choose a partition $\tilde{P}$ such that $U(\tilde{P}) - L(\tilde{P}) < \epsilon/2$. Taking a common refinement P^* we see that $L - \epsilon/2 < L(P^*) \leq L$, $U + \epsilon/2 > U(P^*) \geq U$, and $U(P^*) - L(P^*) < \epsilon/2$. But this is a contradiction if $U \neq L$. □

Before we go ahead and prove that continuous functions are RS integrable, we prove another related lemma.

Lemma 3.2. *1. If for an $\epsilon > 0$, there is a partition so that $U(P) - L(P) < \epsilon$, then this is true for every refinement of the partition. (Already proved this.)*

2. If the above holds for P, and if s_i, t_i are arbitrary points in $[x_{i-1}, x_i]$ then $\sum_{i=1}^{n} |f(s_i) - f(t_i)|\Delta\alpha_i < \epsilon$.

3. If f is RS integrable and the above holds, then $|\sum f(t_i)\Delta\alpha_i - \int_a^b f d\alpha| < \epsilon$.

Proof. 1. Already done.

2. Of course $f(s_i), f(t_i) \in [m_i, M_i]$. Thus $\sum |f(s_i) - f(t_i)|\Delta\alpha_i < \sum(M_i - m_i)\Delta\alpha_i < \epsilon$.

3. Since $L(P) < \sum f(t_i)\Delta\alpha_i < U(P)$, and $L(P) \leq \int f d\alpha \leq U(P)$ we see the result. □

Now we prove that continuous functions are RS integrable :
Proof : The idea is to choose a sufficiently fine partition such that f does not vary much. Therefore $U(P) - L(P)$ will be small. Continuity means that f does not vary much near a point. But in principle, how "near" you have to be to a point p, might potentially depend on p. This is where the concept of uniform continuity comes to our rescue.
Since continuous functions on $[a, b]$ are uniformly continuous, given an $\epsilon > 0$ there exists an n such that for all p, q satisfying $|p - q| < \frac{b-a}{n}$ we have $|f(q) - f(p)| < \frac{\epsilon}{\alpha(b)-\alpha(a)}$. Now choose the partition $x_0 = a, x_1 = a + \frac{b-a}{n}, x_2 = x_1 + \frac{b-a}{n}, \ldots, x_n = b$. By the extreme value theorem there exist $u_i, v_i \in [x_{i-1}, x_i]$ such that $f(v_i) = M_i$ and $f(u_i) = m_i$. By uniform continuity $M_i - m_i < \frac{\epsilon}{\alpha(b)-\alpha(a)}$. Thus $U(P) - L(P) = \sum(M_i - m_i)\Delta\alpha_i < \frac{\epsilon}{\alpha(b)-\alpha(a)} \sum \Delta\alpha_i = \epsilon$.