Some Results from Calculus

Rich Schwartz
January 29, 2015

1 Single Variable Functions

These notes prove some results about functions on $\mathbb{R}^n$. We’ll start with functions of a single variable.

Lemma 1.1 Let $g : \mathbb{R} \rightarrow \mathbb{R}$ be a differentiable function with $g(0) = 0$. Then

$$|g(A)| \leq A \times \sup_{x \in [0,A]} |g'(x)|.$$ 

Proof: This is an immediate consequence of the Fundamental Theorem of Calculus. Here is a proof from scratch. We will establish the more general statement that the inequality

$$(*) \quad |g(a) - g(b)| \geq (1 + \epsilon)|a - b| \sup_{x \in [a,b]} |g'(x)|$$

cannot hold for any $\epsilon > 0$ on any sub-interval $[a, b] \subset [0, A]$. If (*) holds for some interval $I$, then by the triangle inequality it also holds for one of the two intervals obtained by cutting $I$ in half. But then (*) holds on a nested sequence $\{I_n\}$ of intervals, shrinking to a point $x_0$. This means that

$$\frac{|g(a_n) - g(b_n)|}{|a_n - b_n|} \geq (1 + \epsilon)|g'(x_0)|.$$ 

Here $I_n = [a_n, b_n]$. This contradicts the differentiability of $g$ at $x_0$ once $n$ is sufficiently large. ♠
2 Differentiability

A map \(f: \mathbb{R}^n \to \mathbb{R}^m\) is called differentiable at \(p\) if there is some linear map \(L: \mathbb{R}^n \to \mathbb{R}^n\) such that

\[
\lim_{{|h| \to 0}} \frac{|f(p + h) - f(p) - L(h)|}{|h|} = 0.
\]

Here \(h \in \mathbb{R}^n\) is a vector. In this case we write \(df(p) = L\). When \(f\) is differentiable at \(p\), the transformation \(L\) is the usual matrix of partial derivatives of \(f\), evaluated at \(p\).

**Theorem 2.1** Suppose that \(f: \mathbb{R}^n \to \mathbb{R}^m\) is a function whose partial derivatives exist and are continuous. Then \(f\) is differentiable at all points.

**Proof:** Considering the coordinate functions separately, it suffices to consider the case \(m = 1\). Translating the domain and range, it suffices to prove this at 0, under the assumption that \(f(0) = 0\). (Here \(0 \in \mathbb{R}^n\) is shorthand for \((0, ..., 0)\).) Subtracting off a linear functional, we can assume that \(\partial_j f(0) = 0\) for all \(j\).

Let \(v\) be any unit vector. (Here we are thinking that \(h = tv\) in Equation 1.) Let \(h = tv = (h_1, ..., h_n)\). Define

\[
h_0 = (0, ..., 0), \quad h_1 = (h_1, 0, ..., 0), \quad h_2 = (h_1, h_2, 0, ..., 0), \cdots.
\]

We have some constant \(\epsilon_t\) so that \(|f_{x_j}| < \epsilon_t\) for all \(j\). Moreover, \(\epsilon_t \to 0\) as \(t \to 0\). Lemma 1.1 gives us

\[
|f(h_j) - f(h_{j-1})| \leq t \times \epsilon_t.
\]

Summing over \(j\), we get

\[
|f(h)| \leq nt \epsilon_t.
\]

Hence

\[
\frac{|f(h)|}{|h|} \leq \frac{|f(tv)|}{t} \leq \epsilon_t.
\]

This ratio goes to 0 as \(t \to 0\). This shows that \(f\) is differentiable at 0 and \(Df(0)\) is the 0 transformation. ♠
An Example: Choose any smooth $2\pi$-periodic non-constant function $\psi(\theta)$ so that $\psi(k\pi/2) = 0$ for $k = 0, 1, 2, 3$. Now consider the function (in polar coordinates) $f(r, \theta) = r\psi(\theta)$. Also define $f = 0$ at the origin. This function is smooth except at the origin, and vanishes along the $x$-axis and $y$-axis. Hence $f_x$ and $f_y$ exist everywhere, and vanish at the origin. On the other hand, the restriction of $f$ to some line through the origin is a nonzero linear function, meaning that some directional derivative of $f$ at the origin is nonzero.

3 Another View of Differentiation

Define the dilation $D_r(p) = rp$. Consider the sequence of maps

$$f_r = D_r \circ f \circ D_{1/r}. \quad (2)$$

By construction, $f_r(v)$ converges to the directional (vector) derivative $D_v(f)$. Thus, $f$ is differentiable at 0 if and only if $\{f_r\}$ converges, uniformly on compact subsets, to a linear map $M$. This linear map is precisely the matrix of partials $Df(0)$.

This observation leads to the following result.

**Lemma 3.1** Suppose that $f : \mathbb{R}^n \to \mathbb{R}^n$ is a map with $f(0) = 0$. Suppose that $f$ is invertible on the unit ball $U$, and $V = f(U)$ is an open set, and $f$ is differentiable at 0. Then $f^{-1}$ is differentiable at 0 and $D(f^{-1})(0) = (Df(0))^{-1}$.

**Proof:** Replacing $f$ by $Af$ for some linear map $A$, it suffices to consider the case when $Df(0)$ is the identity. In this case, the sequence $\{f_n\}$ converges uniformly on compact subsets to the identity map. Consider the functions

$$f_{n}^{-1} = D_n \circ f^{-1} \circ D_n. \quad (3)$$

Since $V$ is an open set, the map $f^{-1}$ is defined on the disk of radius $\epsilon$ about 0. Hence $f_{n}^{-1}$ is defined on the disk of radius $n\epsilon$. In particular, these maps are eventually defined on any given compact subset $K$. Moreover, these maps converge to the identity. But then $f^{-1}$ is differentiable at 0 and $D(f^{-1})$ is the identity. ♠
4 A Technical Result

In this section we assemble another ingredient for the Inverse Function Theorem. We call \( f \) \textit{nice} if \( f(0) = 0 \) and

\[
\|df(p) - I\| < 10^{-100}. \tag{3}
\]

for all vectors \( v \) with \( \|v\| < 10^{100} \). Here \( I \) is the identity matrix and the norm can be taken to mean the maximum absolute value of a matrix entry of \( df(p) - I \). One property a nice function has is that

\[
\|(p - q) - (f(p) - f(q))\| < \frac{\|p - q\|}{2}. \tag{4}
\]

for all \( p, q \) having norm less than \( 10^{100} \). To prove this, we consider the segment \( \gamma \) connecting \( p \) to \( q \). Then \( f(\gamma) \) is a curve whose tangent vector is everywhere almost equal to \( p - q \).

\textbf{Lemma 4.1} Let \( f \) be a nice function. Let \( B_r \) denote the ball of radius \( r \) centered at the origin. Then \( B_1 \subset f(B_{10}) \).

\textbf{Proof:} If this is false, then there is some \( P \in B_1 - f(B_{10}) \). Note that \( f \) maps every point on the boundary of \( B_{10} \) at least, say, 8 units away from \( p \). For this reason, we can find some \( Q \in B_{10} \) such that

\[
P - f(Q) = \inf_{q \in B_{10}} P - f(q) > 0.
\]

But now consider the new point

\[
\overline{Q} = Q + (P - f(Q)).
\]

We compute

\[
P - f(\overline{Q}) = (\overline{Q} - Q) - (f(\overline{Q}) - f(Q)).
\]

From Equation 4 we get

\[
\|P - f(\overline{Q})\| < \|Q - \overline{Q}\|/2 = \|P - f(Q)\|/2.
\]

This is a contradiction, because \( f(\overline{Q}) \) is closer to \( P \) than is \( f(Q) \) and again \( \overline{Q} \in B_{10} \). \( \spadesuit \)
Say that \( f : \mathbb{R}^n \rightarrow \mathbb{R}^n \) is \( C^\infty \) if all partial derivatives of all orders exist for \( f \). Say that \( f \) is nonsingular at \( p \) if \( df(p) \) is invertible. Given open sets \( U, V \subset \mathbb{R}^n \) suppose \( f(U) = V \). Say that \( f \) is a diffeomorphism from \( U \) to \( V \) if \( f \) is a bijection and both \( f \) and \( f^{-1} \) are \( C^\infty \) and nonsingular at all points of their domains.

**Theorem 5.1 (Inverse Function Theorem)** Suppose that \( f \) is \( C^\infty \) and nonsingular at \( p \). Then there are open sets \( U \) and \( V \) with \( p \in U \) and \( f(p) \in V \) such that the map \( f : U \rightarrow V \) is a diffeomorphism.

Let \( ||q|| \) denote the norm of a point \( q \). We can replace \( f \) by a composition of the form \( AfB \), where \( A \) and \( B \) are invertible affine maps, to arrange that:

- \( p = 0 \) and \( f(p) = 0 \).
- For all \( q \) with \( ||q|| < 10^{100} \), we have \( ||Df_q - I|| < 10^{-100} \).

Here \( I \) is the identity matrix.

Now let \( U \) be the unit disk and let \( V = f(U) \). We will verify all the desired properties through a series of lemmas.

**Lemma 5.2** \( f \) is injective on \( U \).

**Proof:** for any \( q_1, q_2 \in U \), let \( \gamma \) be the line segment connecting \( q_1 \) to \( q_2 \). Consider the curve \( f(\gamma) \). By construction, the tangents \( f(\gamma) \) are nearly parallel equal to \( \gamma \). Hence \( \gamma \) cannot be a loop, and \( f(q_1) \neq f(q_2) \). \( \spadesuit \)

We also note that the argument above gives

\[
||f(q_2) - f(q_1)|| > \frac{||q_1 - q_2||}{2}.
\]  
\( 5 \)

**Lemma 5.3** \( V \) is open.

**Proof:** choose some \( v_0 \in V \) and let \( u_0 \in U \) be such that \( f(u_0) = v_0 \). Composing \( f \) by translations and dilations, we can switch to the case when

- \( u_0 = v_0 = 0 \).
• $f$ is nice.
• $B_{10} \subset U$.
• $B_1 \subset V$.

Then we can apply Lemma 4.1. ♠

Now we know that $V$ open. Consider $f^{-1} : V \to U$. Equation 5 tells us immediately that $f^{-1}$ is continuous. Lemma 3.1, together with symmetry, now tells us that $f^{-1}$ is differentiable and $D(f^{-1}) = (Df)^{-1}$ at each point. Now we have the magic equation

$$Df^{-1}(q) = df \circ f^{-1}(q).$$

(6)

If we know that $f^{-1}$ is $k$ times differentiable, then by the chain rule $Df^{-1}$ is $k$ times differentiable. But then $f^{-1}$ is $k + 1$ times differentiable. By induction $f^{-1}$ is $C^\infty$. 