Lecture 14

14.1 Compact subsets of $\mathbb{R}^n$

We start with the definition of Compactness which refers to a property of sets that captures the notion of being finite or having no "holes".

Definition 14.1.1 ▶ Compact Subset

A subset $K \subseteq \mathbb{R}^n$ is said to be compact if every sequence $\{x_n\} \subseteq K$ has a subsequence $\{x_{n_k}\}$ that is convergent to some $x \in K$.

This is known as the Bolzano-Weierstrass Property.

Observe that a compact subset of $\mathbb{R}^n$ is always **closed**. To see this, note that every sequence $\{x_n\} \subseteq K$, where K is a compact subset of $\mathbb{R}^n$ that converges to some $x \in \mathbb{R}^n$ has a convergent subsequence $\{x_{n_k}\}$ that converges to the same x. Since K is compact, we can say that $x \in K$. So, the convergent sequence $\{x_n\}$ converges to a point in K. Hence, K is closed.

More is true. A compact subset of $\mathbb{R}^n$ is **bounded** too. Assume that a compact subset $K \subseteq \mathbb{R}^n$ is not bounded. Note that, a subset of $\mathbb{R}^n$ is bounded iff it is contained inside an open ball. Since K is unbounded, we can get a sequence $\{x_m\} \subseteq K$ with $\|x_m\| > m$, which doesn't have a convergent subsequence. This shows that K is not compact that contradicts our assumption.

Therefore, a compact subset of $\mathbb{R}^n$ is closed and bounded. What about the converse?

Theorem 14.1.1

A closed and bounded box in $\mathbb{R}^n$ is compact.

Proof. We take a closed and bounded box $K := \prod_{i=1}^{n}[a_i, b_i] \subseteq \mathbb{R}^n$. Fix $i \in [n]$. Consider a sequence $\{x_m\} \subseteq K$. We take its projection on the i^{th} coordinate, i.e., $\{\pi_i(x_m)\} \subseteq [a_i, b_i]$. Consider $i = 1$, by Bolzano-Weierstrass Theorem, it has a convergent subsequence $\{\pi_1(x_{m_t})\} \subseteq [a_1, b_1]$ converging to $\alpha_1 \in [a_1, b_1]$. Now we take $i = 2$ and repeat the process to get a convergent subsequence $\left\{\pi_2(x_{m_{t_l}})\right\} \subseteq [a_2, b_2]$ converging to $\alpha_2 \in [a_2, b_2]$. Continuing this way, we get a convergent subsequence of $\{x_m\}$ converging to $\alpha = (\alpha_1, \ldots, \alpha_n) \in K$. Hence, K is compact. □

Theorem 14.1.2 (Heine-Borel Theorem)

A subset $K \subseteq \mathbb{R}^n$ is compact iff it is closed and bounded.

Proof. $\implies$ Done!

$\Longleftarrow$ Since, K is bounded, it is contained in a closed box, i.e., there exists $r > 0$ such that $K \subseteq [-r, r]^n$. So, Theorem 14.1.1 implies that all sequences in K has a convergent subsequence, which must converge in K because K is closed. Hence, K is compact.

□

Theorem 14.1.3

Let $f : \mathcal{O}_n \to \mathbb{R}^m$ be a continuous map. Then f sends compact sets to compact sets.

In other words, continuous image of a compact set is compact.

Proof. Let $K \in \mathcal{O}_n$ be compact. Take a sequence $\{x_k\} \subseteq K$ with a convergent subsequence $\{x_{k_t}\} \subseteq K$ converging to $x \in K$. Then, $\{f(x_k)\}$ is sequence in $f(K)$ with convergent subsequence $\{f(x_{m_t})\}$ converging to $f(x)$. The last statement about convergence follows from the continuity of f. This shows that $f(K)$ is compact. □

Theorem 14.1.4 (Extreme Value Theorem)

Let $K \subseteq \mathbb{R}^n$ be compact and $f : K \to \mathbb{R}$ a continuous map. Then $\exists\ a, b \in K$ such that $f(a) \le f(x) \le f(b)$ for all $x \in K$.

Proof. By Theorem 14.1.3, $f(K)$ is compact. So f is bounded which implies $\sup_K f, \inf_K f$ exist! Since $f(K)$ is closed, they must exist inside $f(K)$. □

14.2 Inverse Function Theorem

We are now going to study Inverse Function Theorem which relates the differentiability of a function to the differentiability of its inverse, enabling the study of local behavior and solving equations in higher dimensions. But before that we prove a lemma that is essential to prove the theorem.

Lemma 14.2.1

Let $\mathcal{O}_n \subseteq \mathbb{R}^n$ be open and convex. Suppose $f : \mathcal{O}_n \to \mathbb{R}^n$ be a C^1 function. If $\exists\ M > 0$ such that

$$\sup_{x \in \mathcal{O}_n} \left| \frac{\partial f_i}{\partial x_j}(x) \right| \le M \text{ for all } i, j$$

Then $\|f(x) - f(y)\| \le n^2 M \|x - y\|$ for every $x, y \in \mathcal{O}_n$.

Proof. Pick $x, y \in \mathcal{O}_n$ and $i \in [n]$. Then using Mean Value Theorem, we can get $c_i \in L_{x,y}$ such that

$$\begin{aligned}
f_i(x) - f_i(y) &= \boldsymbol{\nabla} f_i(c_i) \cdot (x - y) \\
\implies |f_i(x) - f_i(y)| &= \left| \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(c_i) \cdot (x_i - y_i) \right| \\
&\le \sum_{j=1}^{n} \left| \frac{\partial f_i}{\partial x_j}(c_i) \right| |x_i - y_i| && \text{(Triangle inequality)} \\
&\le M \sum_{j=1}^{n} |x_i - y_i| \le nM \|x - y\|
\end{aligned}$$

The last inequality follows from the inequality $|x_i - y_i| \le \|x - y\|$ which holds for all i.

Using the above,

$$\|f(x)-f(y)\| = \sqrt{\sum_{i=1}^{n} |f_i(x)-f_i(y)|^2} \leq \sqrt{\sum_{i=1}^{n} n^2M^2\|x-y\|^2} \leq n^2M\|x-y\|$$

we obtain the result. □

Theorem 14.2.1 (Inverse Function Theorem)

Let $f : \mathcal{O}_n \to \mathbb{R}^n$ be a C^1 function and $a \in \mathcal{O}_n$. Suppose $Df(a)$ is invertible. Then, there exist open sets V and W containing a and $f(a)$ respectively, such that $f : V \to W$ is invertible. Moreover, the local inverse $f^{-1} \equiv \left(f|_V\right)^{-1} : W \to V$ is differentiable and for all $y \in W$,

$$Df^{-1}(y) = \left((Df)\left(f^{-1}(y)\right)\right)^{-1}$$

i.e., locally, the derivative of the inverse is the matrix inverse of the derivative.

Proof.
We call $L = Df(a)$ which is given to be invertible and take $g(x) := L^{-1}f(x)$. Then,

$$\begin{aligned} Dg(a) &= L^{-1}(f(a)) \cdot (Df)(a) \\ &= \left[L^{-1}\right] \cdot Df(a) = I \end{aligned}$$

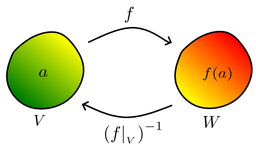

Figure 14.1

As this transformation can be made, without loss of generality, we may assume that $Df(a) = I_n$ which would imply that there exists a closed box U containing a such that for all $x \in U \setminus \{a\}, f(a) \neq f(x)$. To see this, let $f(a) = f(a+h)$ with arbitrarily small $\|h\|$. But then,

$$\frac{1}{\|h\|}\left(f(a+h) - f(a) - Ih\right) = \frac{h}{\|h\|} \neq 0$$

which contradicts the definition of derivative. Note that $\det J_f(a) \neq 0$. So, by continuity $\det J_f(x) \neq 0$ for all $x \in U$ (we may shrink U if necessary). Hence, $Df(x)$ is invertible for all $x \in U$. Again by continuity, for all $x \in U$ (we may shrink U if necessary),

$$\left| \frac{\partial f}{\partial x_j}(x) - \underbrace{\frac{\partial f}{\partial x_j}(a)}_{\delta_{ij}} \right| \leq \frac{1}{2n^2}$$

Now we claim the following,

Claim

For all $x, y \in U$,

$$\|f(x) - f(y)\| \geq \frac{1}{2}\|x-y\|$$

Proof. We take $g(x) = f(x) - x$ for all $x \in U$. Taking derivative, we get

$$\begin{aligned} Dg(x) &= Df(x) - I = Df(x) - Df(a) \\ \implies \frac{\partial g_i}{\partial x_j}(x) &= \frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(a) \\ \implies \left|\frac{\partial g_i}{\partial x_j}(x)\right| &\leq \frac{1}{2n^2} \ \forall\, x \in U \end{aligned}$$

Then by Lemma 14.2.1, for all $x, y \in U$,

$$\begin{aligned} &\|g(x) - g(y)\| \le n^2 \cdot \frac{1}{2n^2}\|x - y\| \\ \implies &\|(f(x) - f(y)) - (x - y)\| \le \frac{1}{2}\|x - y\| \\ \implies &\boxed{\|f(x) - f(y)\| \ge \frac{1}{2}\|x - y\|} \end{aligned} \tag{14.1}$$

Where (14.1) follows from the Triangle inequality. So, we get the claim! It shows that f is injective. ☐

Next we look at the compact set $\partial U \subseteq U$. Since $a \notin \partial U$, we can say $f(x) \neq f(a)$ for all $x \in \partial U$. So, by continuity of f and compactness of ∂U, we can find a $d \in \mathbb{R}_{\ge 0}$ such that

$$\|f(x) - f(a)\| \ge d \quad \forall\, x \in \partial U$$

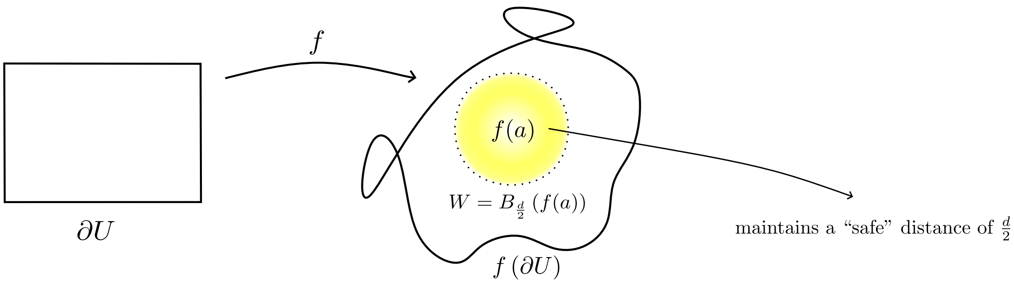

Figure 14.2

Now, we take $W = B_{\frac{d}{2}}(f(a))$. Then, for every $y \in W$ and $x \in \partial U$,

$$\underbrace{\|y - f(a)\|}_{\text{atmost } \frac{d}{2}} < \underbrace{\|y - f(x)\|}_{\text{atleast } d} \tag{14.2}$$

Claim

For a fixed $y \in W, \exists$ a unique $x_0 \in U^\circ$ such that $f(x_0) = y$

Proof. We define a continuous function $g : U \to \mathbb{R}$ with

$$g(x) = \|y - f(x)\|^2 = \sum_{i=1}^{n} (y_i - f_i(x))^2$$

Since, $\inf_U g$ cannot occur at the boundary ∂U, but must occur in U, there exists $x_0 \in U^\circ$ such that

$$\boldsymbol{\nabla} g(x_0) = 0 \text{ i.e., } \frac{\partial g}{\partial x_j}(x_0) = 0\ \forall j$$

Now, the partials of g,

$$\begin{aligned} \frac{\partial g}{\partial x_j}(x) &= \frac{\partial}{\partial x_j} \sum_{i=1}^{n} (y_i - f_i(x))^2 \\ &= -2 \sum_{i=1}^{n} (y_i - f_i(x)) \frac{\partial f_i}{\partial x_j}(x) \end{aligned}$$

At x_0, we get,

$$\sum_{i=1}^{n} (y_i - f_i(x_0)) \frac{\partial f_i}{\partial x_j}(x_0) = 0 \ \forall j$$
$$\implies \underbrace{\left(\frac{\partial f_i}{\partial x_j}(x_0)\right)^t}_{Df(x_0)^t} (y - f_i(x_0)) = 0$$

As Df is invertible in U and $x_0 \in U^\circ$ we obtain $y = f(x_0)$, which shows the existence of x_0. The uniqueness follows from (14.1). □

We now set $V := U \cap f^{-1}(W)$. Since U is closed, $V = U^\circ \cap f^{-1}(W)$. Hence, $f|_V : V \to W$ is invertible!

Claim

$f^{-1} \equiv (f|_V)^{-1} : W \to V$ is continuous.

Proof. (14.1) gives,

$$\|f(x_1) - f(x_2)\| \geq \frac{1}{2}\|x_1 - x_2\| \ \forall \ x_1, x_2 \in V \subseteq U$$

equivalently,

$$2\|y_1 - y_2\| \geq \left\|f^{-1}(y_1) - f^{-1}(y_2)\right\| \qquad \text{(where, } y_i = f(x_i)\text{)}$$

which shows that f^{-1} is Lipschitz, hence continuous. □

Claim

f^{-1} is differentiable.

Proof. We fix $y_0 = f(x_0) \in W$ for some $x_0 \in V$ and take $A = Df(x_0)$. As we know,

$$\lim_{h \to 0} \frac{1}{\|h\|}\left(f^{-1}(y_0 + h) - f^{-1}(y_0) - A^{-1}h\right) = 0$$
$$\iff \lim_{y \to y_0} \frac{1}{\|y - y_0\|}\left(f^{-1}(y) - f^{-1}(y_0) - A^{-1}(y - y_0)\right) = 0 \tag{14.3}$$

We set $\phi(h) = f(x_0 + h) - f(x_0) - Ah$ for h in a neighborhood of 0. Now,

$$\begin{aligned} A^{-1}(f(x_0 + h) - f(x_0)) &= h + A^{-1}\phi(h) \\ &= ((x_0 + h) - x_0) + A^{-1}(\phi((x_0 + h) - x_0)) \end{aligned}$$

Also set $y = f(x_0 + h)$. Then,

$$A^{-1}(y - y_0) = f^{-1}(y) - f^{-1}(y_0) + A^{-1}\left(\phi\left(f^{-1}(y) - f^{-1}(y_0)\right)\right)$$
$$\implies -A^{-1}\left(\phi\left(f^{-1}(y) - f^{-1}(y_0)\right)\right) = f^{-1}(y) - f^{-1}(y_0) - A^{-1}(y - y_0) \tag{14.4}$$

So, it is now enough to prove that,

$$\lim_{h \to 0} \frac{1}{\|h\|} A^{-1}\left(\phi\left(f^{-1}(y) - f^{-1}(y_0)\right)\right) = 0$$
$$\iff \lim_{y \to y_0} \frac{1}{\|y - y_0\|}\left(\phi\left(f^{-1}(y) - f^{-1}(y_0)\right)\right) = 0$$

But,

$$\frac{\left\|\phi\left(f^{-1}(y)-f^{-1}(y_0)\right)\right\|}{\|y-y_0\|} = \underbrace{\frac{\left\|\phi\left(f^{-1}(y)-f^{-1}(y_0)\right)\right\|}{\|f^{-1}(y)-f^{-1}(y_0)\|}}_{0} \cdot \underbrace{\frac{\left\|f^{-1}(y)-f^{-1}(y_0)\right\|}{\|y-y_0\|}}_{\leq 2} = 0$$

Hence, the limit (14.3) is true. □

To show that f^{-1} is C^1, observe that all the partials of f^{-1} are rational polynomial functions (with non-zero denominators) of those of f. This completes the proof. □