# Most continuous functions are nowhere differentiable

math center

This note is to demonstrate an application of the Baire category theorem (BCT), which we state now.

**Theorem 0.1** (Baire, Hausdorff)**.** Let $(X, d)$ be a complete metric space, $\{U_n\}_{n\geq 1}$ be a sequence of dense open subsets of $X$. Then the intersection of all the $U_n$ is dense.

*Proof.* Suppose that $V \neq \varnothing$ is an open set disjoint from $\bigcap U_n$. We inductively define two sequences $\{x_i\}$ and $\{r_i\}$, satisfying

- $B(x_1, r_1) \subset V$,
- $B(x_{i+1}, r_{i+1}) \subset B(x_i, \frac{1}{2}r_i) \cap U_i$,
- $r_i \leq 2^{-i}$ for all $i \geq 1$.

This makes $B(x_i, r_i)$ a strictly descending sequence of open balls.

To conclude the proof, we need to show that $x_i$ is a Cauchy sequence (obvious), that its limit $x$ lies in all $U_n$, and that the limit $x$ lies in $V$. We have

$$x_k \in B(x_i, r_i), \quad \forall k \geq i \geq 1,$$

$$\implies x \in \overline{B(x_i, r_i)} \subset \overline{B(x_{i-1}, \tfrac{1}{2}r_{i-1})} \subset D(x_{i-1}, \tfrac{1}{2}r_{i-1}) \subset B(x_{i-1}, r_{i-1})$$

for $i \geq 2$. Now, the last term in the chain lies in $V$ if $i = 2$, and in $U_{i-2}$ otherwise. □

**Remark 0.2.** The result holds under a different assumption that $X$ is locally compact Hausdorff.

## 1 Main result

Next, we give an application which is the claim in the title: "most" continuous functions are differentiable.

First, we need to find an appropriate notion of being a "big" subset of the space of continuous functions. This space is too *big* (for example, it's not locally compact) to have a good measure, like the Lebesgue measure on $\mathbb{R}^n$. I believe this "issue" is handled to some extent in functional analysis, and namely, by the BCT. What we will prove is the following.

**Theorem 1.1.** Let $\mathcal{A}$ be the subset of $\mathscr{C}^0([0,1], \mathbb{R})$ consisting of the functions that are nowhere differentiable on $(0, 1)$. Under the uniform norm, $\mathcal{A}$ is dense in $\mathscr{C}^0([0,1], \mathbb{R})$.

We need the following "viewpoint":

### Input. more continuous than continous functions

As will be clear from the proof, we want these functions to be at least Lipchitz continuous to do basically anything. Thanks to the following theorem, we can avoid the danger of losing every brain cell from considering general continuous functions.

**Theorem 1.2** (Weierstrass Approximation)**.** Polynomial functions are dense in $\mathscr{C}^0([0,1], \mathbb{R})$.

*Proof.* Given in Section 2. □

### Apply BCT

A function in $\mathscr{C}^0([0,1],\mathbb{R})$ is differentiable at at least one point in (0,1) only if it lies in the union of

$$X_{n,m} := \left\{ f : \exists x \in (0,1) \text{ s.t. } 0 < |y-x| < \tfrac{1}{m} \implies \frac{|f(y)-f(x)|}{|y-x|} \leq n \right\},$$

so $\mathcal{A}$ contains the intersection of all $Y_{n,m} := (X_{n,m})^{\mathrm{c}}$. One checks that the subsets $X_{n,m}$ is closed using a compactness argument, so $Y_{n,m}$ is open. (This is not trivial but also not difficult.) Thus by the BCT, it suffices to show that each $Y_{n,m}$ is dense.

By Theorem 1.2, it suffices to construct for each polynomial $P$ a function $h$ of arbitrarily small norm (say $\|h\| < \varepsilon$) such that $P + h \in Y_{n,m}$. This is in fact easy. For example, we can put $h(x) := \varepsilon \sin(cx)$, and take $c \gg 0$ such that

$$\sup(P+h)' \geq \sup h' - \sup|P'| = \varepsilon c - \text{const} \overset{!}{>} n.$$

This shows that $P + h \in Y_{n,m}$ as desired.

## 2 Proof of Weierstrass Approximation

I'll record the proof here because it's pretty interesting.

We are given a continous function $f : [0,1] \longrightarrow \mathbb{R}$, and we need to construct better and better uniform approximations by polynomial functions $f_n$. This suggests that we can define $f_n$ in a way that only depends on the samples $f(k/n)$ of $f$:

$$f_n(x) := \sum_{k=0}^{n} p_k(x) f\left(\frac{k}{n}\right), \quad p_k : \text{ polynomials},$$

such that $p_k$ converges to the characteristic function of $\{x\}$ in an appropriate sense. We have the following observation.

**Theorem 2.1** (Weak Law of Large Numbers)**.** Let $X_1, \cdots, X_n$ be i.i.d. random variables with finite mean $\mu$. Then the sample mean $\overline{X_n}$ converges in probability (=in measure) to $\mu$.

*Proof.* Obvious. □

We can take $X_i \sim \mathrm{Bern}(x)$ $(0 \leq x \leq 1)$, which implies that the sample mean is $1/n$ times a binomial distribution. Namely,

$$\mathrm{P}\left(\overline{X_i} = \frac{k}{n}\right) = \binom{n}{k} x^k (1-x)^{n-k}, \quad 0 \leq k \leq n.$$

We then define $p_k(x)$ as this quantity. Once this is done, the proof that $\|f - f_n\| \longrightarrow 0$ is easy.

*Proof of Theorem 1.2.* Define $f_n$ as described above; then each $f_n(x)$ is a polynomial in $x$. Fix $\varepsilon > 0$, and let $\delta > 0$ be such that $|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon$. Then

$$\begin{aligned}
(f - f_n)(x) &= f(x) - \sum_{k=0}^{n} p_k(x) f\left(\frac{k}{n}\right) \\
&= \sum_{k=0}^{n} p_k(x) \left(f(x) - f\left(\frac{k}{n}\right)\right) \\
&= \sum_{|x-k/n|<\delta} p_k(x) \left(f(x) - f\left(\frac{k}{n}\right)\right) + \sum_{|x-k/n|\geq\delta} p_k(x) \left(f(x) - f\left(\frac{k}{n}\right)\right)
\end{aligned}$$

The first sum is bounded by $\varepsilon$ and the second sum goes to 0 by Theorem 2.1. □

## The End

All main results proven here feels philosophically pleasing...

Compiled on 2025/11/10.
Home page