Teorema da Aproximação Universal

Caio Lins

14 de março de 2021

1 Teorema da aproximação de Weierstrass

Desejamos mostrar que dada uma função contínua $f : [a, b] \to \mathbb{R}$, podemos aproximá-la arbitrariamente bem por funções polinomiais $p : [a, b] \to \mathbb{R}$.

Em outras palavras, seja $C([a, b])$ o espaço das funções contínuas em $[a, b]$. Indicamos por $\|\varphi\|_\infty$ a norma do supremo de uma função limitada $\varphi : [a, b] \to \mathbb{R}$, ou seja,

$$\|\varphi\|_\infty = \sup \{|\varphi(x)| \ ; \ x \in [a, b]\}.$$

Então é verdade que

Teorema 1.1. *Dada $f \in C([a, b])$, para todo $\varepsilon > 0$ existe um polinômio $p : [a, b] \to \mathbb{R}$ tal que*

$$\|f - p\|_\infty < \varepsilon.$$

Inicialmente, observamos que basta provar o teorema para o caso $f \in C([0, 1])$. De fato, dada $f \in C([a, b])$, considere o homeomorfismo $\varphi : [0, 1] \to [a, b]$ dado por $\varphi(x) = a + (b - a)x$, cuja inversa é $\varphi^{-1} : [a, b] \to [0, 1]$ dada por $\varphi^{-1}(x) = \frac{x-a}{b-a}$. Então a função $g = f \circ \varphi$ pertence a $C([0, 1])$ e, dado $\varepsilon > 0$, se existe um polinômio $p(x)$ com $\|g - p\|_\infty < \varepsilon$, temos também, como φ^{-1} é um polinômio de grau 1,

$$\left\|g \circ \varphi^{-1} - p \circ \varphi^{-1}\right\|_\infty < \varepsilon.$$

Como $g \circ \varphi^{-1} = f$ e $p \circ \varphi^{-1}$ é um polinômio, o resultado vale também para $C([a, b])$.

Em seguida, devemos definir a classe de polinômios que utilizaremos na demonstração.

Definição 1.1. Dada $g : X \to \mathbb{R}$ definimos o n-ésimo polinômio de Bernstein de g como

$$B_n(x, g) \overset{\text{def}}{=} \sum_{k=0}^{n} g\left(\frac{k}{n}\right) \binom{n}{k} x^k (1 - x)^{n-k}. \tag{1}$$

Note a semelhança entre os polinômios de Bernstein e a expansão binomial de $(1 + (1 - x))^n$. De fato, temos $B_n(x, 1) = (1 + (1 - x))^n = 1$. Mais geralmente, para toda constante $c \in \mathbb{R}$ tem-se $B_n(x, c) = c$.

Utilizaremos essa semelhança para obter algumas identidades essenciais para a demonstração do Teorema 1.1. Dados p e q reais, começamos considerando a expansão binomial de $(p+q)^n$:

$$(p+q)^n = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k}.$$

Considerando ambos lados da igualdade como funções de p, podemos derivá-los com relação a essa variável, obtendo

$$n(p+q)^{n-1} = \sum_{k=0}^{n} k \binom{n}{k} p^{k-1} q^{n-k}.$$

Multiplicando ambos lados por p/n, ficamos com

$$p(p+q)^{n-1} = \sum_{k=0}^{n} \frac{k}{n} \binom{n}{k} p^k q^{n-k}. \tag{2}$$

Essa é a primeira identidade, válida para todos $p, q \in \mathbb{R}$. Derivando novamente com relação a p e multiplicando ambos lados por p/n obtemos

$$p^2 \left(1 - \frac{1}{n}\right) (p+q)^{n-2} + \frac{p}{n}(p+q)^{n-1} = \sum_{k=0}^{n} \frac{k^2}{n^2} \binom{n}{k} p^k q^{n-k}, \tag{3}$$

a segunda identidade que utilizaremos.

Como consideramos $f, g \in C([0,1])$, segue da Definição 1.1 que se $f \geq 0$, então $B_n(x, f) \geq 0$ e, se $f \leq g$, então $B_n(x, f) \leq B_n(x, g)$.

Com essas ferramentas, podemos então apresentar a

Demonstração do Teorema 1.1. Observamos inicialmente que como f é uma função contínua definida em um compacto, é uniformemente contínua. Portanto, dado $\varepsilon > 0$, existe $\delta > 0$ tal que se $x, y \in [0,1]$ satisfazem $|x - y| < \delta$ então

$$|f(x) - f(y)| < \frac{\varepsilon}{2}.$$

Agora, definimos $M \overset{\text{def}}{=} \|f\|_\infty$ e fixamos $\xi \in [0,1]$. Logo, se $|x - \xi| \geq \delta$ temos

$$|f(x) - f(\xi)| \leq 2M \leq 2M \left(\frac{x - \xi}{\delta}\right)^2.$$

Combinando as duas últimas desigualdades, concluímos que para todo $x \in [0,1]$ vale

$$|f(x) - f(\xi)| \leq 2M \left(\frac{x - \xi}{\delta}\right)^2 + \frac{\varepsilon}{2}. \tag{4}$$

Vamos aproximar f pelos seus polinômios de Bernstein. Seja $B_n(x, f)$ o n-ésimo polinômio de Bernstein de f, avaliado em x. Então

$$\begin{aligned}
|B_n(x, f) - f(\xi)| &= |B_n(x, f - f(\xi))| && (5)\\
&\leq B_n\left(x, 2M \left(\frac{x - \xi}{\delta}\right)^2 + \frac{\varepsilon}{2}\right) && (6)\\
&= \frac{2M}{\delta^2} B_n(x, (x - \xi)^2) + \frac{\varepsilon}{2} && (7)\\
&+ \frac{2M}{\delta^2} \left(B_n(x, x^2) + B_n(x, -2x\xi) + \xi^2\right) + \frac{\varepsilon}{2} && (8)\\
&= \frac{2M}{\delta^2} \left(B_n(x, x^2) - 2\xi B_n(x, x) + \xi^2\right) + \frac{\varepsilon}{2}. && (9)
\end{aligned}$$

Aqui fizemos uso das propriedades de $B_n(x, f)$ que seguem de $x \in [0, 1]$, discutidas anteriormente. Utilizando as equações (2) e (3), com a substituição $p = x$ e $q = 1 - x$, concluímos que

$$B_n(x, x) = x$$

e que

$$B_n(x, x^2) = x^2 \left(1 - \frac{1}{n}\right) + \frac{x}{n}.$$

Substituindo em (9), ficamos com

$$\frac{2M}{\delta^2}\left(B_n(x, x^2) - 2\xi B_n(x, x) + \xi^2\right) + \frac{\varepsilon}{2} = \frac{2M}{\delta^2}\left(x^2\left(1 - \frac{1}{n}\right) + \frac{x}{n} - 2\xi x + \xi^2\right) + \frac{\varepsilon}{2} \tag{10}$$

$$= \frac{2M}{\delta^2}\left(x^2 + \frac{x - x^2}{n} - 2\xi x + \xi^2\right) + \frac{\varepsilon}{2} \tag{11}$$

$$= \frac{\varepsilon}{2} + \frac{2M}{n\delta^2}(x - x^2) + \frac{2M}{\delta^2}(x - \xi)^2. \tag{12}$$

Sendo assim,

$$|B_n(x, f) - f(\xi)| \leq \frac{\varepsilon}{2} + \frac{2M}{n\delta^2}(x - x^2) + \frac{2M}{\delta^2}(x - \xi)^2. \tag{13}$$

Como essa desigualdade vale para todo $x \in [0, 1]$, em especial é válida para $x = \xi$. Fazendo essa substituição, obtemos

$$|B_n(\xi, f) - f(\xi)| \leq \frac{\varepsilon}{2} + \frac{2M}{n\delta^2}(\xi - \xi^2). \tag{14}$$

Facilmente podemos verificar que $\xi - \xi^2 \leq \frac{1}{4}$ para todo $\xi \in [0, 1]$. Logo,

$$|B_n(\xi, f) - f(\xi)| \leq \frac{\varepsilon}{2} + \frac{M}{2n\delta^2}. \tag{15}$$

Por fim, tomando $n > \frac{M}{\varepsilon\delta^2}$, temos $\frac{M}{2n\delta^2} < \frac{\varepsilon}{2}$ e, assim,

$$|B_n(\xi, f) - f(\xi)| < \varepsilon. \tag{16}$$

Como o valor de n obtido para que essa desigualdade seja satisfeita depende apenas de ε (lembramos que δ depende apenas de ε, pela continuidade uniforme de f), ela é válida para todo $\xi \in [0, 1]$, ou seja,

$$\|B_n(\cdot, f) - f\|_\infty < \varepsilon. \qquad \square$$

2 Um pouco sobre Espaços Métricos

Para estudar os próximos resultados, precisaremos do Teorema da Categoria de Baire. Pelo bem da completude do texto, primeiro introduziremos algumas noções relativas a Espaços Métricos que serão utilizadas na demonstração.

Definição 2.1. Dado um conjunto X qualquer, uma *métrica* em X é uma função $d : X \times X \to \mathbb{R}$ tal que:

i) $d(x, x) = 0$;

ii) $d(x, y) > 0$ se $x \neq y$;

iii) $d(x, y) = d(y, x)$;

iv) $d(x, z) \leq d(x, y) + d(y, z)$.

Definição 2.2. Um *espaço métrico* é um par (X, d) onde X é um conjunto e d é uma métrica em X.

Por vezes, onde não houver prejuízo ao entendimento do texto, utilizaremos apenas o nome do conjunto para nos referirmos ao espaço métrico por ele formado.

Definição 2.3. Um suconjunto M de um espaço métrico (X, d) é dito *limitado* se existe $c \in \mathbb{R}$ tal que $d(x, y) \leq c$ para todos $x, y \in M$. Nesse caso, o definimimos o *diâmetro* de M, denotado por $\operatorname{diam} M$, como $\sup \{d(x, y) \; ; \; x, y \in M\}$. Se M é ilimitado, ou seja, dado $c > 0$ existem $x, y \in M$ com $d(x, y) > c$, dizemos que $\operatorname{diam} M = \infty$.

Definição 2.4. Dado um espaço métrico (X, d) e um ponto $a \in X$, chamamos de *bola aberta de raio r centrada em a*, e denotamos por $B(a, r)$, o conjunto

$$\{x \in X \; ; \; d(x, a) < r\}.$$

Definição 2.5. Dado um espaço métrico (X, d) e um subconjunto $Y \subset X$, chamamos de *interior* de Y, e denotamos por $\operatorname{int} Y$, o subconjunto de Y formado pelos elementos $a \in Y$ tais que existe $r > 0$ satisfazendo $B(a, r) \subset Y$.

Definição 2.6. Um subconjunto A de um espaço métrico (X, d) é dito *aberto* se $A = \operatorname{int} A$.

Definição 2.7. Dado um subconjunto M de um espaço métrico (X, d), um ponto $x \in X$ é dito *aderente* a M se toda bola aberta centrada em x tiver interseção não-vazia com M. Chamamos de *fecho* de M, e denotamos por $\overline{M}$, o conjunto dos pontos de aderência de M.

Definição 2.8. Um subconjunto F de um espaço métrico (X, d) é dito *fechado* se $F = \overline{F}$.

Definição 2.9. Dada uma sequência $(x_n)_{n \in \mathbb{N}}$ de elementos do espaço métrico (X, d), ou seja, uma função $f : \mathbb{N} \to X$, dizemos que (x_n) *converge* para $L \in X$ se, dado $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que, para $n \geq n_0$, vale $d(x_n, L) < \varepsilon$. Se para todo $L \in X$ é falso que $\lim x_n = L$, dizemos que (x_n) é *divergente*.

Definição 2.10. Uma sequência $(x_n)_{n \in \mathbb{N}}$ de elementos do espaço métrico X é dita *de Cauchy* se, dado $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que, para $n, m \geq n_0$, vale $d(x_n, x_m) < \varepsilon$. Equivalentemente, (x_n) é de Cauchy se, para $n \geq n_0$, vale $d(x_n, x_{n+p}) < \varepsilon$ para todo $p \in \mathbb{N}$.

Proposição 1. *Toda sequência convergente $(x_n)_{n \in \mathbb{N}}$ no espaço métrico X é de Cauchy*

Demonstração. Seja $L = \lim x_n$. Dado $\varepsilon > 0$, tome $n_0 \in \mathbb{N}$ de modo que, para $n \geq n_0$, valha $d(n, L) < \varepsilon/2$. Então, se $n, m \geq n_0$ temos

$$d(x_n, x_m) \leq d(x_n, L) + d(x_m, L) = \varepsilon/2 + \varepsilon/2 = \varepsilon. \qquad \square$$

Exemplo 1. Embora toda sequência convergente seja de Cauchy, é falso que dado um espaço métrico qualquer, toda sequência de Cauchy convirja para um ponto pertencente a ele. Por exemplo, considerando o conjunto $\mathbb{Q}$ com a métrica d induzida pela métrica de $\mathbb{R}$, temos que toda sequência de racionais convergindo para um irracional é de Cauchy, mas diverge em $\mathbb{Q}$.

Proposição 2. *Se $(x_n)_{n\in\mathbb{N}}$ em um espaço métrico X é de Cauchy e possui um valor de aderência (ou seja, existe uma subsequência convergente (x_{n_k}) de (x_n)), então (x_n) converge para esse valor de aderência.*

Demonstração. Seja $L \in X$ o limite da subsequência (x_{n_k}). Então, dado $\varepsilon > 0$ conseguimos obter $k_0 \in \mathbb{N}$ tal que, se $k > k_0$, então $d(x_{n_k}, L) < \varepsilon/2$. Também conseguimos $n_0 \in \mathbb{N}$ tal que, se $n, m \geq n_0$ então $d(x_n, x_m) < \varepsilon/2$. Tome $\ell > \max\{n_0, k_0\}$. Então claramente $n_\ell \geq \ell$ e, com isso,

$$d(x_\ell, L) \leq d(x_\ell, x_{n_\ell}) + d(x_{n_\ell}, L) < \varepsilon/2 + \varepsilon/2 = \varepsilon. \qquad \square$$

Definição 2.11. Um espaço métrico (X, d) é dito *completo* se toda sequência de Cauchy em X converge para um elemento de X.

Proposição 3. *Um espaço métrico (X, d) é completo se, e somente se, dada uma sequência decrescente $F_1 \supset F_2 \supset \cdots$ de conjuntos não-vazios fechados em X, tais que $\lim \operatorname{diam} F_n = 0$, existe $a \in X$ com*

$$\{a\} = \bigcap_{n=1}^{\infty} F_i.$$

Demonstração. Suponha que X seja completo e considere $(F_n)_{n\in\mathbb{N}}$ como no enunciado do Teorema. Para cada conjunto F_n, escolha $x_n \in F_n$, formando uma sequência $(x_n)_{n\in\mathbb{N}}$ de Cauchy. De fato, como $\lim \operatorname{diam} F_n = 0$, dado $\varepsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que para $n \geq n_0$, temos $d(x, y) < \varepsilon$ para todos $x, y \in F_n$. De $F_1 \supset F_2 \supset \cdots$ concluímos que $n, m > n_0$ implicam $x_n, x_m \in F_{n_0}$ o que implica $d(x_n, x_m) < \varepsilon$.

Da completude de X concluímos que existe $x = \lim x_n$. Como todos F_n são fechados e, para $m \geq n$ temos $x_m \in F_n$, conclui-se que $x \in F_n$ para todo $n \in \mathbb{N}$, ou seja,

$$x \in \bigcap_{n=1}^{\infty} F_n.$$

Suponha, agora, que X seja um espaço métrico no qual toda sequência de fechados como a do enunciado convirja. Seja $(x_n)_{n\in\mathbb{N}}$ uma sequência de Cauchy em X. Defina, para cada $n \in \mathbb{N}$, o conjunto $F_n = \{x_n, x_{n+1}, \dots\}$. Então $(\overline{F_n})_{n\in\mathbb{N}}$ é uma sequência decrescente de conjuntos fechados tais que $\lim \operatorname{diam} \overline{F_n} = \lim \operatorname{diam} F_n = 0$. Por hipótese, existe $a \in \bigcap \overline{F_n}$. Como a é limite de sequência de pontos de F_k para todo $k \in \mathbb{N}$, para cada k podemos escolher $a_{n_k} \in F_k$ de modo que $d(a, a_{n_k}) < 1/k$ e, assim, $\lim a_{n_k} = a$. Claramente $n_k > k$ para todo $k \in \mathbb{N}$, portanto, passando a uma subsequência se necessário, a_{n_k} é subsequência de (x_n) o que implica, como (x_n) é de Cauchy, $\lim x_n = a$. $\square$

Teorema 2.1 (Teorema da Categoria de Baire). *Se (X, d) é um espaço métrico completo e $A_1, A_2, \dots$ são abertos densos em X, então*

$$A = \bigcap_{i=1}^{\infty} A_i$$

é denso em X.

Demonstração. Devemos mostrar que dado V um conjunto aberto em X, temos $A \cap V \neq \emptyset$. Nossa estratégia será construir uma sequência decrescente $F_1 \supset F_2 \supset \cdots$ de conjuntos fechados não-vazios tais que $\lim \operatorname{diam} F_n = 0$ e, para todo $n \in \mathbb{N}$, $F_n \subset A_n \cap V$. Então, pela Proposição 3, o ponto x que satisfaz $\{x\} = \bigcap F_n$ é tal que $x \in V$ e $x \in F_n \subset A_n$ para todo n, ou seja, $x \in A$ e, portanto, $x \in A \cap V$.

Começamos obeservando que, como A_1 é denso, $A_1 \cap V$ é um conjunto aberto não-vazio. Logo, existe B_1 bola aberta não-vazia de raio menor que 1, tal que $\overline{B_1} \subset A_1 \cap V$. Suponha, agora, definidos $B_1, \ldots, B_n$ de forma que, para todo $1 < k \leq n$, B_k é uma bola aberta não-vazia de raio menor que $1/k$ tal que $\overline{B_k} \subset V \cap A_k \cap B_{k-1}$. Novamente, como A_{n+1} é denso, $A_{n+1} \cap B_n$ é um conjunto aberto não vazio. Logo, definimos B_{n+1} como uma bola aberta não-vazia contida em $A_{n+1} \cap B_n$, de raio menor que $1/(n+1)$ tal que $\overline{B_{n+1}} \subset A_{n+1} \cap B_n \subset A_{n+1} \cap V$.

Com isso, obtemos uma sequência decrescente $B_1 \supset \cdots \supset B_n \supset \cdots$ de bolas abertas não-vazias, com o raio de B_n menor que $1/n$, cujos fechos $\overline{B_1} \supset \cdots \supset \overline{B_n} \supset \cdots$ formam uma sequência decrescente de conjuntos fechados não-vazios, com $\operatorname{diam} \overline{B_n} \leq 1/n$ e $\overline{B_n} \subset A_n \cap V$ para todo $n \in \mathbb{N}$, o que, como apontado anteriormente, termina a prova. □

3 O Décimo Terceiro Problema de Hilbert

Diferentemente do Teorema da Aproximação de Weierstrass, o 13° problema de Hilbert não trata de aproximações, mas de representações exatas de funções. Mais especificamente, Hilbert postulou (utilizando a linguagem matemática de sua época) que existem funções contínuas de $\mathbb{I}^3$ em $\mathbb{R}$, onde $\mathbb{I} = [0, 1]$, que não podem ser expressas por meio da composição e adição de funções de $\mathbb{R}^2$ em $\mathbb{R}$.

Décadas após ser postulada, essa conjectura eventualmente foi demonstrada *falsa*. A prova foi dada por Vladimir Igorevich Arnol'd, 14 anos após a morte de Hilbert. Ele e seu orientador de Doutorado, Andrej Nikolajewitsch Kolmogorov, provaram que, na verdade, toda função contínua $f : \mathbb{I}^n \to \mathbb{R}$ pode ser expressa como composições e adições de funções contínuas de $\mathbb{R} \to \mathbb{R}$. Com o trabalho de outros matemáticos, esse resultado foi generalizado. Uma dessas generalizações é apresentada no Teorema a seguir.

Teorema 3.1 (Kolmogorov, Arnol'd, Kahane, Lorentz e Sprecher)**.** *Para todo $n \in \mathbb{N}$ com $n \geq 2$, existem números reais $\lambda_1, \lambda_2, \ldots, \lambda_n$ e funções contínuas $\varphi_k : \mathbb{I} \to \mathbb{R}$, para $k = 1, \ldots, 2n+1$, com a propriedade de que para toda função contínua $f : \mathbb{I}^n \to \mathbb{R}$ existe uma função contínua $g : \mathbb{R} \to \mathbb{R}$ tal que, para todo $(x_1, \ldots, x_n) \in \mathbb{I}^n$,*

$$f(x_1, \ldots, x_n) = \sum_{k=1}^{2n+1} g(\lambda_1 \varphi_k(x_1) + \cdots + \lambda_n \varphi_k(x_n)). \tag{17}$$

Nos atentaremos ao caso especial em que $n = 2$:

Teorema 3.2. *Existem $\lambda \in \mathbb{R}$ e funções $(\varphi_1, \ldots, \varphi_5) \in [C(\mathbb{I})]^5$ tais que, para toda função $f \in C(\mathbb{I}^2)$ existe $g : \mathbb{R} \to \mathbb{R}$, contínua, satisfazendo, para todos $(x_1, x_2) \in \mathbb{I}^2$,*

$$f(x_1, x_2) = \sum_{k=1}^{5} g(\varphi_k(x_1) + \lambda\varphi_k(x_2)).$$

Para a prova que apresentaremos, necessitamos de alguns lemas

4 O Teorema de Hahn-Banach

4.1 O Lema de Zorn

Aqui apresentaremos alguns conceitos de Teoria dos Conjuntos que serão utilizados na demonstração de Teorema de Hahn-Banach. Dado um conjunto X uma *relação* em X é um subconjunto R de $X \times X$. Escreveremos xRy para indicar que $(x, y) \in R$.

Uma *relação de ordem parcial* em X é uma relação que satisfaz, dados $x, y, z \in X$:

i) xRx (*reflexividade*);

ii) Se xRy e yRx, então $x = y$ (*antisimetria*);

iii) Se xRy e yRz, então xRz (*transitividade*);

O termo *parcial* é utilizado para indicar que podem existir elementos de X não relacionados por essa ordem. Caso tenhamos xRy ou yRx para todos $x, y \in X$, então R é dita uma relação de ordem *total*. Utilizaremos o símbolo $\prec$ para indicar uma ordem parcial. Dizemos então que $(X, \prec)$ é um conjunto parcialmente ordenado. Se $\prec$ também é total, $(X, \prec)$ é totalmente ordenado. Claramente todo subconjunto de X também é parcialmente ordenado, com a ordem induzida por $\prec$.

Um *elemento maximal* de um conjunto ordenado X é um elemento x tal que se $y \in X$ com $x \prec y$, então $y = x$. Uma *cota superior* de um subconjunto $Y \subset X$ é um elemento $x \in X$ tal que $y \prec x$ para todo $y \in Y$. Observe que se $\prec$ não é total, não necessariamente um elemento maximal de $Y \subset X$ é uma cota superior de Y.

Dois exemplos usuais de conjuntos ordenados são $\mathbb{R}$ e $\mathcal{P}(X)$, para um dado conjunto X. O primeiro é um conjunto totalmente ordenado pela ordem usual $\leq$, o qual não possui elemento maximal. O segundo se torna um conjunto ordenado ao dizermos que dados $A, B \subset X$, temos $A \prec B$ se, e somente se $A \subset B$. Essa ordem não é total e o único elemento maximal de $\mathcal{P}(X)$ é o próprio X.

Axioma (Lema de Zorn)**.** *Todo conjunto parcialmente ordenado, tal que todos seus subconjuntos totalmente ordenados possuam cota superior, possui elemento maximal.*

Agora provaremos um resultado que exemplifica como o Lema de Zorn geralmente é utilizado.