Problemas fundamentais da teoria da aproximação funcional

por **Luís G. M. de Albuquerque**

1. Introdução.

Considere-se o seguinte exemplo, bem conhecido, de aproximação funcional:

Seja $f(x)$ uma função de variável real que pode ser desenvolvida em série inteira de x

$$(1.\ 1) \qquad f(x) = \sum_{k=0}^{\infty} a_k x^k$$

em dado intervalo $[-\rho, \rho]$; e designemos por $s_m(x)$ a soma dos $m+1$ primeiros termos de (1. 1):

$$s_m(x) = \sum_{k=0}^{m} a_k x^k .$$

A série (1. 1) é uniformemente convergente em $[-\rho, \rho]$, e sabe-se que, para $\delta > 0$, é possível determinar o inteiro $N(\delta)$ de tal modo que se tem

$$(1.\ 2) \qquad |f(x) - s_m(x)| < \delta$$

qualquer que seja $x \in [-\rho, \rho]$, desde que $m > N(\delta)$. Assim, fixado nestas condições um valor m de n, $s_m(x)$ é um polinómio de grau m em x e coeficientes conhecidos

$$(1.\ 3) \qquad a_k = \frac{1}{k!} f^{(k)}(o) \quad (k = 1, 2, \cdots m);$$

e a função dada, $f(x)$, pode ser substituida em todo o intervalo $[-\rho, \rho]$ por este polinómio, com um erro de módulo inferior a δ.

Neste caso particular o problema da aproximação funcional consiste, portanto, em substituir num intervalo conhecido uma dada função por outra, em geral mais simples, de modo que o erro resultante dessa substituição não exceda uma quantidade prèviamente fixada.

Este exemplo sugere, no entanto, um outro problema. Considere-se a função de variável real, $f(x)$, definada em $[a, b]$, e tome-se o polinómio de grau m e coeficientes quaisquer

$$(1.\ 4) \qquad p_m(x) = \sum_{k=0}^{m} \alpha_k x^k ;$$

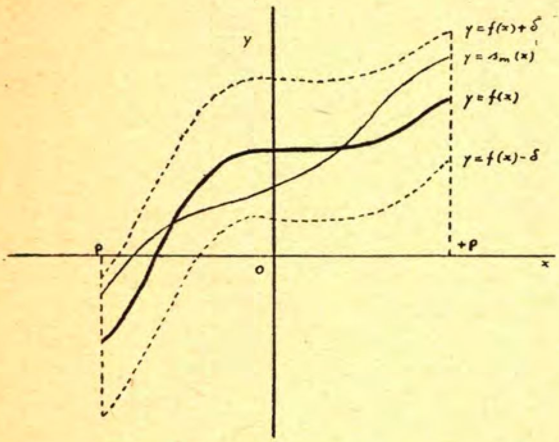
que valores α_k^0 devem ser atribuidos aos coeficientes α_k do polinómio (1. 4) para que, sendo

$$p_m^0(x) = \sum_{k=0}^{m} \alpha_k^0 x^k ,$$

se tenha

$$(1.\ 5) \qquad \max_{a \le x \le b} |f(x) - p_m^0(x)| = \min_{\alpha_k} \max_{a \le x \le b} |f(x) - p_m(x)| \ ?$$

Se existe um polinómio $p_m^0(x)$ nestas condições diremos que ele é a *melhor aproximacão* de $f(x)$ em (a, b) dada por um polinómio de grau m. Se $f(x)$ é desenvolvível em série inteira, $[a, b]$ é o intervalo de convergência e $m > N(\delta)$, o mínimo procurado é decerto inferior a δ. Mas deve se observar que os dois problemas apresentados são bem diferentes. No primeiro caso são conhecidos os coeficientes (1. 3) dos polinómios a construir, e pretende-se determinar o grau de um polinómio que faz o módulo (1. 2) inferior a uma quantidade dada inicialmente. À solução $s_m(x)$ corresponde uma curva $y = s_m(x)$ contida entre $y = f(x) + \delta$ e $y = f(x) - \delta$ em $[-\rho, \rho]$ (figura); e todo o polinómio de coeficientes (1. 3) e grau superior a m satisfaz ao problema. No segundo caso, pelo contrário, fixa-se o grau do polinómio $p_m(x)$, e procuram-se os valores dos coeficientes



desse polinómio que tornam mínimo o valor $\max\limits_{a \lesseqgtr x \lesseqgtr b} |f(x) - p_m(x)|$; à solução $p_m^0(x)$ pode corresponder uma curva que saia fora da faixa limitada por $y = f(x) + \delta$ e $y = f(x) - \delta$.

O primeiro destes problemas entra no quadro da teoria das séries de funções. Quanto ao segundo, que foi resolvido pelo primeiro de dois célebres teoremas de WEIERSTRASS demonstrados em 1885, constitue juntamente com os resultados obtidos por TCHEBICHEF, um dos primeiros passos da teoria da aproximação funcional.

Notemos agora que designando por distância de $f(x)$ a $p_m(x)$ em (a, b) o

$$\max_{a \lesseqgtr x \lesseqgtr b} |f(x) - p_m(x)|,$$

o problema resolvido pelo teorema de WEIERSTRASS pode-se enunciar assim: dada uma função $f(x)$, determinar o polinómio de grau m para o qual é mínimo o valor daquela distância.

Nada impede, porém, que a noção de distância entre $f(x)$ e $p_m(x)$ não seja compreendida doutro modo. Se admitirmos, por exemplo, que se designa por distância de $f(x)$ a $p_m(x)$ em (a, b), o valor do integral

$$(1.\ 6) \qquad \int_a^b [f(x) - p_m(x)]^2 \, dx.$$

o enunciado do problema tem outra significação; e se existe um polinómio $p_m(x)$ que torna mínimo o integral (1. 6), diremos que ele é *a melhor aproximação em média* de $f(x)$ em (a, b).

O que fica escrito já nos permite enunciar com toda a generalidade o problema fundamental da teoria da aproximação funcional (1).

Consideremos definidas sobre um dado conjunto de pontos P *de um espaço com qualquer número de dimensões, as funções* f(P) *e* $\varphi(P; \lambda_1, \cdots \lambda_n)$, *dependendo esta última de* n *parâmetros arbitrários* $\lambda_1, \cdots \lambda_n$; *pretende-se determinar o valor destes parâmetros de tal modo que a distância entre as funções* f(P) *e* $\varphi(P; \lambda_1, \cdots, \lambda_n)$ *seja mínima no conjunto considerado.*

É claro que a solução do problema assim enunciado dependerá da noção de distância entre duas funções que se tenha convencionado introduzir.

2. Espaços vectoriais.

Tomemos o conjunto E constituído pelos elementos $x, y, \cdots$, a que chamaremos pontos ou vectores, a respeito dos quais está definida uma relação de tal modo que, dados $x, y \in E$, se tenha $x = y$ ou $x \neq y$, gozando $x = y$ das propriedades reflexiva, simétrica e transitiva (2). Seja ainda $\mathfrak{K}$ o corpo dos números complexos $\alpha, \beta, \cdots$.

O conjunto E é um *espaço vectorial* (ou *linear*) a respeito do corpo $\mathfrak{K}$ quando:

(1) N. I. ACHIESER, *Vorlesungen über Approximations-theorie,* Berlim, 1953, pág. 1.

(2) Conf. o artigo de J. DIONÍSIO: *Os espaços métricos e a Análise clássica: o método do ponto fixo, Gazeta de Matemática,* n.os 62 e 63, § 1.

v_1) se define sobre E um operador $+$, adição de vectores, a respeito do qual E é um grupo abeliano, isto é, de tal maneira que se verificam os seguintes axiomas:

v_1') se $x, y \in E$, o vector $z = x + y \in E$ fica univocamente determinado;

v_1'') $x + y = y + x$;

v_1''') há em E um vector zero, $0 \in E$, tal que $x + 0 = x$ qualquer que seja $x \in E$.

v_2) se define um produto αx dos elementos de $\mathfrak{K}$ pelos vectores de E, de tal modo que, para $\alpha, \beta \in \mathfrak{K}$ e $x, y \in E$, se tem:

v_2') $\alpha(x + y) = \alpha x + \alpha y$ e $(\alpha + \beta) x = = \alpha x + \beta x$;

v_1'') $\alpha \cdot (\beta x) = (\alpha \cdot \beta) x$;

v_2''') existe em $\mathfrak{K}$ uma unidade, $1 \in \mathfrak{K}$, tal que $1 \cdot x = x$, para todo o $x \in E$.

Como exemplos de espaços vectoriais citaremos:

I) *O espaço* $\mathrm{E_n}$ *de elementos definidos por* n *números complexos ou reais*

$$\mathrm{x} \equiv \{\mathrm{a}_1, \cdots \mathrm{a}_n\}, \mathrm{y} \equiv \{\mathrm{b}_1, \cdots \mathrm{b}_n\}, \cdots$$

com n *qualquer, finito,* desde que definamos a adição de dois elementos x e y por:

$$x + y = \{a_1 + b_1, \cdots a_n + b_n\}$$

e o produto de qualquer $\alpha \in \mathfrak{K}$ por um elemento x assim:

$$\alpha x \equiv \{\alpha \cdot a_1, \cdots \alpha \cdot a_n\},$$

é um espaço vectorial: as condições v_1) e v_2) verificam-se desde que $0 \equiv \{0, \cdots, 0\}$ seja considerado como vector zero de E_n.

Em particular, os vectores do plano euclidiano e do espaço euclidiano constituiem espaços vectoriais, E_2 e E_3, relativamente ao corpo de números reais, quando a adição definida por v_1) é a adição vectorial, e o produto dado por v_2) tem o sentido corrente do produto de um escalar por um vector.

II) *O espaço* C *das funções contínuas* $\mathrm{x(t)}, \mathrm{y(t)}, \cdots$ *definidas num dado conjunto fechado e limitado,* D, *é um espaço vectorial relativamente ao corpo dos números reais* [1]. Para que os axiomas v_1) e v_2) sejam verificados bastará tomar como vector zero a função $x(t) \equiv 0$ em D, e manter as definições habituais de adição de funções e de produto de uma função por um número real.

III) *O espaço* P *dos polinómios de coeficientes reais ou imaginários* é também num espaço vectorial, relativamente ao corpo correspondente, desde que se mantenha o sentido algébrico na adição de dois polinómios ou no produto de um número por um polinómio.

3. Bases. Variedades lineares.

Dados num espaço vectorial E n vectores não nulos $x_k \in E\,(k = 1, \cdots, n)$, sejam $\alpha_k\,(k = = 1, \cdots, n)$ n números quaisquer do corpo $\mathfrak{K}$; se a relação linear

$$\text{(3. 1)} \qquad \sum_{k=1}^{n} \alpha_k x_k = 0$$

tem soluções distintas da solução trivial $\alpha_k = 0\,(k = 1, \cdots, n)$, diremos que os vectores x_k são *linearmente dependentes.* No caso contrário os x_k são *linearmente independentes,* constituindo um *sistema linearmente independente de ordem* n (n é o número de vectores).

Um sistema de infinitos vectores $x_k\,(k = = 1, 2, \cdots)$ é linearmente independente quando qualquer número finito de vectores do sistema é linearmente independente.

Teorema A. *É condição necessária e suficiente para que os vectores não nulos* $\mathrm{x_k}$ $(\mathrm{k} = 1, 2, \cdots \mathrm{n})$ *sejam linearmente dependentes,*

---

[1] J. J. Dionísio no artigo citado, Exemplo 2, § 2, mostrou que C é um espaço métrico.

que algum x_j, *com* $2 \leqq j \leqq n$, *seja uma combinação linear dos anteriores.*

Demonstração. Algum x_j há-de ser linearmente dependente com os vectores que o precedem, pois em virtude da hipótese esta afirmação tem ao menos lugar para $j = n$; seja x_r o primeiro vector nestas condições $(r \leqq n)$

$$\sum_{i=1}^{r} \alpha_i x_i = 0 \tag{3. 2}$$

não pode ser $\alpha_r = 0$, de contrário (3. 2) dava $\sum_{i=1}^{r-1} \alpha_i x_i = 0$, e x_i não seria o primeiro vector com a propriedade indicada; com $\alpha_r \neq 0$ tira-se de (3. 2):

$$x_r = - \sum_{i=1}^{r-1} \frac{\alpha_i}{\alpha_r} x_i$$

e esta relação mostra que x_r é uma combinação linear dos $x_i (i = 1, \cdots, r-1)$ que o antecedem; assim, a condição é necessária. Mas também é suficiente. Se o primeiro vector nas condições do enunciado é x_n, a afirmação é trivial. Seja então $x_r (r < n)$ o primeiro vector que é combinação linear dos anteriores; é claro que os $x_i (i = 1, \cdots, r)$ são linearmente dependentes e $\sum_{i=1}^{r} \alpha_i x_i$ tem a solução não trivial $\alpha_i = \alpha_i^0 (i = 1, 2, \cdots r)$; nestas condições também $\sum_{i=1}^{n} \alpha_i x_i = 0$ tem a solução não trivial

$$\alpha_i = \alpha_i^0 (i = 1, \cdots, r), \alpha_i = 0 (i = r+1, \cdots, n),$$

e os $x_i (i = 1, 2, \cdots n)$ são linearmente dependentes.

Suponhamos que é n a maior ordem dos sistemas linearmente independentes contidos num dado espaço vectorial E; e seja

$$B \equiv \{x_1, \cdots, x_n\} \quad \text{ou} \quad \{x_k\}_{k=1,\cdots,n} \tag{3. 3}$$

um desses sistemas; nestas condições chamaremos a B uma *base* do espaço E, dizendo-se que ele tem *dimensão* n.

Qualquer vector $x \in E$ é linearmente dependente com os vectores da base

$$\sum_{i=1}^{n} \alpha_i x_i + \alpha x = 0$$

e como não pode ser $\alpha = 0$ (de contrário os vectores da base seriam linearmente dependentes) tem-se, com $\beta_i = -\alpha_i / \alpha$

$$x = \sum_{i=1}^{n} \beta_i x_i \tag{3. 4}$$

isto é: *cada vector* $x \in E$ *pode-se exprimir como combinação linear dos vectores de qualquer base do espaço.* A representação (3. 4) do vector x nos vectores da base é única: pois se tivessemos $x = \sum_{i=1}^{n} \gamma_i x_i$, subtraindo de (3. 4) obtinhamos $\sum_{i=1}^{n} (\beta_i - \gamma_i) x_i = 0$, que só pode ter a solução trivial $\beta_i = \gamma_i$. Os coeficientes $\beta_i (i = 1, \cdots, n)$, univocamente determinados, chamam-se as componentes do vector $x \in E$ relativos à base B.

Por outro lado, considerando o espaço vectorial E (dimensão n), de base $\{x_k\}_{k=1,2,\cdots n}$, conclui-se não ser possível exprimir linearmente todos os vectores do espaço em menos de n vectores linearmente independentes $y_1, \cdots y_m (m < n)$; porque se assim fosse, os vectores $x_k (k = 1, \cdots n)$ podiam-se definir como combinações lineares dos $y_i (i = 1, \cdots m)$; nesse caso teriamos os x_k como n combinações lineares distintas de $m < n$ vectores, e eles seriam linearmente dependentes, (como o leitor verificará recorrendo à teoria dos sistemas de equações lineares homogéneas) em discordância com a hipótese de ser $\{x_k\}_{k=1,\cdots n}$ uma base.

Admitamos agora que, por maior que seja n, é sempre possível destacar do espaço E um sistema de n vectores linearmente independentes. Diz-se então que E *tem dimensão infinita*, e qualquer sistema infinito de vectores linearmente independentes constitui uma base (infinita) do espaço. Esta definição não nos autoriza, contudo, a substituir para este caso a soma por uma série na igualdade (3. 4); só em espaços vectoriais que verificam novas condições isso é legitimo, como teremos oportunidade de referir adiante.

Seja E um espaço vectorial de dimensão finita ou infinita, e $x_1, x_2, \cdots x_r$ ($r \leqq n$, se n é a dimensão finita do espaço) r vectores linearmente independentes de E; seja V o conjunto de todos os vectores de E que se podem exprimir como combinações lineares dos $x_k (k=1, 2, \cdots r)$:

$$x \in V \text{ quando } x = \sum_{k=1}^{r} \alpha_k x_k ;$$

é $V \subset E$, e como $O \in V$ (combinação dos x_k que corresponde a $\alpha_k = 0$, $k=1, 2, \cdots r$) é *V um sub-espaço vectorial de dimensão* r *e com a base* $\mathrm{B} = \{\mathrm{x}_1, \cdots \mathrm{x}_r\}$; diremos que *$V$ é uma variedade linear do espaço vectorial* E, *gerada pela base* B.

Sejam agora M_1 e M_2 dois sistemas de vectores de E, e $B_1 \subset M_1$, $B_2 \subset M_2$ os conjuntos constituídos pelo maior número de vectores linearmente independentes de M_1 e M_2; os dois sistemas considerados dizem-se *equivalentes* quando se identificam as variedades lineares geradas por B_1 e B_2. É claro que esta circunstância tem lugar quando, e só quando, cada vector de um dos sistemas M_1 ou M_2 se pode exprimir como combinação linear de vectores do outro.

Consideremos alguns exemplos. O espaço E_n considerado em I) do número precedente tem a dimensão n, pois pode-se tomar como base de E_n o sistema de vectores linearmente independentes

$$e_1 = \{1, 0, \cdots 0\}, \ e_2 = \{0, 1, \cdots 0\}, \cdots \cdots e_n = \{0, 0, \cdots 1\};$$

as componentes de qualquer vector

$$x = \{a_1 a_2, \cdots a_n\}$$

nesta base são $a_i (i = 1, 2, \cdots n)$ pois

$$x = a_1 e_1 + a_2 e_2 + \cdots + a_n e_n .$$

Do mesmo modo: o conjunto dos polinómios de grau $< n$ e coeficientes quaisquer, P_{n-1}, é um espaço vectorial de dimensão n com a base

$$(3.\ 5) \qquad B \equiv \{1, t, t^2, \cdots, t^{n-1}\} .$$

Como exemplo de espaço com dimensão infinita serve o espaço P dos polinómios, referido no exemplo III) do número 2. Qualquer que seja n a combinação linear $\sum_{k=0}^{n} \alpha_k t^k = 0$ só tem a solução trivial $\alpha_k = 0 (k=0, \cdots, n)$; deste modo os vectores

$$1, t, t^2, \cdots t^n, \cdots$$

constituem um sistema infinito linearmente independente, que pode ser tomado como base de P.

Finalmente: $P_n \subset P$ é uma variedade linear deste último espaço, gerada pela base (3. 5).

4. Espaços vectoriais normados. Espaços separáveis e completos. Espaço de Banach.

Um espaço vectorial E, definido sobre o corpo $\mathfrak{K}$, diz-se *normado* quando a cada $x \in E$ se faz corresponder um número real $\| x \| \geq 0$, a *norma de* x, tal que se verifiquem as seguintes condições:

n_1) $\quad \| x \| = 0$ se e só se $x = 0$;

n_2) $\quad \| \alpha \cdot x \| = | \alpha | \cdot \| x \|$, com $\alpha \in \mathfrak{K}$;

n_3) $\quad \| x + y \| \leqq \| x \| + \| y \|$.

Se na última condição a igualdade tem lugar quando e só quando $y = \alpha x$, diz-se que a *norma é forte*.

Todo o espaço vectorial normado é metrizável, isto é, pode-se nele definir uma função de distância sobre cada par ordenado de elementos $x, y \in E$, de tal modo que se verifiquem as condicões δ_1) e δ_2)(1) exigidas para que E seja um espaço métrico: na verdade, pondo $\delta(x, y) = \| x - y \|$ tem-se $\delta(x, y) = 0$ quando e só quando $x = y$, e δ_1) é válida; por outro lado, $x - y = (x - z) + (z - y)$ donde, em virtude de n_1) e n_2) $\delta(x, y) \leqq \delta(x, z) + \delta(y, z)$, e δ_2) também se verifica.

Assim, a sucessão $\{x_n\}$ de vectores de um espaço vectorial normado E tem por limite um vector x quando $\delta(x, x_n) < \varepsilon$ para $n > N(\varepsilon)$; ou, o que é o mesmo, quando $\| x - x_n \| < \varepsilon$ para $n > N(\varepsilon)$.

Também se conclui logo que a norma $\| x \|$ é uma função contínua de x, quer dizer, se $x_n \to x$ também $\| x_n \| \to \| x \| \, (n \to \infty)$; pois se $x_n \to x$ é $\| x_n - x \| \to 0$ e, para $n > N(\varepsilon)$, $\| x - x_n \| < \varepsilon$; mas como $x_n = x + (x_n - x)$, em virtude de n_3) é $\| x_n \| \leqq \| x \| + \| x_n - x \|$ ou $\| x_n \| - \| x \| \leq \| x_n - n \|$; anàlogamente $\| x \| - \| x_n \| \leqq \| x_n - x \|$. É portanto

$$\Big| \| x \| - \| x_n \| \Big| < \varepsilon$$

para $n > N(\varepsilon)$ como desejávamos.

No artigo citado (§ 1), J. J. Dionísio indicou que dada uma sucessão $\{x_n\}$ de elementos de um espaço métrico E, a condição de Cauchy

(4. 1) $\delta(x_n, x_m) < \varepsilon$ para $n, m > N(\varepsilon)$

não implicava necessàriamente a existência de um limite da sucessão. Porém, se a condição de Cauchy é suficiente para garantir que o limite de $\{x_n\}$ existe, diz-se que o *espaço é completo*(2). Note-se que a condição para que um espaço vectorial normado, E, seja completo pode ser expressa doutro modo: *o espaço é completo se para toda a sucessão de vectores* $\{\mathrm{x_n}\}$ *de* E *que satisfaça à condição* $\lim\limits_{m, n \to \infty} \| \mathrm{x_n} - \mathrm{x_m} \| = 0$ *existe um vector* $\mathrm{x} \in \mathrm{E}$ *tal que* $\lim\limits_{n \to \infty} \| \mathrm{x} - \mathrm{x_n} \| = 0$.

(1) V. J. J. Dionísio, art. cit. § 1.

(2) No § 2 daquele trabalho o Autor dá exemplos de espaços métricos nestas condições.

Um espaço vectorial normado completo diz-se um *espaço de* Banach. De todo o espaço vectorial normado E se pode obter um espaço de Banach: basta juntar aos vectores de E os elementos limites que não pertencem ao espaço e são definidos por sucessões $\{x_n\}$ de E que satisfazem à condição de Cauchy. O espaço de Banach assim obtido é o fecho $\bar{E}$ de E (sendo E denso em $\bar{E}$).

Um espaço vectorial normado é *separável* quando E contém um sub-conjunto numerável, $N \subset E$, denso em E. Ou seja: quando existe uma sucessão numerável de elementos de E, $\{x_n\}$, tal que sendo $\delta > 0$ e $x \in E$ arbitràriamente escolhidos, pelo menos um elemento $x_{n'}$, de $\{x_n\}$ verifica a condição:

$$\delta(x, x_{n'}) = \| x - x_{n'} \| < \delta .$$

Vamos indicar alguns exemplos:

II) *O espaço* C *das funções contínuas definidas num dado conjunto limitado e fechado,* D, (já indicado no número 2) é um espaço vectorial normado desde que se tome como norma de $x(t) \in C$

$$\| x(t) \| = \sup_{x \in D} | x(t) | .$$

Esta definição respeita a métrica introduzida no art. cit. de J. J. Dionísio (§ 2, exemplo (3)).

O espaço C é completo quando se tome a convergência uniforme como noção de convergência; pois toda a série uniformemente convergente de funções contínuas tem por soma uma função contínua.

IV) *O espaço* m *das sucessões limitadas,* $\mathrm{x} = \{a_i\}_1^\infty$ *com* $|a_i| < \mathrm{M}$, *é um espaço de* Banach. J. J. Dionísio mostrou (1) que m é um espaço

(1) Art. cit., § 2, exemplo (2).

métrico *completo*; resta-nos, portanto, indicar que também é um espaço *vectorial normado*: para o que definiremos como norma de qualquer $x \in m$ o valor $\| x \| = \sup_i | a_i |$; o leitor verificará fàcilmente que tem lugar as condições n_1), n_2) e n_3), e que é respeitada a métrica definida no artigo citado.

V) *O espaço* $P[a, b]$ *dos polinómios de variável real e coeficientes quaisquer definidos num dado intervalo limitado e fechado* $[a, b]$, *é um espaço vectorial normado separável.* Como em III) verificariamos que $P[a, b]$ é um espaço vectorial; mas podemos definir a norma de qualquer $p(t) \in P[a, b]$ de modo análogo à do exemplo II):

$$p(t) = \sup_{t \in [a, b]} | p(t) | .$$

Como o conjunto dos polinómios cujos coeficientes têm parte real e parte imaginária racionais é numerável e denso em $P[a, b]$, o espaço é separável.

5. Espaço de Hilbert.

Chama-se *espaço de* HILBERT a todo o espaço vectorial E, definido sobre um corpo $\mathfrak{K}$, onde se defina uma operação denominada *produto escalar (interno ou hermítico)* do seguinte modo: a todo o par ordenado de vectores $x, y \in E$ corresponde um número complexo (x, y) — o seu *produto escalar* — que verifica as condições:

e_1) $(y, x) = \overline{(x, y)}$ ($\overline{a}$ conjugado de a);

e_2) $(\alpha x + \beta y, z) = \alpha (x, z) + \beta (y, z)$ com $\alpha, \beta \in \mathfrak{K}$;

e_3) $(x, x) \geqq 0$, tendo lugar a iguadade quando e só quando $x = 0$.

e_2) exprime a linearidade do produto escalar relativamente ao primeira factor; recorrendo a e_1 o leitor poderá estabelecer que a linearidade em relação ao segundo factor se exprime por:

e'_2) $(x, \alpha y + \beta z) = \overline{\alpha} (x, y) + \overline{\beta} (x, z)$

O produto escalar definido por e_1), e_2) e e_3) goza de propriedades importantes. Assim:

$$(5.1) \qquad (\alpha x, \alpha x) = | \alpha |^2 \cdot (x, x);$$

de facto, empregando e_1) e e_2) tem-se:

$$(\alpha x, \alpha x) = \alpha (x, \alpha x) = \alpha \overline{(\alpha x, x)} = \alpha \cdot \overline{\alpha} \, \overline{(x, x)} = = | \alpha |^2 \cdot (x, x)$$

pois se é $\alpha = \rho e^{i\theta}, \overline{\alpha} = \rho \cdot e^{-i\theta}$, e portanto $\alpha \cdot \overline{\alpha} = \rho^2$. Por outro lado:

$$(5.2) \qquad | (x, y) | \leqq \sqrt{(x, x)} \cdot \sqrt{(y, y)};$$

tome-se um vector z tal que $(z, z) = 1$ e seja $(x, z) = \alpha$; tem-se:

$$0 \leqq (x - \alpha z, x - \alpha z) = (x, x) - (x, \alpha z) - - (\alpha z, x) + | \alpha |^2 \cdot (z, z)$$

ou $\qquad 0 \leqq (x, x) - | \alpha |^2 \qquad (*)$

pois $(z, z) = 1$ e $(x, \alpha z) + (\alpha z, x) = 2 \cdot \alpha \cdot \overline{\alpha} = = 2 | \alpha |^2$; de (*) tira-se

$$| \alpha |^2 = | (x, z) |^2 \leqq (x, x)$$

e pondo $z = \dfrac{y}{\| y \|}$, qualquer que seja y, fica

$$\left| \left(x, \frac{y}{\| y \|} \right) \right|^2 \leqq (x, x)$$

ou $\qquad | (x, y) |^2 \leqq (x, x) \cdot \| y \|^2 = (x, x) \cdot (y, y)$

como desejavamos demonstrar [1]

Note-se que em (5.2) o sinal de igualdade só tem lugar quando $x = \alpha z$ (em virtude de e_3), e portanto quando $x = \dfrac{\alpha}{\| y \|} y = \beta y$; isto é: na expressão (5.2) é válido o sinal de

[1] C. J. EVERETT e H. J. RYSER, «The Gram Matrix and HADAMARD Theorem», *American Math. Montly*, vol. 53 (1946), p. 21.

igualdade quando x e y são linearmente dependentes. (5. 2) é correntemente designada por *desigualdade de* CAUCHY-BUNJAKOWSKI.

A definição de produto escalar e a desigualdade (5. 2) dão ainda lugar a

$$(5.\,3)\qquad \sqrt{(x+y\,,x+y)} \leqq \sqrt{(x\,,x)}+\sqrt{(y\,,y)}\,,$$

que deixamos ao leitor o trabalho de estabelecer a partir de $0 \leqq (x+y\,,x+y)$. Importa assinalar que o sinal de igualdade é válido em (5. 3) nas mesmas condições em que aparece em (5. 2), isto é, quando x e y são linearmente dependentes.

A metrização de um espaço de HILBERT é imediata: basta que nele se introduza como norma de qualquer $x \in E$ o valor real

$$(5.\,4)\qquad \|x\| = \sqrt{(x\,,x)}\,;$$

esta função satisfaz, com efeito, às condições que fixamos no número 4 para a definição de uma norma: sendo $(x\,,x)=0$ quando e só quando $x=0$, ter-se-à $\|x\|=0$ nas mesmas condições, e tem lugar $n_1)$; por outro lado, em consequência de $e_1)$ vem $\|\alpha x\| = \sqrt{(\alpha x\,,\alpha x)} = \sqrt{|\alpha|^2\cdot(x\,,x)} = |\alpha|\cdot\|x\|$, e verifica-se $n_2)$; finalmente, em virtude de (5. 3) pode-se escrever $\|x+y\| = \sqrt{(x+y\,,x+y)} \leqq \sqrt{(x\,,x)}+\sqrt{(y\,,y)} = \|x\|+\|y\|$, e é respeitado o axioma $n_3)$. Deste modo, o espaço de HILBERT é um espaço vectorial normado e, como tal, metrizável. E atendendo à observação que fizemos a respeito do sinal de igualdade em (5. 3), segue-se que em $n_3)$ só será, neste caso, válido o sinal de igualdade desde que se tenha $x=\beta y$; ou seja: *no espaço de* HILBERT *a norma definida por* (5. 4) *é sempre forte.*

Obs. Alguns autores (por exemplo: N. I. ACHIESER e I. M. GLASSMANN, *Theorie der linearen Operatoren im* HILBERT-RAUM, Berlim, 1954) designam pelo nome de espaços de HILBERT os espaços vectoriais normados metrizáveis pela noção de produto escalar e *completos.*

Exemplos de espaços de HILBERT:

V) *O espaço* l^2 *das sucessões infinitas de números complexos*

$$\mathrm{x} = \{a_n\}_1^\infty\,, y = \{\mathrm{b}_n\}_1^\infty\,,\cdots$$

para os quais

$$\sum_1^\infty |\mathrm{a}_n|^2 < \infty\,,\quad \sum_1^\infty |\mathrm{b}_n|^2 < \infty\,,\cdots$$

é um espaço de HILBERT *completo e separável.* (Os números $a_n\,,b_n\,,\cdots(n=1\,,2\,,\cdots)$ serão designados por componentes dos vectores $x\,,y\,,\cdots)$.

Na verdade pode-se definir a adição de dois vectores $x\,,y \in l^2$ assim $x+y=\{a_n+b_n\}_1^\infty$, pois a série $\sum_1^\infty |a_n+b_n|^2$ é convergente em virtude da desigualdade

$$(1)\qquad |\alpha+\beta|^2 \leqq 2|\alpha|^2+2|\beta|^2\,(^1)$$

e as condições $v_1)$ verificam-se desde que se tome $0=\{0\,,0\,,\cdots 0\,,\cdots\}$. O produto de um vector $x \in l^2$ por um número α (real ou complexo) fica definido por $\alpha x = \{\alpha a_n\}_1^\infty$, sendo válidas as condições $v_2)$, como é manifesto. Finalmente, pode-se introduzir como produto escalar de quaisquer dois vectores $x\,,y \in l^2$ o número

$$(5.\,5)\qquad (x\,,y) = \sum_1^\infty a_n \overline{b}_n$$

pois a série que aparece no segundo membro é convergente em virtude de ser $|\alpha\cdot\beta| \leqq$

(1) Tem-se $|\alpha+\beta| \leqq |\alpha|+|\beta|$ donde

(2) $\qquad |\alpha+\beta|^2 \leqq |\alpha|^2+|\beta|^2+2\cdot|\alpha|\cdot|\beta|;$

mas para a e b reais $(a-b)^2 = a^2+b^2-2ab \geqq 0$, portanto $2ab \leqq a^2+b^2$ e $2|\alpha|\cdot|\beta| \leqq |\alpha|^2+|\beta|^2$ (3) o que, substituindo em (2) dá a desigualdade (1), como desejavamos.

$\overline{<}\frac{1}{2}|\alpha|^2+\frac{1}{2}|\beta|^2$ (desigualdade (3) da nota anterior) e as condições e_1), e_2) e e_3) são verificadas: as duas últimas são imediatas e a primeira resulta de:

$$\sum_1^\infty a_n\overline{b}_n = \lim_{r\to\infty}\sum_1^r a_n\overline{b}_n = \lim_{r\to\infty}\overline{\sum_1^r \overline{a}_n b_n} = \overline{\sum_1^\infty \overline{a}_n b_n}.$$

Em consequência de (5. 4) e (5. 5) a norma de qualquer vector $x \in \{a_n\}_1^\infty \in l^2$ é dada por:

$$(5.6)\qquad \|x\| = \sqrt{(x,x)} = \sqrt{\sum_1^\infty a_n\overline{a}_n} = \sqrt{\sum_1^\infty |a_n|^2}.$$

Assim, l^2 é um espaço de HILBERT. Para mostrarmos que ele é separável basta considerar o conjunto N dos vectores de l^2, distintos de 0, para os quais as componentes são números complexos racionais (isto é da forma $\alpha+i\beta$, com α e β racionais); como N é numerável e denso em l^2, este espaço é separável.

Falta provar que l^2, é completo. Seja $x^{(k)}=\{a_n^{(k)}\}_{n=1}^\infty (k=1,2,\cdots)$ uma sucessão de vectores de l^2 que verifica a condição de CAUCHY

$$(4)\qquad \|x^{(r)}-x^{(s)}\| = \sqrt{\sum_1^\infty |a_n^{(r)}-a_n^{(s)}|^2} < \varepsilon$$

para $r, s > k(\varepsilon)$; ora qualquer que seja o número inteiro e positivo m

$$|a_m^{(r)}-a_m^{(s)}| \overline{<} \sqrt{\sum_{n=1}^\infty |a_n^{(r)}-a_n^{(s)}|^2} < \varepsilon$$

para $r, s > k(\varepsilon)$; e assim, cada sucessão dos valores de qualquer componente, $\{a_m^{(k)}\}_{k=1}^\infty$, verifica também a condição de CAUCHY e, como tal, tem um limite a_m

$$(5)\qquad \lim_{k\to\infty} a_m^{(k)} = a_m \quad (m=1,2,\cdots).$$

Da desigualdade (4) tira-se, qualquer que seja o inteiro e positivo v e com $r, s > k(\varepsilon)$,

$$\sqrt{\sum_{n=1}^v |a_n^{(r)}-a_n^{(s)}|^2} < \varepsilon;$$

fazendo $s \to \infty$ e tendo em atenção (5) vem

$$\sqrt{\sum_{n=1}^v |a_n^{(r)}-a_n|^2} \overline{<} \varepsilon,$$

desigualdade que é válida qualquer que seja v, o que permite escrever

$$(6)\qquad \|x^{(r)}-x\| = \sqrt{\sum_{n=1}^\infty |a_n^{(r)}-a_n|^2} \overline{<} \varepsilon,$$

onde é $x=\{a_n\}_1^\infty$, o que prova ser $\lim_{r\to\infty} x^{(r)}=x$; mas como, qualquer que seja $r>k(\varepsilon)$, é $x^{(r)}-x=\{a_n^{(r)}-a_n\}\in l^2$ porque, em virtude de (6), se tem $\sum_{n=1}^\infty |a_n^{(r)}-a_n|^2 \overline{<} \varepsilon^2$, segue-se $x \in l^2$ e o espaço é completo (1).

VI) Seja $L^2[a,b]$ *o conjunto das funções* $x(t)$ *definidas no intervalo finito* $[a,b]$, *mensuráveis e cujo quadrado do valor absoluto é* L-*integrável.* $L^2[a,b]$ *é um espaço de* HILBERT *completo,* pois: Sendo $x(t), y(t) \in L^2[a,b]$, também $x(t)+y(t) \in L^2[a,b]$, visto que a soma de duas funções mensuráveis é uma função mensurável, e $|x(t)+y(t)|^2$ é L-integrável em $[a,b]$ por ser (2)

$$|x(t)+y(t)|^2 \overline{<} 2|x(t)|^2+2|y(t)|^2;$$

como vector nulo do espaço torna-se uma função que seja nula em *quase todo* o intervalo $[a,b]$.

Por outro lado, se $x(t)\in L^2[a,b]$ é $\alpha x(t) \in L^2[a,b]$, qualquer que seja o número α; e fica definido o produto de um vector por um número.

(1) Esta demonstração encontra-se, por exemplo, em N. I. ACHIESER e I. M. GLASSMANN, *loc. cit.*, pág. 7.

(2) Conf. a desigualdade (1) do exemplo anterior.

O produto escalar de dois vectores $x(t), y(t) \in L^2[a,b]$ é dado por

$$(5.7) \qquad (x,y) = \int_a^b x(t) \cdot y(t)\, dt$$

pois (desigualde (3) do exercício anterior)

$$|x(t) \cdot y(t)| \leqq \frac{1}{2} |x(t)|^2 + \frac{1}{2} |y(t)|^2$$

garante a existência do integral do segundo membro de

$$\left| \int_a^b x(t) \cdot y(t)\, dt \right| \leqq \int_a^b |x(t) \cdot y(t)|\, dt$$

e, portanto, a existência do integral que intervem em (5.7).

Conclui-se que $L^2[a,b]$ é um espaço de HILBERT, ficando a norma de qualquer $x(t) \in L^2[a,b]$ determinada por

$$\| x(t) \| = \sqrt{\int_a^b |x(t)|^2\, dt}$$

Falta-nos verificar que $L^2[a,b]$ é um espaço completo[1]. Seja $\{x_n(t)\}_1^\infty \in L^2[a,b]$ uma sucessão que satisfaz à condição de CAUCHY

$$\| x_n(t) - x_m(t) \| < \sqrt{\varepsilon} \quad \text{ou}$$
$$\int_a^b |x_n(t) - x_m(t)|^2\, dt < \varepsilon$$

para $n, m > N(\varepsilon)$; considere-se a sucessão crescente de números inteiros e positivos $k_1 < k_2 < \cdots < k_i < \cdots$ tal que

$$\int_a^b |x_{k_{i+1}}(t) - x_{k_i}(t)|^2\, dt < \frac{1}{8^i} \quad (i = 1, 2, \cdots);$$

o conjunto A_i dos pontos de $[a,b]$ para os quais

$$|x_{k_{i+1}}(t) - x_{k_i}(t)| \geqq 1/2^i$$

tem medida $m(A_i) < 1/2^i$, pois sendo $A_i \subset \subset [a,b]$ vem:

$$\frac{1}{8^i} > \int_a^b |x_{k_{i+1}}(t) - x_{k_i}(t)|^2\, dt \geqq$$
$$\geqq \int_{A_i} |x_{k_{i+1}}(t) - x_{k_i}(t)|^2\, dt \geqq \frac{1}{4^i} m(A_i)$$

donde se tira, como desejavamos

$$(1) \qquad m(A_i) < \frac{1}{2^i}.$$

As desigualdades

$$|x_{k_{s+1}}(t) - x_{k_s}(t)| < \frac{1}{2^s}$$
$$|x_{k_{s+2}}(t) - x_{k_{s+1}}(t)| < \frac{1}{2^{s+1}}$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$

são simultâneamente verificadas no intervalo $I_s \subset [a,b]$ dado por

$$(2) \qquad I_s = [a,b] - (A_s + A_{s+1} + \cdots)$$

em virtude da propriedade aditiva generalizada da medida-L e de (1):

$$m([a,b] - I_s) = m\left(\sum_{n=s}^{\infty} A_n\right) =$$
$$= \sum_{n=s}^{\infty} m(A_n) < \sum_{n=s}^{\infty} \frac{1}{2^n},$$

ou seja

$$(3) \qquad m([a,b] - I_s) < \frac{1}{2^{s-1}};$$

mas de (2) tira-se

$$I_s \subset I_{s+1} \subset \cdots \subset [a,b]$$

donde

$$\lim_{s \to \infty} I_s = I \subset [a,b]$$

sendo, em resultado de (3)

$$(4) \qquad m([a,b] - I) = \lim_{s \to \infty} m([a,b] - I_s) = 0.$$

[1] Seguimos N. I. ACHIESER e I. M. GLASSMANN, *loc. cit.*, pág. 23. O leitor encontra outra demonstração em I. P. NATANSON, *Theorie der Funktionen einer reeller Veranderlichen*, Berlim 1954, pág. 169.

Como para $n > m > s$ é

$$|x_{k_n}(t) - x_{k_m}(t)| \leqq \sum_{r=m}^{n-1} |x_{k_{r+1}}(t) - x_{k_r}(t)|$$

$$< \sum_{r=m}^{n-1} \frac{1}{2^r} < \frac{1}{2^{m-1}}$$

segue-se que a sucessão $\{x_{k_r}(t)\}_{r=1}^{\infty}$ converge uniformemente em cada um dos conjuntos I_s; então converge uniformemente em I, ou seja (conf. (4)) em quase todos os pontos de $[a, b]$.

Definamos então uma função $x(t)$ assim:

$$(5) \quad \begin{cases} x(t) = \lim_{r \to \infty} x_{k_r}(t) & \text{para } t \in I \\ x(t) = 0 & \text{para } t \in [a, b] - I \end{cases}$$

Ora a sucessão considerada verifica a condição de CAUCHY e por isso $(I_s \subset [a, b])$

$$\int_{I_s} |x_m(t) - x_{k_r}(t)|^2 \cdot dt \leqq$$

$$\leqq \int_a^b |x_m(t) - x_{k_r}(t)|^2 dt < \varepsilon$$

desde que $m, k_r > N(\varepsilon)$; mas dada a convergência uniforme de $\{x_{k_r}(t)\}_{r=1}^{\infty}$ em I_s obtem-se tomando o limite quando $r \to \infty$ e tendo em atenção (5):

$$\int_{I_s} |x_m(t) - x(t)|^2 dt < \varepsilon$$

qualquer que seja I_s; ou

$$\| x_m - x \| = \sqrt{\int_a^b |x_m(t) - x(t)|^2 \cdot dt} < \sqrt{\varepsilon}$$

Assim se conclui que $x_m(t) - x(t) \in L^2[a, b]$, portanto $x(t) \in L^2[a, b]$; e, ao mesmo tempo, que $\lim_{m \to \infty} x_m(t) = x(t)$, como desejavamos concluir [1].

[1] Esta demonstração ainda podia servir, com pequenas alterações, no caso em que $[a, b]$ è infinito (N. I. ACHIESER e I. M. GLASSMANN, *loc. cit.*, pág. 24).

PEDAGOGIA

Notas sobre o ensino da matemática em Portugal

*por **Hugo Ribeiro***

As seguintes notas foram escritas em Outubro de 1956 quando interrompemos por umas horas a nossa visita à família e amigos em Portugal para responder à questão [1], que nos foi verbalmente posta, das nossas impressões sobre a situação da Matemática em Portugal e as soluções imediatas a dar-lhe. Não se fundamenta em nenhum estudo sistemático do problema, mas antes em observações feitas em parte de longe e ligadas a reminiscências muito vivas e directas de tempos passados.

1. Já não há, com a mesma agudeza, o problema fundamental reconhecido por ANTÓNIO MONTEIRO e que o próprio, a J. I. M. e o I. A. C. tentaram resolver criando entre os jóvens estudiosos portugueses a confiança nas suas capacidades para a investigação matemática.

[1] Não foi a Redacção nem nenhum dos elementos que a constitui que entrevistou o Prof. HUGO RIBEIRO, simplesmente este, como dedicado Amigo da Gazeta ofereceu à consideração dos leitores o resultado dessa entrevista.