JOURNAL OF APPROXIMATION THEORY 13, 229–234 (1975)

Sur le Théorème de Superposition de Kolmogorov

JEAN-PIERRE KAHANE

Mathématiques, Centre d'Orsay, Université Paris Sud, 91-Orsay, France

Communicated by P. L. Butzer

EN HOMMAGE À G. G. LORENTZ

En 1957, Kolmogorov a démontré le théorème suivant [6]

Toute fonction continue réelle définie sur I^n ($I = [0, 1]$ et n entier $\geqslant 2$) est représentable sous la forme

$$f(x_1, x_2, \ldots, x_n) = \sum_{q=1}^{2n+1} g_q \Big(\sum_{p=1}^{n} \varphi_{pq}(x_p) \Big), \tag{1}$$

où les g_p et les φ_{pq} sont des fonctions continues d'une variable réelle; de plus, les φ_{pq} sont croissantes sur I, et ne dépendent pas de f.

Ce résultat a été précisé dans plusieurs directions:

1. On peut choisir les φ_{pq} dans une classe Lip α $(0 < \alpha < 1)$ [8, Chapter 11] et même dans la classe Lip 1 (Fridman [2]; voir aussi [9]); mais on ne peut pas les construire dans la classe C^1 [4, 10].
2. On peut choisir les $\varphi_{pq}(x)$ de la forme $\lambda_p \varphi_q(x)$ (Sprecher; voir [8, Chapter 11; 9]).
3. On peut choisir les g_q toutes égales ([8, Chapter 11]).

Il est d'autre part intéressant, même dans le cas $n = 1$, de préciser dans quelles classes de fonctions on peut choisir les g_q (par exemple: transformées de Fourier de fonctions sommables, valeurs au bord de fonctions analytiques, etc.). En effet, l'interprétation géométrique du théorème de Kolmogorov est que toute fonction continue sur l'ensemble $E \subset I^{2n+1}$ défini paramétriquement par

$$X_q = \sum_{p=1}^{n} \varphi_{pq}(x_p) \qquad ((x_1, \ldots, x_n) \in I^n, \qquad q = 1, 2, \ldots, 2n + 1) \tag{2}$$

peut être prolongée sur I^{2n+1} en une fonction de la forme

$$g_1(X_1) + g_2(X_2) + \cdots + g_{2n+1}(X_{2n+1}) \tag{3}$$

et—sous la forme plus précise de Lorentz—en une fonction de la forme

$$g(X_1) + g(X_2) + \cdots + g(X_{2n+1}). \tag{4}$$

Ainsi E apparaît comme un "ensemble d'interpolation" dans le sens suivant: toute fonction continue sur E peut être prolongée sur I^{2n+1} en une fonction appartenant à une classe convenable [1; 3; 5; 7, Chapter 4].

Le présent article se propose quelques variantes et commentaires sur ce thème.

Remarquons d'abord que le résultat de Fridman est une conséquence immédiate du théorème de Kolmogorov. En effet, l'ensemble E défini par (2) est la somme algébrique de n courbes Γ_p d'équations

$$X_q = \varphi_{pq}(x) \qquad (x \in I, q = 1, 2,..., 2n + 1).$$

Chacune de ces courbes est rectifiable et peut être paramétrée au moyen de la longueur de l'arc normalisée (de façon que la longueur totale de Γ_p soit 1). Ainsi l'équation de Γ_p prend la forme

$$X_q = \psi_{pq}(s) \qquad (s \in I, q = 1, 2,..., 2n + 1)$$

où chaque fonction ψ_{pq} est croissante et *lipschitzienne*, et l'ensemble E est maintenant défini par

$$X_q = \sum_{p=1}^{n} \psi_{pq}(s_p) \qquad ((s_1 ,..., s_n) \in I^n, \qquad q = 1, 2,..., 2n + 1).$$

Le théorème de Kolmogorov signifie que toute fonction F continue sur I^n s'écrit

$$F(s_1 , s_2 ,..., s_n) = \sum_{q=1}^{2n+1} g_q \left(\sum_{p=1}^{n} \psi_{pq}(s_q)\right). \qquad \text{C.Q.F.D.}$$

Donnons maintenant une démonstration du théorème de Kolmogorov à l'aide de la théorie de Baire; cela facilite les variantes que nous donnerons ensuite.

Si A est un espace métrique complet, nous dirons qu'une propriété a lieu pour *quasi tout* $a \in A$ si elle a lieu sur une intersection dénombrable d'ouverts denses dans A. Il est facile de vérifier que si A et B sont deux espaces métriques complets ayant chacun une base dénombrable d'ouverts, et si une propriété a lieu pour quasi tout $(a, b) \in A \times B$, alors, pour quasi tout a, elle est vraie pour quasi tout b.

Soit Φ l'ensemble des applications φ croissantes et continues de I dans I, telles que $\varphi(0) = 0$ et $\varphi(1) = 1$. La distance usuelle fait de Φ un espace métrique complet. Quasi toute φ est strictement croissante, car visiblement, pour tout couple de rationnels (ρ, ρ') tels que $0 \leqslant \rho < \rho' \leqslant 1$, l'ensemble des φ qui vérifient $\varphi(\rho + 0) < \varphi(\rho' - 0)$ $(= \varphi(\rho'))$ est un ouvert dense dans Φ.

Soit maintenant $\lambda_1, ..., \lambda_n$ des nombres tous distincts, strictement positifs, de somme 1. Nous allons montrer ceci.

Pour quasi tout $(\varphi_1, \varphi_2, ..., \varphi_{2n+1}) \in \Phi^{2n+1}$, *toute fonction* f_0 *continue sur* I^n *est représentable sous la forme*

$$f_0(x_1, ..., x_n) = \sum_{q=1}^{2n+1} g\left(\sum_{p=1}^{n} \lambda_p \varphi_q(x_p)\right) \tag{5}$$

où g *est une fonction continue sur* I.

Pour le moment, on peut penser à f_0 et à g comme à des fonctions à valeurs réelles ou complexes.

Démonstration. Soit $\epsilon > 0$ un nombre assez petit, qu'on fixera plus tard. Pour chaque $f \in C(I^n)$, $f \not\equiv 0$, considérons l'ensemble $\Omega(f)$ des $(\varphi_1, ..., \varphi_{2n+1}) \in \Phi^{2n+1}$ tels qu'il existe $h \in C(I)$ avec

$$\begin{aligned} &\text{(i)} \quad \|h\| \leqslant \|f\|, \\ &\text{(ii)} \quad \left\| f(x_1, ..., x_n) - \sum_{q=1}^{2n+1} h\left(\sum_{p=1}^{n} \lambda_p \varphi_q(x_p)\right) \right\| < (1 - \epsilon)\|f\|, \end{aligned} \tag{6}$$

(les normes étant prises dans $C(I)$ et dans $C(I^n)$). Visiblement $\Omega(f)$ est un ouvert dans Φ^{2n+1}. Nous allons montrer qu'il y est dense; admettons-le pour le moment. Soit F un ensemble dénombrable dense dans $C(I^n)$, ne contenant pas la fonction nulle. Supposons enfin $(\varphi_1, ..., \varphi_{2n+1}) \in \bigcap_{f \in F} \Omega(f)$ (ce sera la signification de "pour quasi tout $(\varphi_1, ..., \varphi_{2n+1})$"). Alors, pour tout $f_0 \in C(I^n)$, $f_0 \neq 0$, il existe une fonction $f \in F$ telle que $\|f^3\| \leqslant \|f_0\|$ et $\|f - f_0\| < (\epsilon/2)\|f_0\|$, et une fonction h telle que (6) ait lieu; soit $h = \gamma(f_0)$. Posons $\gamma(0) = 0$. Définissons pour $j = 0, 1, ...,$ par induction, $h_j = \gamma(f_j)$, et

$$f_{j+1}(x_1, ..., x_n) = f_j(x_1, ..., x_n) - \sum_{q=1}^{2n+1} h_j\left(\sum_{p=1}^{n} \lambda_p \varphi_q(x_p)\right).$$

D'après (6) la série $\sum_{j=0}^{\infty} h_j$ converge dans $C(I)$, et sa somme g vérifie (5).

Reste à démontrer que $\Omega(f)$ est dense dans Φ^{2n+1}. Soit G un ouvert dans Φ^{2n+1}, et $\delta = \delta(G, \epsilon, f) > 0$ (à définir plus tard). Notons

$$I_q = I_q(j) = [q\delta + (2n + 1)j\delta, q\delta + (2n + 1)j\delta + 2n\delta]$$

($q = 1, 2, \ldots, 2n + 1$; $j \in \mathbf{Z}$). Remarquons que 1°) pour q fixé, les intervalles $I_q(j)$ sont disjoints, et séparés par des intervalles de longueur δ 2°) tout point de $I = [0, 1]$ appartient à un I_q pour chaque valeur de q sauf au plus une. Désignons par P_q tout cube de la forme

$$P_q = P_q(j_1, j_2, \ldots, j_n) = I_q(j_1) \times I_q(j_2) \times \cdots \times I_q(j_n)$$

et remarquons que tout point de I^n appartient à un P_q pour chaque valeur de q sauf au plus n (c'est-à-dire pour $n + 1$ valeurs de q au moins). Soit Δ l'ensemble des $(\varphi_1, \ldots, \varphi_{2n+1}) \in \Phi^{2n+1}$ telles que chaque φ_q soit constante sur chaque intervalle I_q, et linéaire entre deux intervalles consécutifs $I_q(j)$ et $I_q(j + 1)$. On choisit $\delta = \delta(G, \epsilon, f)$ de facon que 1°) l'oscillation de f sur chaque P_q ne dépasse pas $\epsilon \|f\|$ 2°) $G \cap \Delta \neq \varnothing$.

Soit $(\varphi_1, \ldots, \varphi_{2n+1}) \in G \cap \Delta$. Quitte à modifier de très peu la valeur des φ_q (de façon à rester dans $G \cap \Delta$), on peut supposer que, pour chaque q, la fonction

$$\chi_q(x_1, \ldots, x_n) = \sum_{p=1}^{n} \lambda_p \varphi_q(x_p)$$

(qui est constante sur chaque P_q) prend des valeurs différentes sur les différents P_q, et aussi que les $\chi_q(P_q)$ sont distincts pour des q différents (on utilise ici l'hypothèse que les λ_j sont distincts); en d'autres termes, l'application $(q, j_1, \ldots, j_n) \to \chi_q(P_q(j_1, \ldots, j_n))$ est injective.

Désignons par $\mathscr{M}(P_q)$ la valeur moyenne de f sur P_q, posons, pour tous les pavés $P_q(j_1, j_2, \ldots, j_n)$

$$h(\chi_q(P_q)) = 2\epsilon \mathscr{M}(P_q)$$

et prolongeons h sur I de manière à avoir $\|h\| \leqslant 2\epsilon \|f\|$ (à cela près, quelconque). Soit $x = (x_1, \ldots, x_n) \in I^n$. Si $x \in P_q$, on a

$$h(\chi_q(x)) = 2\epsilon f(x) + \rho, \qquad |\rho| \leqslant 2\epsilon^2 \|f\|.$$

Comme x appartient à au moins $n + 1$ cubes P_q, on a, pour $\epsilon < 1/2(n + 1)$,

$$\begin{aligned}\left| f(x) - \sum_{1}^{2n+1} h(\chi_q(x)) \right| &\leqslant (1 - 2(n + 1)\epsilon) |f(x)| + 2(n + 1)\epsilon^2 \|f\| + 2n\epsilon \|f\| \\ &\leqslant (1 - 2\epsilon + 2(n + 1)\epsilon^2) \|f\| \\ &\leqslant (1 - \epsilon) \|f\|.\end{aligned}$$

On a donc (6), ce qui achève la démonstration.

La démonstration ci-dessus vaut encore si f_0, au lieu d'être à valeurs réelles ou complexes, a ses valeurs dans un espace de Banach; il suffit de remplacer les valeurs absolues par des normes. Et le résultat vaut encore si f_0 prend ses valeurs dans un espace de Fréchet; en effet, le fait que pour chaque semi-norme $\|\ \|_{[\nu]}$ on ait quasi-sûrement $\| f_0(x_1, ..., x_n) - \sum_{q=1}^{2n+1} g(\sum_{p=1}^{n} \lambda_p \varphi_q(x_p)) \|_{[\nu]} = 0$ pour tout $(x_1, ..., x_n) \in I^n$ entraîne qu'on a (5) quasi-sûrement.

L'énoncé vaut donc si f_0 et g prennent leurs valeurs dans un espace de Fréchet.

Revenons au cas où f_0 est réelle ou complexe, g réelle ou complexe. Désignons par $B(I)$ un espace de Banach contenu dans $C(I)$ et ayant la la propriété suivante: il existe un $c > 0$ tel que, quels que soient N, les points distincts $t_1, t_2, ..., t_N$ de I, et les nombres $u_1, u_2, ..., u_N$ de module $\leqslant 1$, il existe $g \in B(I)$ telle que $\| g \|_{C[I]} \leqslant 1$, $\| g \|_{B[I]} \leqslant c$, et $g(t_j) = u_j$ $(j = 1, 2, ..., N)$

On peut alors, dans l'énoncé du théorème, prendre $g \in B(I)$.

Comme application, si nous identifions I et le cercle $\mathbf{T} = \mathbf{R}/\mathbf{Z}$, l'ensemble $E \subset \mathbf{T}^{2n+1}$ défini par

$$\chi_q = \sum_{p=1}^{n} \lambda_p \varphi_q(x_p) \qquad ((x_1, ..., x_n) \in I^n, \qquad q = 1, 2, ..., 2n + 1), \quad (7)$$

est quasi-sûrement un ensemble d'interpolation dans le sens suivant: toute fonction continue sur E s'écrit sous la forme (4), g étant valeur au bord d'une fonction analytique (c'est-à-dire $g(t) \sim \sum_0^\infty \hat{g}(n) \exp(2\pi int)$).

Identifions toujours I et $\mathbf{T}$. Il est connu qu'aucun ensemble E de la forme (2) n'est d'interpolation pour $A(\mathbf{T}^{2n+1})$ (classe des fonctions $g(t_1, t_2, ..., t_{2n+1})$ dont la série de Fourier est absolument convergente); cela tient au fait que E contient la somme algébrique de deux ensembles infinis. On ne peut donc pas, dans l'énoncé du théorème, prendre $g \in A(\mathbf{T})$. Mais, quitte à déformer convenablement l'ensemble E donné par (7), il devient un ensemble d'interpolation pour $A(\mathbf{T}^{2n+1})$; cette observation est due à R. Doss. Contentons-nous d'énoncer un résultat facile à démontrer en adaptant la méthode de Baire.

Soit $\{m_k\}$ une suite strictement croissante d'entiers positifs (par exemple $m_k = (((k!)!)!)$. Pour quasi tout $(\varphi_1, ..., \varphi_{2n+1}) \in \Phi^{2n+1}$ et quasi tout homéomorphisme ψ de classe C^∞ de $\mathbf{T}$ sur $\mathbf{T}$, toute fonction continue sur l'ensemble $E' \subset \mathbf{T}^{2n+1}$ d'équation

$$t_q = \psi\left(\sum_{p=1}^{n} \lambda_p \varphi_q(x_p)\right) \qquad ((x_1, ..., x_n) \in \mathbf{T}^n, \qquad q = 1, 2, ..., 2n + 1)$$

s'écrit sous la forme

$$g(t_1) + g(t_2) + \cdots + g(t_{2n+1})$$

avec $g \in A(\mathbf{T})$ et $g(t) = \sum_1^\infty \hat{g}(m_k) \exp(2\pi\, \mathrm{i} m_k t)$.

Bibliographie

1. R. Doss, Représentation des fonctions continues de deux variables réelles, *C. R. Acad. Sci., Paris* **277** (1973), 1169–1170. Voir aussi "Representation of functions of several variables," A paraître dans *Amer. J. Math.*
2. B. L. Fridman, Improvement in the smoothness of functions in the Kolmogorof superposition theorem, *Soviet Math. Dokl.* **8** (1967), 1550–1553; *Dokl. Akad. Nauk SSSR* **177** (1967), 5.
3. T. Hedberg, Sur les réarrangements de fonctions de la classe A et les ensembles d'interpolation pour $A(D^2)$, *C. R. Acad. Sci. Paris* **270** (1970), 1491–1494.
4. G. M. Henkin, Linear superpositions of continuously differentiable functions, (en russe) *Dokl. Akad. Nauk. SSSR* **157** (1964), 288–290.
5. J.-P. Kahane, Sur les réarrangements de fonctions de la classe A, *Studia Math.* **31** (1968), 287–293.
6. A. N. Kolmogorov, On the representation of continuous functions of several variables by superpositions of continuous functions of one variable and addition, (en russe) *Dokl. Akad. Nauk SSSR* **114** (1957), 679–681; *Amer. Math. Soc. Transl.* **28** (1963), 55–59.
7. L.-Å. Lindahl and F. Poulsen, "Thin Sets in Harmonic Analysis," Dekker, New York, 1971.
8. G. G. Lorentz, "Approximation of Functions," Holt, Reinhart and Winston, New York, 1966.
9. D. Sprecher, An improvement in the superposition theorem of Kolmogorov, *J. Math. Anal. Appl.* **38** (1972), 208–213.
10. A. G. Vituškin, Proof of existence of analytic functions of several variables, not representable by linear superpositions, (en russe) *Dokl. Akad. Nauk SSSR* **156** (1964), 1258–1261.