ISSN 1064-5624, Doklady Mathematics, 2011, Vol. 83, No. 3, pp. 1–3. © Pleiades Publishing, Ltd., 2011.
Original Russian Text © G.G. Magaril-Il'yaev, K.Yu. Osipenko, 2011, published in Doklady Akademii Nauk, 2011, Vol. 438, No. 3, pp. 300–302.

MATHEMATICS

Hardy–Littlewood–Paley Inequality and Recovery of Derivatives from Inaccurate Data

G. G. Magaril-Il'yaev[a, b] **and K. Yu. Osipenko**[b, c]

Presented by Academician V.A. Il'in December 29, 2010

Received January 26, 2011

DOI: 10.1134/S1064562411030203

The Hardy–Littlewood–Paley inequality [1] is the sharp inequality

$$\|x^{(k)}(\cdot)\|_{L_2(\mathbb{R})} \le \|x(\cdot)\|_{L_2(\mathbb{R})}^{1-k/n}\|x^{(n)}(\cdot)\|_{L_2(\mathbb{R})}^{k/n}, \tag{1}$$

which holds for all functions $x(\cdot)$ from the Sobolev space $\mathcal{W}_2^n(\mathbb{R}) = \{x(\cdot) \in L_2(\mathbb{R}) | \, x^{(n-1)}(\cdot)$ is locally absolutely continuous, $x^{(n)}(\cdot) \in L_2(\mathbb{R})\}$, where k and n are positive integers such that $k < n$.

The sharpness of inequality (1) means that that, for any $\delta_1 > 0$ and $\delta_2 > 0$, the value of the problem (i.e., the supremum of the maximized functional)

$$\|x^{(k)}(\cdot)\|_{L_2(\mathbb{R})} \to \max, \quad \|x(\cdot)\|_{L_2(\mathbb{R})} \le \delta_1, \quad \|x^{(n)}(\cdot)\|_{L_2(\mathbb{R})} \le \delta_2 \tag{2}$$

is $\delta_1^{1-k/n}\delta_2^{k/n}$.

Problem (2) is closely related to the optimal recovery of the kth derivative of a function $x(\cdot) \in \mathcal{W}_2^n(\mathbb{R})$ from inaccurate data on the function itself and its nth derivative. Assume that we are given (observe) functions $y_1(\cdot) \in L_2(\mathbb{R})$ and $y_2(\cdot) \in L_2(\mathbb{R})$ such that $\|x(\cdot) - y_1(\cdot)\|_{L_2(\mathbb{R})} \le \delta_1$ and $\|x^{(n)}(\cdot) - y_2(\cdot)\|_{L_2(\mathbb{R})} \le \delta_2$. The optimal recovery problem is to find the error of optimal recovery

$$E(D^k, \mathcal{W}_2^n(\mathbb{R}), \delta_1, \delta_2) = \inf_m \sup_{\substack{x(\cdot) \in \mathcal{W}_2^n(\mathbb{R}), \\ y_i(\cdot) \in L_2(\mathbb{R}),\, i = 1, 2 \\ \|x(\cdot) - y_1(\cdot)\|_{L_2(\mathbb{R})} \le \delta_1, \\ \|x^{(n)}(\cdot) - y_2(\cdot)\|_{L_2(\mathbb{R})} \le \delta_2,}} \|x^{(k)}(\cdot) - m(y_1(\cdot), y_2(\cdot))(\cdot)\|_{L_2(\mathbb{R})}, \tag{3}$$

where the infimum is taken over all mappings m: $L_2(\mathbb{R}) \times L_2(\mathbb{R}) \to L_2(\mathbb{R})$, and to find those m for which the infimum is attained, which are called optimal recovery methods.

It turns out that the optimal recovery error is equal to the value of problem (2) and there are whole families of optimal methods for recovering the kth derivative. More precisely, the following result holds.

Theorem. *Let k and n be positive integers such that $k < n$, an let $\delta_1 > 0$, $\delta_2 > 0$, and*

$$\lambda_1 = \frac{n-k}{n}\left(\frac{\delta_1}{\delta_2}\right)^{-2k/n}, \quad \lambda_2 = \frac{k}{n}\left(\frac{\delta_1}{\delta_2}\right)^{2(n-k)/n}.$$

Then

$$E(D^k, \mathcal{W}_2^n(\mathbb{R}), \delta_1, \delta_2) = \delta_1^{1-k/n}\delta_2^{k/n},$$

and, if $a(\cdot)$ is a function on $\mathbb{R}$ such that

$$\left|a(\xi) - \frac{\lambda_1 (i\xi)^k}{\lambda_1 + \lambda_2\xi^{2n}}\right| \le \frac{\sqrt{\lambda_1\lambda_2}|\xi|^n}{\lambda_1 + \lambda_2\xi^{2n}}\sqrt{-\xi^{2k} + \lambda_1 + \lambda_2\xi^{2n}}, \tag{4}$$

for a.e. $\xi \in$, then the method m_a: $L_2(\mathbb{R}) \times L_2(\mathbb{R}) \to L_2(\mathbb{R})$ defined by the formula

$$m_a(y_1(\cdot), y_2(\cdot))(\cdot) = \Lambda_1 y_1(\cdot) + \Lambda_2 y_2(\cdot), \tag{5}$$

is optimal; here, Λ_i: $L_2(\mathbb{R}) \times L_2(\mathbb{R})$, $i = 1, 2$, are linear continuous operators whose Fourier transforms are defined as $F\Lambda_1 y_1(\xi) = a(\xi)Fy_1(\xi)$ and $F\Lambda_2 y_2(\xi) = (i\xi)^{-n}((i\xi)^k - a(\xi))Fy_2(\xi)$.

The operators Λ_1 and Λ_2 are convolution operators whose kernels are, generally speaking, distributions (generalized functions). Consider a two-parameter family of such operators such that their kernels are usual functions with a fairly simple description.

[a] *Moscow State Institute of Radio Engineering, Electronics, and Automation (Technical University), pr. Vernadskogo 78, Moscow, 117454 Russia*

[b] *Southern Mathematical Institute, Vladikavkaz Research Center, Russian Academy of Sciences, ul. Markusa 22, Vladikavkaz, 362027 Republic of North Ossetia-Alania, Russia*

[c] *Moscow State Aviation Technological University, ul. Orshanskaya 3, Moscow, 121552 Russia*
e-mail: magaril@mirea.ru, kosipenko@yahoo.com

Corollary. *Let k and n be positive integers such that $k < n$, and let $\delta_1 > 0$, $\delta_2 > 0$, and*

$$\hat{\sigma}_1 = \left(1 - \frac{k}{n}\right)^{1/(2k)}\left(\frac{\delta_1}{\delta_2}\right)^{-1/n},$$

$$\hat{\sigma}_2 = \left(\frac{n}{k}\right)^{1/(2(n-k))}\left(\frac{\delta_1}{\delta_2}\right)^{-1/n}.$$

Then, for any pair $(\sigma_1, \sigma_2) \in [0, \hat{\sigma}_1] \times [\hat{\sigma}_2, \infty]$, the method

$$m_{\sigma_1,\sigma_2}(y_1(\cdot), y_2(\cdot)) = (K^1_{\sigma_1,\sigma_2} * y_1)(\cdot) + (K^2_{\sigma_1,\sigma_2} * y_2)(\cdot)$$

is optimal; here, the kernels $K^1_{\sigma_1,\sigma_2}(\cdot)$ and $K^2_{\sigma_1,\sigma_2}(\cdot)$ are such that their Fourier transforms have the form

$$FK^1_{\sigma_1,\sigma_2}(\xi) = \begin{cases} (i\xi)^k, & |\xi| < \sigma_1, \\ (i\xi)^k\left(1 + \dfrac{k}{n-k}\left(\dfrac{\delta_1}{\delta_2}\right)^2\xi^{2n}\right)^{-1}, & \sigma_1 \le |\xi| < \sigma_2, \end{cases}$$

and $FK^1_{\sigma_1,\sigma_2}(\xi) = 0$ for $|\xi| \ge \sigma_2$ if $\sigma_2 < \infty$;

$$FK^2_{\sigma_1,\sigma_2}(\xi) = \begin{cases} 0, & |\xi| < \sigma_1, \\ (i\xi)^{k-n}\left(1 + \dfrac{n-k}{k}\left(\dfrac{\delta_1}{\delta_2}\right)^{-2}\xi^{-2n}\right)^{-1}, & \sigma_1 \le |\xi| < \sigma_2, \end{cases}$$

and $FK^2_{\sigma_1,\sigma_2}(\xi) = (i\xi)^{k-n}$ for $|\xi| \ge \sigma_2$ if $\sigma_2 < \infty$.

Note that an optimal method for recovering the kth derivative is linear and means that the "observations" $y_1(\cdot)$ and $y_2(\cdot)$ are "smoothed" (i.e., contracted with the kernels $K^1_{\sigma_1,\sigma_2}(\cdot)$ and $K^2_{\sigma_1,\sigma_2}(\cdot)$, respectively, and are then added.

Similar in spirit, the problem of optimally recovering a function and its derivatives from an inaccurate spectrum of the function was considered in [2, 3].

Proof of the theorem. First, we show that the optimal recovery error is no smaller than the value of problem (2). Indeed, let $x(\cdot)$ be an admissible function in (2). Then, obviously, the function $-x(\cdot)$ is also admissible and, for any m: $L_2(\mathbb{R}) \times L_2(\mathbb{R}) \to L_2(\mathbb{R})$, we have

$$2\|x^{(k)}(\cdot)\|_{L_2(\mathbb{R})} \le \|x^{(k)}(\cdot) - m(0)(\cdot)\|_{L_2(\mathbb{R})} + \|-x^{(k)}(\cdot) - m(0)(\cdot)\|_{L_2(\mathbb{R})}$$

$$\le 2 \sup_{\|x(\cdot)\|_{L_2(\mathbb{R})} \le \delta,\ \|x^{(n)}(\cdot)\|_{L_2(\mathbb{R})} \le 1} \|x^{(k)}(\cdot) - m(0)(\cdot)\|_{L_2(\mathbb{R})}$$

$$\le 2 \sup_{\substack{y_i(\cdot) \in L_2(\mathbb{R}) \\ \|x(\cdot) - y_i(\cdot)\|_{L_2(\mathbb{R})} \le \delta_i \\ i = 1, 2,\ \|x^{(n)}(\cdot)\|_{L_2(\mathbb{R})} \le 1}} \|x^{(k)}(\cdot) - m(y_1(\cdot), y_2(\cdot))(\cdot)\|_{L_2(\mathbb{R})}.$$

Passing to the supremum over all admissible functions in (2) on the left-hand side and to the infimum over all methods m on the right-hand side, we obtain the required result.

The since the recovered operator (kth derivative) is translation-invariant, it is natural to search for optimal methods among such operators. In terms of Fourier transforms, the action of a translation-invariant operator from $L_2(\mathbb{R})$ to $L_2(\mathbb{R})$ is multiplication by a function from $L_\infty(\mathbb{R})$ (see [4]). Therefore, optimal methods are sought in the form

$$m(y_1(\cdot), y_2(\cdot))(\cdot) = \Lambda_1 y_1(\cdot) + \Lambda_2 y_2(\cdot), \tag{6}$$

where $F\Lambda_i y_i(\cdot) = a_i(\cdot)Fy_i(\cdot)$ and $a_i(\cdot) \in L_\infty(\mathbb{R})$ for $i = 1, 2$.

It follows from (3) that the optimality of a method m means that the value of the problem

$$\begin{gathered} \|x^{(k)}(\cdot) - \Lambda_1 y_1(\cdot) - \Lambda_2 y_2(\cdot)\|_{L_2(\mathbb{R})} \to \max, \\ \|x(\cdot) - y_1(\cdot)\|_{L_2(\mathbb{R})} \le \delta_1, \\ \|x^{(n)}(\cdot) - y_2(\cdot)\|_{L_2(\mathbb{R})} \le \delta_2,\ y_i(\cdot) \in L_2(\mathbb{R}),\ i = 1, 2, \\ x(\cdot) \in \mathcal{W}^n_2(\mathbb{R}), \end{gathered} \tag{7}$$

is $E(D^k, \mathcal{W}^n_2(\mathbb{R}), \delta_1, \delta_2)$.

Method (6) must be sharp for functions from $\mathcal{W}^n_2(\mathbb{R})$; i.e., the following identity must hold:

$$x^{(k)}(\cdot) = \Lambda_1 x(\cdot) + \Lambda_1 x^{(n)}(\cdot) \quad \forall x(\cdot) \in \mathcal{W}^n_2(\mathbb{R}), \tag{8}$$

Otherwise, the value of problem (7) is equal to infinity. Indeed, if (8) does not hold for some $x_0(\cdot) \in \mathcal{W}^n_2(\mathbb{R})$, then the maximized functional can be made arbitrarily large by setting $x(\cdot) = Cx_0(\cdot)$, $y_1(\cdot) = Cx_0(\cdot)$, and $y_2(\cdot) = Cx_0^{(n)}(\cdot)$ in (7) and choosing $C \in \mathbb{R}$.

By applying the Plancherel theorem, taking into account identity (8) (in terms of Fourier transforms), and introducing $z_1(\xi) = Fx(\xi) - Fy_1(\xi)$ and $z_2(\xi) = (i\xi)^n Fx(\xi) - Fy_2(\xi)$, it is easy to see that the squared value of problem (7) is equal to the value of the problem

$$\begin{gathered} \frac{1}{2\pi}\int_{\mathbb{R}} |a_1(\xi)z_1(\xi) + a_2(\xi)z_2(\xi)|^2 d\xi \to \max, \\ \frac{1}{2\pi}\int_{\mathbb{R}} |z_1(\xi)|^2 d\xi \le \delta_1^2, \quad \frac{1}{2\pi}\int_{\mathbb{R}} |z_2(\xi)|^2 d\xi \le \delta_2^2. \end{gathered} \tag{9}$$

Let us estimate the maximized functional from above. For any $\lambda_i > 0$, $i = 1, 2$, and any $\xi \in \mathbb{R}$, we have the Cauchy–Schwarz inequality

$$|a_1(\xi)z_1(\xi) + a_2(\xi)z_2(\xi)|^2 \le \left(\frac{|a_1(\xi)|^2}{\lambda_1} + \frac{|a_2(\xi)|^2}{\lambda_2}\right)(\lambda_1|z_1(\xi)|^2 + \lambda_2|z_2(\xi)|^2).$$

Integrating it (with $S_{a_1,a_2}(\cdot)$ denoting the first (parenthesized) multiplier on the right-hand side) and taking into account the constraints in problem (9), we find that the value of the latter does not exceed

$$\|S_{a_1,a_2}(\cdot)\|_{L_\infty(\mathbb{R})}(\lambda_1\delta_1^2+\lambda_2\delta_2^2).$$

Choosing $a_i(\cdot) \in L_\infty(\mathbb{R})$ and $\lambda_i > 0$, $i = 1, 2$, such that $\|S_{a_1,a_2}(\cdot)\|_{L_\infty(\mathbb{R})} \le 1$ and

$$\lambda_1\delta_1^2+\lambda_2\delta_2^2 = \delta_1^{2(1-k/n)}\delta_2^{2k/n}, \qquad (10)$$

we conclude (in view of the lower bound obtained for $E(D^k, \mathcal{W}_2^n(\mathbb{R}), \delta_1, \delta_2))$ that the corresponding method is optimal and

$$E(D^k, \mathcal{W}_2^n(\mathbb{R}), \delta_1, \delta_2) = \delta_1^{1-k/n}\delta_2^{k/n}.$$

By virtue of (8), passing to Fourier transforms, we see that $a_1(\cdot)$ and $a_2(\cdot)$ are related by the formula $(i\xi)^k = a_1(\xi) + (i\xi)^n a_2(\xi)$ for a.e. $\xi \in \mathbb{R}$. Replacing $a_2(\cdot)$ in $S_{a_1,a_2}(\cdot)$ by its expression in terms of $a_1(\cdot)$ yields

$$S_{a_1,a_2}(\xi) = \frac{|a_1(\xi)|^2}{\lambda_1} + \frac{|(i\xi)^k - a_1(\xi)|^2}{\lambda_2\xi^{2n}}.$$

Then it is easy to see that the condition $\|S_{a_1,a_2}(\cdot)\|_{L_\infty(\mathbb{R})} \le 1$ can be written as inequality (4) (with $a_1(\cdot)$ replaced by $a(\cdot)$). The numbers $\lambda_1 > 0$ and $\lambda_2 > 0$ must be such that (in addition (10)) the expression under the root on the right-hand side of (4) is nonnegative on $\mathbb{R}$. If λ_2 is expressed as a function of λ_1 from (10) and the result is substituted under the root in (4), then the nonnegativity of the expression under the root means that

$$-\xi^{2k} + \lambda_2\left(\xi^{2n} - \left(\frac{\delta_1}{\delta_2}\right)^{-2}\right) + \left(\frac{\delta_1}{\delta_2}\right)^{-2k/n} \ge 0. \qquad (11)$$

for allξ $\in$. Obviously, the function on the left-hand side vanishes at the point $\left(\frac{\delta_1}{\delta_2}\right)^{-1/n}$. For inequality (11) to hold, this point must be a minimizer of the function. From this condition, we easily find λ_2. An expression for λ_1 follows from (10). Thus, with these λ_1 and λ_2, the expression under the root in (4) is nonnegative on $\mathbb{R}$.

It follows from (4) that the functions $\xi \mapsto a(\xi)$ and $\xi \mapsto (i\xi)^{-n}((i\xi)^k - a(\xi))$ belong to $L_\infty(\mathbb{R})$; thus, Λ_1 and Λ_2 are linear continuous operators.

Proof of the corollary. It follows directly from the theorem that the function

$$a_1(\xi) = \frac{\lambda_1(i\xi)^k}{\lambda_1+\lambda_2\xi^{2n}}$$

defines an optimal method. Setting $a_2(\xi) = (i\xi)^{-n}((i\xi)^k - a_1(\xi))$ and substituting the expressions for λ_i ($i = 1, 2$) into those for $a_1(\cdot)$ and $a_2(\cdot)$, we find that $a_i(\cdot) = FK^i_{0,\infty}(\cdot)$, $i = 1, 2$. It is easy to see that $a_i(\cdot) \in L_2(\mathbb{R})$, $i = 1, 2$. Therefore, method (5) can be written as convolutions with the kernels $K^1_{0,\infty}(\cdot)$ and $K^2_{0,\infty}(\cdot)$. This proves the corollary for the case of $\sigma_1 = 0$, $\sigma_2 = \infty$.

Let $\sigma_1, \sigma_2) \in (0, \hat{\sigma}_1] \times [\hat{\sigma}_2, \infty)$. We show that the functions $a_i(\cdot) = FK^i_{\sigma_1,\sigma_2}(\cdot)$ with the same λ_i ($i = 1, 2$) also define an optimal method. Indeed, $a_1(\xi) = (i\xi)^k$ on the interval $|\xi| < \sigma_1$. Hence, $S_{a_1,a_2}(\xi) = \frac{|\xi|^{2k}}{\lambda_1}$. Since $\sigma_1 \le \hat{\sigma}_1$, it is straightforward that $S_{a_1,a_2}(\xi) \le 1$ if $|\xi| < \sigma_1$. On the interval $\sigma_1 \le |\xi| < \sigma_2$, the functions $a_i(\cdot)$ are the restrictions of $FK^i_{0,\infty}(\cdot)$, $i = 1, 2$, to this interval. Therefore, by virtue of what was proved above, $S_{a_1,a_2}(\xi) \le 1$ if $\sigma_1 \le |\xi| < \sigma_2$. Finally, if $|\xi| \ge \sigma_2$, then

$$S_{a_1,a_2}(\xi) = \frac{|\xi|^{2(k-n)}}{\lambda_2} \le \frac{1}{\hat{\sigma}_2^{2(n-k)}\lambda_2} = 1.$$

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, project nos. 10-01-00188 and 10-01-90002.

REFERENCES

1. G. H. Hardy, J. E. Littlewood, and G. Polya, *Inequalities* (Cambridge Univ. Press, Cambridge, 1934; Inostrannaya Literatura, Moscow, 1948).
2. G. G. Magaril-Il'yaev and K. Yu. Osipenko, Funkts. Anal. Ego Prilozh. **37**, 51–64 (2003).
3. G. G. Magaril-Il'yaev and K. Yu. Osipenko, Funkts. Anal. Ego Prilozh. **44**, 76–79 (2010).
4. L. Hörmander, "Estimates for Translation Invariant Operators in L_p Spaces," Acta Math. **104**, 93–140 (1960) (Inostrannaya Literatura, Moscow, 1962).

SPELL OK