

On the Correctness of Finite-Rank Approximations by Series of Shifted Gaussians

S. M. Sitnik^{1*}, A. S. Timashov^{1**}, and S. N. Ushakov^{2***}

(Submitted by A. M. Elizarov)

¹Belgorod State University, Belgorod, 308015 Russia

²Voronezh State University, Voronezh, 394018 Russia

Received November 26, 2019; revised December 5, 2019; accepted December 6, 2019

Abstract—In this paper we consider interpolation problem connected with series by integer shifts of Gaussians. Known approaches for these problems met numerical difficulties. Due to it another method is considered based on finite-rank approximations by linear systems. The main result for this approach is to establish correctness of the finite-rank linear system under consideration. And the main result of the paper is to prove correctness of the finite-rank linear system approximation. For that an explicit formula for the main determinant of the linear system is derived to demonstrate that it is non-zero.

DOI: 10.1134/S1995080220030166

Keywords and phrases: *integer shifts, Gaussian, linear systems, quadratic exponents, Vandermonde matrices.*

1. INTRODUCTION AND PROBLEM STATEMENT

For a long period the main instrument for approximations was based on expansions by complete orthogonal systems. But nowadays in different fields of mathematics and applications occurred more and more problems solutions of which needs expansions by incomplete, overdetermined or non-orthogonal systems. Such problems are widely considered for electric or optic signals, filtration, holography, tomography and medicine. We have such examples of overdetermined systems as frames, and non-orthogonal as wavelets, Gabor systems and coherent states, Rvachev's systems and so on.

Consider a problem of approximation for arbitrary function defined by its values at integer points. And stress once more that functions considered are really arbitrary as we work only with function values at integer points. Let consider approximations by special type series in integer shifts of Gaussians (quadratic exponents with parameters). This system is incomplete in standard spaces but all the same very effective. For the history of this class of approximations, basic results and multiple applications see [1–3].

Exactly we will study the next main

Interpolation problem: consider arbitrary function $f(x)$ defined on the line $x \in \mathbb{R}$, and a parameter $\sigma > 0$, which in probabilistic applications means the standard deviation. We seek for an interpolating function $\tilde{f}(x)$ also defined on the line $x \in \mathbb{R}$ in the form of the series by integer shifts of Gaussian

$$\tilde{f}(x) \sim \sum_{k=-\infty}^{\infty} f_k e^{-\frac{(x-k)^2}{2\sigma^2}} \quad (1)$$

*E-mail: sitnik@bsu.edu.ru

**E-mail: loaderrus@gmail.com

***E-mail: ushakoww@yandex.ru

which coincides with $f(x)$ at integer points

$$f(m) = \tilde{f}(m), \quad m \in \mathbb{Z}. \quad (2)$$

Some approaches to this problem were studied. In [1] the solution is derived with using special functions, exactly Jacobi theta-functions. But as it was demonstrated in [2–3] in spite of theoretical value such approach has not computational perspectives as it involves division on very small denominators which leads to large numerical errors. Another approach based on discrete Fourier transform (DFT) is effective but for very small range of parameters [3]. To overcome these difficulties in [4] another approach was considered based on a reduction to finite-rank linear system solution. A serious obstacle for this last approach was that it was not grounded as a correctness of this method was not established. In this paper we eliminate this gap by proving that a finite-rank linear system is correctly solved. We prove that in a direct and clear way by calculating explicitly the main system determinant and show that it is non-zero. This result is a necessary theoretical ground for a development of numerical methods and algorithms and also rid of difficulties with special functions and DFT.

2. REDUCING TO AN INFINITE SYSTEM OF LINEAR EQUATIONS

Let introduce a convenient notation

$$e(\sigma, x, k) = e^{-\frac{(x-k)^2}{2\sigma^2}}. \quad (3)$$

Solving of the problem formulated is reduced to determination of a sequence of unknown coefficients f_k from (1). For this following the standard way for interpolation problems we need to find nod functions for every interpolation integer nodes $x = m$, $m \in \mathbb{Z}$. For our case we have to find a single *basic node function* for the node $x = 0$ which we will seek in the form

$$G(\sigma, x) = \sum_{k=-\infty}^{\infty} g_k e(\sigma, x, k). \quad (4)$$

We shortly consider steps to derive an infinite linear system for finding coefficients of the basic node function (4), cf. [2–3].

It follows from (2) that this basic node function satisfies for all $m \in \mathbb{Z}$:

$$G(\sigma, m) = \sum_{k=-\infty}^{\infty} g_k e(\sigma, m, k) = \sigma_{m0}, \quad (5)$$

where σ_{m0} is the Kroneker symbol

$$\sigma_{m0} = \begin{cases} 1, & m = 0, \\ 0, & m \neq 0. \end{cases}$$

If such a function $G(\sigma, x)$ obeying a condition (5) is found then it is easy to find a formal solution of the problem considered. In fact the function $G_l(\sigma, x) = G(\sigma, x - l)$ is the node function for the node $x = l$ as for all m $G_l(\sigma, m) = G(\sigma, m - l) = \sigma_{ml}$. Then one solution of the interpolating problem considered will be a function

$$\tilde{f}(x) = \sum_{l=-\infty}^{\infty} f(l) G_l(\sigma, x), \quad (6)$$

because for $x = m$ in the sum (6) only one term is non-zero: $f(m) G_m(\sigma, m) = f(m) \cdot 1 = f(m)$. To move from solution representation (6) to the one we seek in the form (1) do a necessary substitution. By it we derive using (4):

$$\tilde{f}(x) = \sum_{l=-\infty}^{\infty} f(l) G_l(\sigma, x) = \sum_{l=-\infty}^{\infty} f(l) G(\sigma, x - l) = \sum_{l=-\infty}^{\infty} f(l) \sum_{k=-\infty}^{\infty} g_k e(\sigma, x - l, k).$$

Let introduce a new summation index $j = l + k$ instead of $l = j - k$ and formally change the order of summation. We derive

$$\begin{aligned} \tilde{f}(x) &= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} f(j-k)g_k e(\sigma, x-j, k) \\ &= \sum_{l=-\infty}^{\infty} \left\{ \sum_{l=-\infty}^{\infty} f(j-k)g_k \right\} e(\sigma, x-j, k) = \sum_{j=-\infty}^{\infty} f_j e(\sigma, x, j), \end{aligned} \tag{7}$$

where coefficients to find are represented in the form (after index change $j \rightleftharpoons k$) to agree with (1)

$$f_k = \sum_{j=-\infty}^{\infty} f(k-j)g_j, \tag{8}$$

and $f(m)$ are values of the given function at integer points, and g_j are coefficients of the basic node function series (4).

Now transform the system of equations:

$$\sum_{k=-\infty}^{\infty} g_k e(\sigma, m, k) = \sigma_{m0}, \quad m \in \mathbb{Z}.$$

For that let introduce a new variable $q = e^{-\frac{1}{2\sigma^2}}$. And after that we derive

$$\sum_{k=-\infty}^{\infty} g_k q^{(m-k)^2} = \sigma_{m0}, \quad m \in \mathbb{Z}. \tag{9}$$

The next important step is to consider and study finite-rank cuts of this infinite system of equations (9).

3. REDUCING TO A FINITE-RANK SYSTEM OF LINEAR EQUATIONS

As it was considered before the key moment in the interpolation approach is to define the node function. Now we pass to study finite-rank approximations of the initial problem. These approximations are naturally finite-rank cuts of infinite system of equations (9). This direct approach was studied in [4]. Being in some way restrictive this approach allows to avoid difficulties from other approaches in [1–3]. And the most important is that our method allows to construct effective numerical calculations based on it, cf. [4].

So let us seek for node function (4) $G(\sigma, x)$ approximations of the form $H(\sigma, x)$ as finite sums

$$H(n, x, \sigma) = \sum_{k=-n}^n d_k q^{(x-k)^2}, \quad q = \exp\left(-\frac{1}{2\sigma^2}\right), \quad 0 < q < 1, \tag{10}$$

after that an infinite system (9) is reduced to a finite one, and number of equations may be greater than a number of unknown quantities:

$$H(n, m, j, \sigma) = \delta_{0j}, \quad j = -m, \dots, 0, \dots, m, \quad m \geq n. \tag{11}$$

The system of linear equations which follows from conditions (10)–(11) there are $2m + 1$ equations and $d_k, -n \leq k \leq n$ unknown numbers.

Transform the system following from (10)–(11) in the matrix form

$$A\bar{d} = \bar{y}, \tag{12}$$

with matrix and vector elements

$$a_{ij} = q^{(i-j)^2}, \quad y_j = \delta_{0j}, \quad i = -n, \dots, 0, \dots, n, \quad j = -m, \dots, 0, \dots, m.$$

For considered for $m = n$ coefficients d_k of approximate node function we denote as $H(n, x, \sigma)$, and for $m > n$ as $H(n, m, x, \sigma)$.

Let introduce some notation. Vandermond determinant $n \times n$ we denote as $W(x_1, \dots, x_n)$, this determinant without line l and column k as $W_{l,k}(x_1, \dots, x_n)$. For illustrating further considerations let use $n = 5$ as examples of arbitrary dimensions, e.g.

$$W(x_1, x_2, x_3, x_4, x_5) = \begin{vmatrix} 1 & x_1 & x_1^2 & x_1^3 & x_1^4 \\ 1 & x_2 & x_2^2 & x_2^3 & x_2^4 \\ 1 & x_3 & x_3^2 & x_3^3 & x_3^4 \\ 1 & x_4 & x_4^2 & x_4^3 & x_4^4 \\ 1 & x_5 & x_5^2 & x_5^3 & x_5^4 \end{vmatrix}, \quad W_{3,2}(x_1, x_2, x_3, x_4, x_5) = \begin{vmatrix} 1 & x_1^2 & x_1^3 & x_1^4 \\ 1 & x_2^2 & x_2^3 & x_2^4 \\ 1 & x_4^2 & x_4^3 & x_4^4 \\ 1 & x_5^2 & x_5^3 & x_5^4 \end{vmatrix}.$$

It is known that

$$W_{l,k}(x_1, \dots, x_n) = \sum x_{\alpha_1} x_{\alpha_2} \dots x_{\alpha_{n-k}} \prod_{n \geq i > j \geq 1, i \neq l, j \neq l} (x_i - x_j), \tag{13}$$

with summation by all sets of $n - k$ numbers $\alpha_1, \alpha_2, \dots, \alpha_{n-k}$ from $1, 2, \dots, n$. We have also to mention that considered problems are deeply connected with the third Jacobi Theta-function [5]

$$\vartheta_3(z, q) = 1 + 2 \sum_{k=1}^{\infty} q^{k^2} \cos(2kz),$$

and Jacobi–Poisson transformation formula

$$\sum_{k=-\infty}^{+\infty} e^{-a(t+\pi k)^2} = \frac{1}{\sqrt{\pi a}} \sum_{k=-\infty}^{+\infty} e^{-\frac{k^2}{a}} e^{i2kt}. \tag{14}$$

For these connections see [1–3].

4. A MAIN CASE THEN NUMBER OF EQUATIONS EQUALS TO A NUMBER OF UNKNOWN COEFFICIENTS

4.1. A Correct Solvability and Formula for the Determinant for the $n \times n$ System

Now the main object of our study is a finite system of linear equations (12) with square matrices. We use notation introduced in (10)–(12).

Let prove our main result of correctness of system solution for square matrix. From it follows that the system is always uniquely solved.

Theorem 1. *The matrix A for $m = n$ is non-degenerate and its determinant equals to*

$$|A| = q^{\frac{2n(n+1)(2n+1)}{3}} W(q^{-2n}, \dots, 1, \dots, q^{2n}). \tag{15}$$

Proof.

$$|A| = \begin{vmatrix} 1 & q & q^4 & q^9 & q^{16} \\ q & 1 & q & q^4 & q^9 \\ q^4 & q & 1 & q & q^4 \\ q^9 & q^4 & q & 1 & q \\ q^{16} & q^9 & q^4 & q & 1 \end{vmatrix}.$$

Elements of this determinant is possible to factorize $a_{ij} = q^{(i-j)^2} = q^{i^2} q^{-2ij} q^{j^2}$. Consequently it is possible to factor out from i -th line a value q^{i^2} and from j -th column— q^{j^2} . Repeat this operation for all lines and columns

$$|A| = q^4 \cdot q \cdot 1 \cdot q \cdot q^4 \cdot \begin{vmatrix} q^{-4} & q^{-3} & 1 & q^5 & q^{12} \\ 1 & q^{-1} & 1 & q^3 & q^8 \\ q^4 & q & 1 & q & q^4 \\ q^8 & q^3 & 1 & q^{-1} & 1 \\ q^{12} & q^5 & 1 & q^{-3} & q^{-4} \end{vmatrix} = q^{20} \begin{vmatrix} q^{-8} & q^{-4} & 1 & q^4 & q^8 \\ q^{-4} & q^{-2} & 1 & q^2 & q^4 \\ 1 & 1 & 1 & 1 & 1 \\ q^4 & q^2 & 1 & q^{-2} & q^{-4} \\ q^8 & q^4 & 1 & q^{-4} & q^{-8} \end{vmatrix}.$$

Now elements of intermediate determinant are q^{-2ij} , further factor out from i -th lines multipliers q^{2ni} , $i = -2n, \dots, 2n$ with a unit product:

$$q^{20} \begin{vmatrix} q^{-8} & q^{-4} & 1 & q^4 & q^8 \\ q^{-4} & q^{-2} & 1 & q^2 & q^4 \\ 1 & 1 & 1 & 1 & 1 \\ q^4 & q^2 & 1 & q^{-2} & q^{-4} \\ q^8 & q^4 & 1 & q^{-4} & q^{-8} \end{vmatrix} = q^{20} \begin{vmatrix} 1 & q^4 & q^8 & q^{12} & q^{16} \\ 1 & q^2 & q^4 & q^6 & q^8 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & q^{-2} & q^{-4} & q^{-6} & q^{-8} \\ 1 & q^{-4} & q^{-8} & q^{-12} & q^{-16} \end{vmatrix} \\ = q^{20} \begin{vmatrix} 1 & q^4 & (q^4)^2 & (q^4)^3 & (q^4)^4 \\ 1 & q^2 & (q^2)^2 & (q^2)^3 & (q^2)^4 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & q^{-2} & (q^{-2})^2 & (q^{-2})^3 & (q^{-2})^4 \\ 1 & q^{-4} & (q^{-4})^2 & (q^{-4})^3 & (q^{-4})^4 \end{vmatrix}.$$

In the general case the result is:

$$\det A = \left(\prod_{i=-n}^n q^{i^2} \right)^2 \left(\prod_{j=-n}^n q^{j^2} \right)^2 \prod_{i=-n}^n q^{2ni} W(q^{-2n}, \dots, 1, \dots, q^{2n}) \\ = q^{\frac{2n(n+1)(2n+1)}{3}} W(q^{-2n}, \dots, 1, \dots, q^{2n}) = q^{\frac{2n(n+1)(2n+1)}{3}} \prod_{i,j=-n, i \neq j}^n (q^{-2i} - q^{-2j}).$$

Now it is obvious that the main system determinant is non-zero due to inequalities under considered restrictions $0 < q < 1$, because of the obvious fact all multipliers in the last formula for the determinant are non-zeros.

Theorem 1 is completely proved.

Remark. The problem to prove correctness of the finite linear system (12) and so to ground strictly the method of interpolation by shifted Gaussian was first formulated by L.A. Minin and S.M. Sitnik. After rather a time on the basis of numerical computations the formula for the determinant (15) was found by A.S. Timashov. The strict proof was found by S.N. Ushakov. After that the correctness of the finite linear system (12) was established and it is a basis for numerical computer methods and algorithms.

4.2. Further Properties in case of $n \times n$ System: Palindromes, Symmetry and Dimension Reduction

To formulate further results we define the notion of vector palindrome, this notion is connected with vector components symmetry with respect to its “middle” components.

Definition 1. The vector \bar{x} in n dimensions is called a palindrome if the next relation for its components is fulfilled $\bar{x}_i = \bar{x}_{n+1-i}$, $i = 1, \dots, n$.

Theorem 2. Let consider the next linear system of equations

$$A\bar{x} = \bar{b}, \quad (16)$$

where A is a non-degenerate matrix $n \times n$ and for its elements the next relation is valid

$$a_{i,j} = a_{n+1-i,n+1-j} \quad \forall i, j = 1, \dots, n,$$

and vector \bar{b} is a palindrome. Then the system solution \bar{x} is also a palindrome.

In fact the condition from Theorem 2 means that every matrix A line and column are also palindromes.

Proof. As the matrix A is non-degenerate then the unique solution \bar{x} exists. Every i -th system line may be written as

$$\sum_{j=1}^n a_{i,j} x_j = b_i.$$

Let us prove that a vector $\bar{y} = (x_n, x_{n-1}, \dots, x_1)$ is also a solution to (16), and from it due to uniqueness property the statement of the Theorem 2 follows.

Really for the i -th line

$$\sum_{j=1}^n a_{i,j} y_j = \sum_{j=1}^n a_{n+1-i,n+1-j} y_j = \sum_{j=1}^n a_{n+1-i,j} x_j = b_{n+1-i} = b_i.$$

And from it the Theorem 2 follows.

Now let us demonstrate the importance of the palindrome property. Using it and symmetry of the system it is possible to decrease a number of calculations essentially, approximately twice. It leads to more effective and robust numerical algorithms.

Corollary 1. For the system (12) the solution is symmetrical, namely $d_k = d_{-k}$.

Really let us note that for the system (12) conditions of the Theorem 2 is valid:

$$a_{ij} = q^{(i-j)^2} = q^{(n+1-i-(n+1-j))^2} = a_{n+1-i,n+1-j}.$$

It follows that $d_k = d_{-k}$.

This corollary leads to a very important conclusion. Due to it we can reduce a number of equations to solve numerically in fact twice, and it also reduce gaps in values of coefficients. Because of that the numerical difficulties of the problem and computational time are reduced essentially. It permits to solve twice larger systems for the same time as systems without using palindrome symmetry.

Another important fact for the system analysis is that we can find coefficients for approximations of the node function d_k explicitly, they are defined by (10) and are solutions to the system (11)–(12).

Theorem 3. For coefficients d_k the next formula is valid:

$$d_k = (-1)^k q^{-k^2} \frac{W_{k,n+1}(q^{-2n}, \dots, q^0, \dots, q^{2n})}{W(q^{-2n}, \dots, q^0, \dots, q^{2n})}. \quad (17)$$

Proof. Really by Cramer's rule $d_k = \Delta_k / |A|$. Manipulating with Δ_k in the same way as in the proof of the Theorem 1 we derive

$$\Delta_k = (-1)^{n+1+k+1+n} q^{-k^2} q^{\frac{2n(n+1)(2n+1)}{3}} W_k(q^{-2n}, \dots, q^0, \dots, q^{2n}).$$

So

$$d_k = (-1)^k q^{-k^2} \frac{W_k(q^{-2n}, \dots, q^0, \dots, q^{2n})}{W(q^{-2n}, \dots, q^0, \dots, q^{2n})}.$$

The Theorem 3 is proved.

For example let illustrate the determinant formula (17) by determinants 5×5 as we do in this paper for clarity and reader’s convenience.

$$\begin{aligned} \Delta_1 &= \begin{vmatrix} 1 & q & q^4 & 0 & q^{16} \\ q & 1 & q & 0 & q^9 \\ q^4 & q & 1 & 1 & q^4 \\ q^9 & q^4 & q & 0 & q \\ q^{16} & q^9 & q^4 & 0 & 1 \end{vmatrix} = q^{20} \begin{vmatrix} q^{-8} & q^{-4} & 1 & 0 & q^8 \\ q^{-4} & q^{-2} & 1 & 0 & q^4 \\ 1 & 1 & 1 & q^{-1} & 1 \\ q^4 & q^2 & 1 & 0 & q^{-4} \\ q^8 & q^4 & 1 & 0 & q^{-8} \end{vmatrix} = q^{20} q^{-1} \begin{vmatrix} q^{-8} & q^{-4} & 1 & 0 & q^8 \\ q^{-4} & q^{-2} & 1 & 0 & q^4 \\ 1 & 1 & 1 & 1 & 1 \\ q^4 & q^2 & 1 & 0 & q^{-4} \\ q^8 & q^4 & 1 & 0 & q^{-8} \end{vmatrix} \\ &= q^{20} q^{-1} \begin{vmatrix} q^{-8} & q^{-4} & 1 & 0 & q^8 \\ q^{-4} & q^{-2} & 1 & 0 & q^4 \\ 1 & 1 & 1 & 1 & 1 \\ q^4 & q^2 & 1 & 0 & q^{-4} \\ q^8 & q^4 & 1 & 0 & q^{-8} \end{vmatrix} = -q^{20} q^{-1} \begin{vmatrix} 1 & q^4 & (q^4)^2 & (q^4)^4 \\ 1 & q^2 & (q^2)^2 & (q^2)^4 \\ 1 & q^{-2} & (q^{-2})^2 & (q^{-2})^4 \\ 1 & q^{-4} & (q^{-4})^2 & (q^{-4})^4 \end{vmatrix}. \end{aligned}$$

And finally let us mention that to problems considered in this paper the method of transmutations may be applied [6, 7]. An idea of applying transmutations is to represent a set of Gauss functions as transmuted ones from a set of more standard functions, e.g. linear exponentials. It means that for some transmutation operator T it is valid that

$$\exp(-(x - k)^2) = T(\exp(-(x - k))), \quad k \in \mathbb{Z}. \tag{18}$$

So the series in shifted Gaussians may be transformed to series in shifted linear exponents. Such transmutation operators T may be found explicitly but are rather complicated.

5. FINAL CONCLUSIONS AND SHORT PAPER RESUME

In this paper we consider interpolation problem connected with series by integer shifts of Gaussians. Known approaches for these problems met numerical difficulties. Due to it the direct method is considered based on finite-rank approximations by linear systems. The main result of the paper (Theorem 1) is to prove correctness of the finite-rank linear system approximation. For that an explicit formula for the main determinant of the linear system is derived to demonstrate that it is non-zero. More we consider symmetry properties of the system based on the palindrome notion. It leads to the reduction of the considered system essentially, almost twice in dimension (Theorem 2). And at last we derive an explicit formula for the system solution as a ratio of two Vandermond determinants (Theorem 3).

REFERENCES

1. V. Mazya and G. Schmidt, *Approximate Approximations*, AMS Mathematical Surveys and Monographs (Am. Math. Soc., 2007).
2. E. A. Kisilev, L. A. Minin, I. Ya. Novikov, and S. M. Sitnik, “On the Riesz constants for systems of integer translates,” *Math. Notes* **96**, 228–238 (2014).
3. M. V. Zhuravlev, E. A. Kisilev, L. A. Minin, and S. M. Sitnik, “Jacobi theta-functions and systems of integer shifts of Gaussians,” *Mod. Math. Appl.* **67**, 107–116 (2010).
4. S. M. Sitnik and A. S. Timashov, “Finite-rank mathematical model and its computation for the problem of quadratic exponential interpolation,” *Vestn. Belgor. Univ., Mat. Fiz.* **32**, 184–186 (2013).
5. E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis* Cambridge Mathematical Library (Cambridge Univ. Press, 1996).
6. V. V. Katrakhov and S. M. Sitnik, “The transmutation method and boundary-value problems for singular elliptic equations,” *Contemp. Math. Fundam. Direct.* **64**, 211–426 (2018).
7. S. M. Sitnik and E. L. Shishkina, *Method of Transmutations for Differential Equations with Bessel Operators* (Fizmatlit, Moscow, 2019) [in Russian].