Bull. Nov. Comp. Center, Num. Anal., 7 (1996), 7–18
© 1996 NCC Publisher

Tensor preconditioners in decomposition methods*

A.Yu. Bezhaev

Introduction

One of the most effective approach in solution of mesh and finite-element SLAEs

$$Au = f, \tag{1}$$

arrising in approximation of two-dimensional (or multi-dimensional) problems is the decomposition method. The essense of the method consists in special choice of easy-invertible linear transformation H and in a successive realization of iterating process

$$u^{k+1} = u^k - H^{-1}(Au^k - f). \tag{2}$$

The transformation H is often taken from the reasons, reflecting special properties of the initial problem, which is approximated by SLAE. From the point of linear algebra these reasons may be expressed with the help of the following example. Let us approximate the matrix A by the product $H = A_1 \cdot A_2$ of matrices having more simple structure (upper- and lower band). The inversion of the product matrix H on an arbitrary vector is done easily with the help of the sweep Cholessky method. It causes high efficiency of iterative process (2). Matrix H can also be used as a preconditioner in the method of conjugate gradients, this additionally speeds up the solution of the problem (1).

Usually preconditioners are chosen on the base of theoretical spectral assimptotic estimations of convergence rates of iterating process. Note, that the theoretical estimates on a little number of iterations do not take place. However, in most cases the practically satisfactory solution for a problem is obtained for a number of iterations k, which is much smaller than the dimension N of the matrix A. For this reason it is naturally to suppose the existence of essentially more effective preconditioners than the theoretical ones, if we will not assume obtaining large accuracy in solving system. We find such an preconditioner in Section 1, having stated and approximately

*Supported by the Russian Foundation of Basic Research under Grant 95–01–00949.

solved the problem of parameter optimization for a well-known splitting method, minimizing arrising error with the given smoothness of the solution and the given number of iterations k.

In the subsequent sections we suggest a new general algebraic approach for searching preconditioners for splitting methods on the basis of tensor decomposition of matrices, having effective realization. Actually the sections are not connected with the content of the first section, except identical model example (the Dirichlet problem for Poisson's equation).

1. Optimal choice of parameter τ in the assumption of smoothness

Assume some representation of the matrix of SLAE in the form $A = A_1 + A_2$. Then the factorized matrix H of decomposition method (2) can be chosen in the following known form:

$$H = H_\tau = \frac{1}{\tau}(E + \tau A_1)(E + \tau A_2). \tag{3}$$

Optimal parameter for this method is equal $\tau = 1/\sqrt{Mm}$. Here m and M are, respectively, least among minimum and largest among maximum eigen-values of matrices A_1 and A_2. In this case optimal parameter is determined from a condition of minimization of the spectral norm of the transition matrix

$$(E + \tau A_2)^{-1}(E - \tau A_2)(E + \tau A_1)^{-1}(E - \tau A_1).$$

Optimal parameter τ gives the best solution on all spectrum: one iteration equally diminishes the components of an error, corresponding to the minimum and maximum eigen-value and vector of matrix A. The components corresponding to intermediate average eigen-value $(m + M)/2$ are reduced best of all. But in practice it is usual the situation, when in a vector of solution, hence, in an initial vector of error, the low frequency amplitudes are great in comparison with high-frequency. The given fact is a consequence of the following property of smooth functions decomposition in the Fourier harmonics. If the vector u is a trace of function of the space $W_2^n[0,1]$ to the equal spaced mesh, then the following representation is valid:

$$u = \sum_{m=1}^{N} u_m \phi_m = \sum_{m=1}^{N} \frac{c_m}{m^n} \phi_m,$$

where $c_m < C$, constant C is positive. It is thus possible to state optimization problem: at given distribution of amplitudes find the optimal parameter τ, at which norm of an error's vector, or, that the same, sum of squares of amplitudes is best diminished.

Assume that the error belongs to the space of traces $W_2^{\frac{1}{2}}[0,1]$ on a mesh. Then at iterations of a decomposition method it will change as follows:

$$\xi^k = \sum_{m=1}^{N} \Big(\frac{1-\tau\lambda_m}{1+\tau\lambda_m}\Big)^k \cdot \frac{c_m}{m}\phi_m,$$

and square of its A-norm –

$$\|\xi^k\|^2 = \sum_{m=1}^{N} \Big(\frac{1-\tau\lambda_m}{1+\tau\lambda_m}\Big)^{2k} \frac{c_m^2}{m^2}.$$

Here λ_m, ϕ_m are eigen-vectors of the matrix A. For determination of parameter τ giving the best convergence for k iterations of decomposition method one needs to find the minimum of the last function. Find the approximate solution of the task in the following manner. As far as at $\tau > 0$ the function

$$\varphi_\tau = \frac{1-\tau\lambda}{1+\tau\lambda}$$

monotonously decreases on variable λ and its module does not surpass unit, and the coefficients of the error decrease too, a good condition of an optimal for k iterations will be the equality of amplitudes at the first and last harmonic on k-th iteration:

$$\Big(\frac{1-\tau m}{1+\tau m}\Big)^{2k} = \frac{1}{N^2}\Big(\frac{1-\tau M}{1+\tau M}\Big)^{2k}.$$

This equation is solvable on τ. It is reduced to the following:

$$\frac{1-\tau m}{1+\tau m} = -\frac{1}{\sqrt[k]{N}}\frac{1-\tau M}{1+\tau M},$$

and further, if to denote $\beta = \frac{1}{\sqrt[k]{N}}$, it is reduced to square equation

$$(1-\tau m)(1+\tau M) = -\beta(1+\tau m)(1-\tau M).$$

Rewrite it in the form, convenient for a finding of roots

$$\tau^2(-1-\beta)mM + \tau((1-\beta)(M-m)) + \beta + 1 = 0.$$

Then it is easy to determine the discriminant

$$D = (1-\beta)^2(M-m)^2 + 4(\beta+1)^2 mM$$

and the positive root

$$\tau = \frac{(1-\beta)(M-m)+\sqrt{D}}{2(\beta+1)mM}.$$

The latter result can be generalized to a splitting method with preconditioner (3) for two-dimensional case. Assume the initial function to be of the form

$$u = \sum_{m=1}^{100}\sum_{n=1}^{100}\frac{c_{mn}}{mn}\phi_m\phi_n,$$

the matrix A to be arising in approximation of the Dirichlet problem to the Laplase equation on unit square. For this case eigen-values and eigen-vectors have the form

$$\lambda_{st} = 4\Big(\sin^2\frac{s\pi}{2(N+1)}+\sin^2\frac{t\pi}{2(N+1)}\Big)$$

and

$$u_{st,ij} = 2\sin\frac{s\pi i}{N+1}\sin\frac{t\pi j}{N+1}.$$

Numerical results can be seen from the table.

Two-dimensional problem with $100 * 100$ unknowns, $m = 0.00097$, $M = 3.99903$

k	τ	Min	Max	NORM	NORM1	NORMT
1	1013.2022	0.0100	−0.9995	0.24743	0.779	0.940
6	378.8816	0.4635	−0.9987	0.06057	0.465	0.688
11	214.4200	0.6564	−0.9977	0.03065	0.322	0.504
16	149.4312	0.7474	−0.9967	0.01904	0.231	0.370
21	115.1040	0.7996	−0.9957	0.01291	0.168	0.271
26	94.0302	0.8332	−0.9947	0.00916	0.122	0.198
31	79.8546	0.8566	−0.9938	0.00666	0.089	0.145
36	69.7153	0.8736	−0.9929	0.00492	0.065	0.106
41	62.1359	0.8866	−0.9920	0.00366	0.048	0.078
46	56.2783	0.8967	−0.9912	0.00274	0.035	0.057
∞	16.0772	0.9694	−0.9694			

Conclusions from numerical experiments

There is the significant acceleration of error's decreasing with the new optimal parameters for splitting method if the solution is smooth.

Classic choice of optimal parameter τ_{opt} does not provide accelerated decreasing of the error for the smooth solutions. The theoretical estimates of error's norm reduction with the optimal choice τ are well confirmed on individual function despite of their validity on a class.

2. Tensor factorization of matrices at construction of iterative processes

Generally speaking the system of equations (1) for two-dimensional problems has a sparse matrix. If one places nonzero elements around of the diagonal, then the width of a band will be wider only in two times, than for one-dimensional problem. Nevertheless, to achieve this with the help of replacement variable and rearrangement of equations is not possible. The application for the solution of a Cholessky method becomes not such effective as in one-dimensional case, as far as at any replacement of variables and rearrangement of equations the width of a band does not manage to make less than $O(N)$.

Cholessky sweep method helps to construct effective algorithms in two-dimensional and multi-dimensional problems when the matrix of system A is represented in a tensor product of two (and more) band matrices. Since it is not true in general, then it appears an idea of choice of a preconditioning matrix in a kind of tensor product $H = A_1 \otimes A_2$, that seems especially natural for the problems of mathematical physics in the rectangular domains when the tensor product spaces and regular grids arise. In these cases the matrix A has a structure from blocks with band matrices, and the blocks will also form a band structure, that is A has such a structure as well as tensor product of band matrices.

Other argument serving the basis for choice of such preconditioner is the algorithm suggested by the author in this article. It is intended for the solution of the following optimization problem:

$$||A - H|| = \min ||A - A_1 \otimes A_2|| \tag{4}$$

for the spherical matrix norm. It is proved that the problem is reduced to construction of the best singular decomposition for a matrix, received from A by rearrangement of elements.

As an important argument for the benefit of tensor preconditioner choice we shall put the fact, that the offered algorithm for the optimization problem (4), applicable for the matrix A of a general kind, is essentially accelerated for matrices with block band structure for the account of essential reduction the cost of calculations. Here the optimal matrices A_1 and A_2 are received band.

The attained results are illustrated on a Dirichlet problem for the Poisson's equation. Naturally, they can easily be applied to a problem of spline-smoothing on scattered meshes, which could be the subject of further investigations and applications.

3. Best approximation by a matrix of the rank 1

Let F be a rectangular matrix of dimension $N \times M$. Consider a problem of its best approximation in the spherical norm by matrices of the rank 1:

$$\min_{u\in\mathbf{R}^N,\, v\in\mathbf{R}^M} \|F - uv^T\|^2 = \min \sum_{i=1,j=1}^{N,M} (f_{ij} - u_i v_j)^2. \tag{5}$$

Here u and v are columns, their product gives a matrix of the rank 1, moreover, any matrix of the rank 1 can be decomposed in such product. The following theorem is the well-known fact in linear algebra.

Theorem 1. *Consider a problem of eigen-values*

$$\begin{bmatrix} 0 & F \\ F^* & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \lambda \begin{bmatrix} u \\ v \end{bmatrix}. \tag{6}$$

Let the components u_*, v_* compose normal eigen-vector, respecting to the maximal eigen-value λ_* for problem (6). *Then, vectors*

$$u = \sqrt{2\lambda_*}u_*, \qquad v = \sqrt{2\lambda_*}v_* \tag{7}$$

are the solution to the problem of best approximation (5).

4. Best decomposition of a matrix in tensor product

Let B and C be square matrices of dimensions $n \times n$ and $m \times m$, respectively. Remember definition of the tensor product of this matrices

$$B \otimes C = \begin{bmatrix} b_{11}C & b_{12}C & \dots & b_{1n}C \\ b_{21}C & b_{22}C & \dots & b_{2n}C \\ \dots & \dots & \dots & \dots \\ b_{n1}C & b_{n2}C & \dots & b_{nn}C \end{bmatrix}$$

having dimension $(nm) \times (nm)$. We state the problem of best representation of the matrix A in the tensor product form

$$\|A - A_1 \otimes A_2\| = \min_{U\in M_n,\, V\in M_m} \|A - U \otimes V\|. \tag{8}$$

Evidently, A_1, A_2 are the matrices of the best solution to the problem (8), the set M_n consists of any matrices of dimension $n \times n$.

Consider the matrix A as four-indexed:

$$\{a_{i_1 j_1, i_2, j_2}\}, \quad \text{where } 1 \le i_1, i_2 \le n,\ 1 \le j_1, j_2 \le m. \tag{9}$$

Having made rearrangement of indexes $h_{i_1 i_2, j_1, j_2} = a_{i_1 j_1, i_2, j_2}$ we shall proceed to other four-index matrix, which corresponds to the rectangular matrix H of dimension $n^2 \times m^2$.

Theorem 2. *The elements of matrices A_1 and A_2 which are the solutions to the best decomposition problem* (8) *are received from elements of vectors u, v, determined in Theorem 1, with the help of their two-dimensional ordering.*

Proof. Problem (8) is equivalent to the following:

$$\min \sum_{\substack{1 \le i_1, i_2 \le n \\ 1 \le j_1, j_2 \le m}} \left(a_{i_1 j_1, i_2, j_2} - u_{i_1 i_2} \cdot v_{j_1 j_2}\right)^2,$$

or, after introducing the notations – to the following problem:

$$\min \sum_{\substack{1 \le i_1, i_2 \le n \\ 1 \le j_1, j_2 \le m}} \left(h_{i_1 i_2, j_1, j_2} - u_{i_1 i_2} \cdot v_{j_1 j_2}\right)^2. \tag{10}$$

Here $u_{i_1 i_2}$ and $v_{j_1 j_2}$ are elements of matrices U and V respectively. Assuming $N = n^2$, $M = m^2$, we are really convinced that the problem of the best tensor representation is reduced to the approximation problem by a matrix of the rank 1. □

Remark 1. The matrix of the tensor product $A_1 \otimes A_2$ is not a matrix of the rank 1, moreover, in many cases it is nonsingular simmetrical positively defined matrix and it can be used as the preconditioner in iterative process.

5. Tensor decomposition of a block sparse matrix

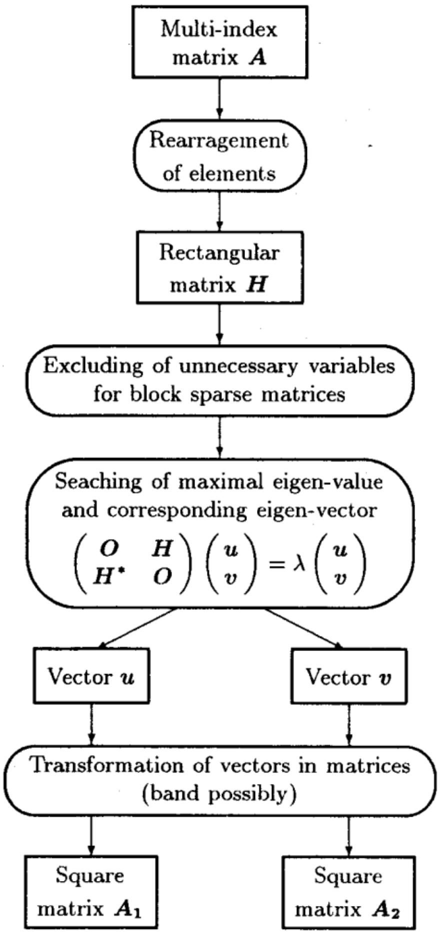
Multi-index matrix A, arising in various methods of numerical mathematics (difference, variational difference, finite element) has *special* block structure with band matrices in blocks determined by the following character of sparseness. Usually there are two positive numbers k_1, k_2 such, that the elements $a_{i_1 j_1, i_2, j_2}$ are not zerous, only if $|i_1 - i_2| \le k_1$ and $|j_1 - j_2| \le k_2$.

Theorem 3. *The optimal matrices A_1 and A_2 of the best tensor decomposition of the matrix A with special block structure are band.*

Proof. The theorem elementary follows from the representation of problem (10). Since the elements $h_{i_1 i_2, j_1, j_2}$ are equal to zero at $|i_1 - i_2| > k_1$, or $|j_1 - j_2| > k_2$, then without fail we receive $u_{i_1 i_2} = 0$ at $|i_1 - i_2| > k_1$, and $v_{j_1 j_2} = 0$ at $|j_1 - j_2| > k_2$. □

Remark 2. Tensor product decomposition of general kind matrix was required finding an eigen-vector and eigen-value for a matrix of the order $(n^2 + m^2)$ (see (6)). Special block structure matrix can be reduced to a matrix of the smaller order $O(n + m) < (2 * k_1 + 1)n + (2 * k_2 + 1) * m$.

The block-scheme illustrating the algorithm of tensor approximation is shown in the figure.


Multi-index matrix A
Rearragement of elements
Rectangular matrix H
Excluding of unnecessary variables for block sparse matrices
Seaching of maximal eigen-value and corresponding eigen-vector
$\begin{pmatrix} O & H \\ H^* & O \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \lambda \begin{pmatrix} u \\ v \end{pmatrix}$
Vector u
Vector v
Transformation of vectors in matrices (band possibly)
Square matrix A_1
Square matrix A_2

The scheme of construction of the best tensor preconditioner $A_1 \otimes A_2$ for the matrix A

6. Explicit construction of preconditioners for Dirichlet problem

In Dirichlet problem for Poisson's equation, at its approximation by finite defferences, there arises the following system of equations with block three diagonal matrix of dimension $n^2 \times n^2$:

$$A = \left(\begin{array}{ccc}
\begin{pmatrix} 4 & -1 & & 0 \\ -1 & 4 & \ddots & \\ & \ddots & \ddots & -1 \\ 0 & & -1 & 4 \end{pmatrix} & \begin{pmatrix} -1 & & 0 \\ & \ddots & \\ 0 & & -1 \end{pmatrix} & \\
\begin{pmatrix} -1 & & 0 \\ & \ddots & \\ 0 & & -1 \end{pmatrix} & \begin{pmatrix} 4 & -1 & & 0 \\ -1 & 4 & \ddots & \\ & \ddots & \ddots & -1 \\ 0 & & -1 & 4 \end{pmatrix} & \begin{pmatrix} -1 & & 0 \\ & \ddots & \\ 0 & & -1 \end{pmatrix} \\
 & \begin{pmatrix} -1 & & 0 \\ & \ddots & \\ 0 & & -1 \end{pmatrix} & \begin{pmatrix} 4 & -1 & & 0 \\ -1 & 4 & \ddots & \\ & \ddots & \ddots & -1 \\ 0 & & -1 & 4 \end{pmatrix}
\end{array}\right).$$

We present here only 9 blocks of this matrix, since the other blocks repeat these. Find the best tensor factorization for his matrix. According to Theorem 2 it is necessary for this purpose to rearrange elements of matrix A. In our case the new matrix H looks like as follows:

$$\left(\begin{array}{ccc|cccc|ccccc|c|ccc}
4 & -1 & 0\dots0 & -1 & 4 & -1 & 0\dots0 & 0 & -1 & 4 & -1 & 0\dots0 & & 0\dots0 & -1 & 4 \\
-1 & 0 & 0\dots0 & 0 & -1 & 0 & 0\dots0 & 0 & 0 & -1 & 0 & 0\dots0 & \dots & 0\dots0 & 0 & -1 \\
0 & 0 & 0\dots0 & 0 & 0 & 0 & 0\dots0 & 0 & 0 & 0 & 0 & 0\dots0 & & 0\dots0 & 0 & 0 \\
\vdots & & & & & & & & & & & & & & & \\
0 & 0 & 0\dots0 & 0 & 0 & 0 & 0\dots0 & 0 & 0 & 0 & 0 & 0\dots0 & & 0\dots0 & 0 & 0 \\
\hline
-1 & 0 & 0\dots0 & 0 & -1 & 0 & 0\dots0 & 0 & 0 & -1 & 0 & 0\dots0 & & 0\dots0 & 0 & -1 \\
4 & -1 & 0\dots0 & -1 & 4 & -1 & 0\dots0 & 0 & -1 & 4 & -1 & 0\dots0 & \dots & 0\dots0 & -1 & 4 \\
-1 & 0 & 0\dots0 & 0 & -1 & 0 & 0\dots0 & 0 & 0 & -1 & 0 & 0\dots0 & & 0\dots0 & 0 & -1 \\
0 & 0 & 0\dots0 & 0 & 0 & 0 & 0\dots0 & 0 & 0 & 0 & 0 & 0\dots0 & & 0\dots0 & 0 & 0 \\
\vdots & & & & & & & & & & & & & & & \\
\vdots & & & & & & & & & & & & & & &
\end{array}\right).$$

Here the blocks of matrix A are extensive in a line of matrix H: the block 1×1 – in the first line, block 1×2 – in the second line, further the matrix A has $n - 1$-th zero blocks, and the matrix F accordingly $n - 1$-th zero lines, after this situation is repeated. It is obvious, that the matrix F has very much zero columns and lines. According to Theorem 3 for such kind of special matrices A of a block structure ($k_1 = k_2 = 1$) matrix H can be reduced to a matrix of smaller dimension, namely to a matrix of the following system of linear algebraic equations:

$$\left(\begin{array}{rrrrrrr|c} 4&-1&-1&4&-1&-1&4& \\ -1&0&0&-1&0&0&-1& \\ -1&0&0&-1&0&0&-1& \\ 4&-1&-1&4&-1&-1&4&\cdots \\ -1&0&0&-1&0&0&-1& \\ -1&0&0&-1&0&0&-1& \\ 4&-1&-1&4&-1&-1&4& \\ \hline \multicolumn{7}{c|}{\cdots} & \cdots \end{array}\right)\left(\begin{array}{c}\alpha\\-\beta\\-\beta\\\alpha\\-\beta\\-\beta\\\alpha\\\hline\vdots\end{array}\right)=\lambda\left(\begin{array}{c}\alpha\\-\beta\\-\beta\\\alpha\\-\beta\\-\beta\\\alpha\\\hline\vdots\end{array}\right). \tag{11}$$

In generall case we need to solve the spectral problem (6). However, in our case matrix H is symmetric, therefore, the optimal vectors u and v should coincide and be the solution of spectral problem $Hu = \lambda u$. Actually, the reduced matrix H has rank 2, it has only two different linear independent lines, other are repeated. Therefore, the eigen-vector also has only two various components, we shall denote them α and $-\beta$. Thus, find the solution to (11). We have two equations:

$$\begin{cases} 4n\alpha + 2(n-1)\beta = \lambda\alpha, \\ n\alpha \qquad\qquad\quad = \lambda\beta. \end{cases}$$

From the latter equation we have $\alpha = \lambda\beta n^{-1}$ and receive a square equation relatively λ

$$\lambda^2 - 4n\lambda - 2n(n-1) = 0$$

with the following roots

$$\lambda_{\max} = 2n + \sqrt{6n^2 - 2n}, \quad \lambda_{min} = 2n - \sqrt{6n^2 - 2n}. \tag{12}$$

From Theorem 2 it follows that for the optimal decomposition it is necessary to find the eigen-vector u_* with the following norm condition $\|u_*\|^2 = 1/2$. At first, we find an eigen-vector appropriate to the maximum eigen-value. For this purpose we put $\beta = 1$ and receive $\alpha = \lambda_{\max}/n$. Hence, substituting components of required vector $\alpha_* = q\alpha$, $\beta_* = q\beta$ into the norm condition, we have

$$q^2(n(\lambda_{\max}/n)^2 + 2(n-1)) = 1/2.$$

Then, we find the unknown factor

$$q = \sqrt{\frac{n}{2(\lambda_{\max}^2 + 2n(n-1))}}, \tag{13}$$

for the components of a normalized vector

$$\alpha = \lambda_{\max} q n^{-1}, \quad \beta = q,$$

and obtain the components of the optimal vector

$$\alpha = \lambda_{\max} q n^{-1} \sqrt{2\lambda_{\max}}, \quad \beta = q\sqrt{2\lambda_{\max}}. \tag{14}$$

Thus we have received the following statement.

Theorem 4. *Factors of the best tensor product for the matrix of the Dirichlet problem on square grid are positively defined 3-diagonal matrices of the following form:*

$$A_1 = A_2 = \begin{pmatrix} \alpha & -\beta & & & & \\ -\beta & \alpha & -\beta & & & \\ & -\beta & \alpha & -\beta & & \\ & & \ddots & \ddots & \ddots & \\ & & & -\beta & \alpha & -\beta \\ & & & & -\beta & \alpha \end{pmatrix},$$

where α and β are determined by (12)–(14).

Remark 3. Being engaged matrix decomposition of a Dirichlet problem, the author has received its following beautiful decomposition in a form of a difference of two tensor products of three diagonal matrices:

$$\otimes^2 \begin{pmatrix} 2\beta + \frac{1}{2\beta} & -\beta & & \\ -\beta & \ddots & \ddots & \\ & \ddots & \ddots & -\beta \\ & & -\beta & 2\beta + \frac{1}{2\beta} \end{pmatrix} - \otimes^2 \begin{pmatrix} 2\beta - \frac{1}{2\beta} & -\beta & & \\ -\beta & \ddots & \ddots & \\ & \ddots & \ddots & -\beta \\ & & -\beta & 2\beta - \frac{1}{2\beta} \end{pmatrix}$$

valid for any $\beta > 0$.

Remark 4. Best spectral preconditioners for the matrices of the simple structure $A = A_1 \otimes E + E \otimes A_2$ (as in case of Dirichlet problem) can be constructed explicitly. They can be received on the basis of the following exact decomposition in the difference of products of lower- and upper triangular matrices

$$\begin{aligned} A_1 \otimes E + E \otimes A_2 = {} & \frac{1}{2\tau}(E + \tau A_1 \otimes E)(E + \tau E \otimes A_2) - \\ & \frac{1}{2\tau}(E - \tau A_1 \otimes E)(E - \tau E \otimes A_2) \end{aligned}$$

valid for all $\tau > 0$. Actually this decomposition is also the tensor one, since it can be rewritten in the equivalent form

$$\begin{aligned} A_1 \otimes E + E \otimes A_2 = & \frac{1}{2\tau}(E+\tau A_1)\otimes(E+\tau A_2) - \\ & \frac{1}{2\tau}(E-\tau A_1)\otimes(E-\tau A_2). \end{aligned} \tag{15}$$

Using this decomposition and the known spectrum, one can construct optimal value τ for convergence when utilizing the first component

$$\frac{1}{2\tau}(E+\tau A_1)\otimes(E+\tau A_2), \tag{16}$$

as the preconditioner.

Remark 5. Note the following analogous decomposition for the Laplace operator at the differential level

$$\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}=\frac{1}{2\tau}\Big(I+\tau\frac{d^2}{dx^2}\Big)\otimes\Big(I+\tau\frac{d^2}{dy^2}\Big)-\frac{1}{2\tau}\Big(I-\tau\frac{d^2}{dx^2}\Big)\otimes\Big(I-\tau\frac{d^2}{dy^2}\Big).$$