Institute for Computational Mathematics
Hong Kong Baptist University

ICM Research Report
08-12

Matrix Inversion Cases with Size-Independent Tensor Rank Estimates

Ivan Oseledets [1], Eugene Tyrtyshnikov [1], Nickolai Zamarashkin [1]

Institute of Numerical Mathematics, Russian Academy of Sciences, Gubkin Street, 8, Moscow 119333

Abstract

Let M be a matrix of order $n = pq$. Then the tensor rank of M is defined as the minimal possible ρ in expressions of the form $M = \sum_{t=1}^{\rho} U_t \otimes V_t$, where U_t and V_t are matrices of order p and q, respectively. Let M be a nonsingular matrix of tensor rank 3 and, moreover, of the form

$$M = I + A \otimes X + Y \otimes B$$

with $\text{rank} X = \text{rank} Y = 1$. Then, it is discovered and proved that the tensor rank of M^{-1} is bounded from above by 5 independently of p and q, the estimate being sharp. Some related and extended results are also given.

AMS classification: 15A12; 65F10; 65F15

Key words: Tensor ranks; Kronecker product; Low-rank matrices; Inverse matrices; Multilevel matrices; Toeplitz matrices; Circulant matrices.

1 Introduction

Kronecker (tensor) product decompositions and approximations are proving to be very useful in the design of fast algorithms [1,7] and modern numerical technologies for large-scale problems (for example, see [3,4,6,16]), especially in higher dimensions [2,11]. However, tensor-rank estimates are available

Email addresses: `ivan@bach.inm.ras.ru` (Ivan Oseledets), `tee@inm.ras.ru` (Eugene Tyrtyshnikov), `kolya@bach.inm.ras.ru` (Nickolai Zamarashkin).

[1] Supported by RFBR grants 05-1-00721, 06-01-08052 and Priority Research Programme of Department of Mathematical Sciences of Russian Academy of Sciences.

only for tensor approximations to special classes of function-related matrices [4,13,14], and next to nothing seems to be known in a pure algebraic context (for one of very few see [15]). In this paper we present some general estimates for matrices with no reference to functions and grids.

Given a field $\mathbb{F}$, we suppose that the entries of all matrices belong to $\mathbb{F}$. The general findings presented below are valid for an arbitrary (maybe finite) field. However, in some proofs we must needs assume that $\mathbb{F} = \mathbb{C}$.

Let M be a matrix of order $n = pq$. Then the tensor rank of M is defined as the minimal possible ρ in expressions of the form

$$M = \sum_{t=1}^{\rho} U_t \otimes V_t,$$

where U_t and V_t are matrices of order p and q, respectively. In this case we shall write

$$\text{tRank}(M) = \rho.$$

It might be worthy to emphasize that here we deal with the two-adic (two-factor) case, which differs in many respects from the case of more factors.

If M is nonsingular and $\text{tRank}(M) = 2$, then a nontrivial estimate proposed in [15] for the two-adic case reads

$$\text{tRank}(M^{-1}) \leq \min\{p, q\}.$$

If $\text{tRank}(M) = 3$, then no nontrivial estimate is known as yet.

Here, nevertheless, we consider the case $\text{tRank}(M) = 3$ and present an estimate of the form

$$\text{tRank}(M^{-1}) \leq 5.$$

It excels in being independent of p and q but capitalizes on a certain additional structure in M:

$$M = I + A \otimes X + Y \otimes B,$$

where

$$\text{rank}X = \text{rank}Y = 1.$$

This kind of structure emerges from recent approximation constructions for Toeplitz matrices [12,17] and, as we envisage, can provide some grounds for the application of approximate iterations [5] for the inversion of two-level Toeplitz matrices [9,10].

2 Main results

Lemma 2.1 *Let the entries of all matrices belong to an arbitrary field $\mathbb{F}$. Assume that a nonsingular matrix K is of the form*

$$K = I + A \otimes \left(uv^\top\right), \tag{1}$$

where A is $p \times p$ and u, v are column vectors with q entries. Then

$$K^{-1} = I + \tilde{A} \otimes \left(uv^\top\right) \tag{2}$$

with

$$\tilde{A} = -A\left(I + (v^\top u)A\right)^{-1}. \tag{3}$$

Proof. We can regard the given matrix

$$K = I + (A \otimes u)(I \otimes v^\top)$$

as a rank p modification of the identity matrix I. On application of the Sherman–Morrison–Woodbury formula,

$$K^{-1} = I - (A \otimes u)W^{-1}(I \otimes v^\top),$$

where

$$W = I + (I \otimes v^\top)(A \otimes u) = I + A \otimes (v^\top u) = I + (v^\top u)A.$$

Since

$$(A \otimes u)W^{-1} = (A \otimes u)(W^{-1} \otimes [1]) = AW^{-1} \otimes u,$$

we obtain

$$K^{-1} = I - (A \otimes u)W^{-1}(I \otimes v^\top) = I - (AW^{-1} \otimes u)(I \otimes v^\top) = I - AW^{-1} \otimes uv^\top,$$

which gives us up to (2) with $\tilde{A} = -AW^{-1}$. □

Lemma 2.2 *Let the entries of all matrices belong to an arbitrary field $\mathbb{F}$. Assume that a nonsingular matrix M is of the form*

$$L = I + \left(gh^\top\right) \otimes B, \tag{4}$$

where B is $q \times q$ and g, h are column vectors with p entries. Then

$$L^{-1} = I + \left(gh^\top\right) \otimes \tilde{B} \tag{5}$$

with

$$\tilde{B} = -(I + (h^\top g)B)^{-1}B. \tag{6}$$

Proof. It suffices to observe that $L = I + gh^\top \otimes B$ reduces to $I + B \otimes gh^\top$ by a permutation similarity transformation. Another way is a straightforward modification of the previous proof. $\square$

Theorem 2.1 *Let the entries of all matrices belong to an arbitrary field $\mathbb{F}$. Let M be a nonsingular matrix of the form*

$$M = I + A \otimes X + Y \otimes B, \tag{7}$$

where A, Y are $p \times p$ and X, B are $q \times q$, and assume that

$$\text{rank} X = \text{rank} Y = 1. \tag{8}$$

Assume additionally that the matrices $I + A \otimes X$ and $I + Y \otimes B$ are both nonsingular. Then

$$M^{-1} = I + \hat{A} \otimes X + Y \otimes \hat{B} + Y_1 \otimes X_1 + Y_2 \otimes X_2 \tag{9}$$

with some matrices $\hat{A}$, $\hat{B}$ of possibly full rank and other matrices X_1, X_2, Y_1, Y_2 of rank at most 1.

Proof. For some column vectors u, v and g, h we can write

$$X = uv^\top, \quad Y = gh^\top.$$

Then, on the base of Lemmas 2.1 and 2.2,

$$K^{-1} = I + \tilde{A} \otimes \left(uv^\top\right),$$

$$L^{-1} = I + \left(gh^\top\right) \otimes \tilde{B},$$

and, by a direct calculation,

$$KL = M + \left(Agh^\top\right) \otimes \left(uv^\top B\right).$$

Consequently,

$$M = K\left(I - K^{-1}\left(Agh^\top \otimes uv^\top B\right) L^{-1}\right) L. \tag{10}$$

Now, let us have a closer look at the matrix in between of K and L in the right-hand side. Above all, it is naturally expressed in the form $I - F$. Then, observe that

$$Agh^\top \otimes uv^\top B = (Ag \otimes u)(h^\top \otimes v^\top B)$$

is a matrix of rank 1. Hence,

$$\text{rank} F = 1,$$

and using the Sherman–Morrison–Woodburry formula we conclude that

$$(I-F)^{-1} = I + cF$$

for some scalar value c. More precisely,

$$\begin{aligned} F &= (I+\tilde{A}\otimes uv^\top)(Ag\otimes u)(h^\top\otimes v^\top B)(I+gh^\top\otimes\tilde{B}) \\ &= \left(Ag\otimes u + (v^\top u)\tilde{A}Ag\otimes u\right)\left(h^\top\otimes v^\top B + (h^\top g)h^\top\otimes v^\top B\tilde{B}\right) \\ &= (Sg\otimes u)(h^\top\otimes v^\top T) = Sgh^\top\otimes uv^\top T, \end{aligned}$$

where

$$S = A + (v^\top u)\tilde{A}A, \quad T = B + (h^\top g)B\tilde{B}.$$

Due to (3) and (6) we have

$$S = -\tilde{A}, \quad T = -\tilde{B},$$

which implies that

$$F = \tilde{A}gh^\top\otimes uv^\top\tilde{B}$$

and, therefore,

$$(I-F)^{-1} = I + c\tilde{A}gh^\top\otimes uv^\top\tilde{B}.$$

In chime with (10) and the above formulas for K^{-1} and L^{-1},

$$\begin{aligned} M^{-1} &= L^{-1}(I-F)^{-1}K^{-1} \\ &= \left(I+gh^\top\otimes\tilde{B}\right)\left(I+c\tilde{A}gh^\top\otimes uv^\top\tilde{B}\right)\left(I+\tilde{A}\otimes uv^\top\right) \\ &= I + gh^\top\otimes\tilde{B} + \tilde{A}\otimes uv^\top + c\tilde{A}gh^\top\otimes uv^\top\tilde{B} + \\ &\quad \alpha gh^\top\otimes\tilde{B}uv^\top + gh^\top\tilde{A}\otimes\tilde{B}uv^\top + \beta\tilde{A}gh^\top\tilde{A}\otimes uv^\top + \alpha gh^\top\tilde{A}\otimes\tilde{B}uv^\top, \end{aligned}$$

where

$$\alpha = ch^\top\tilde{A}g, \quad \beta = cv^\top\tilde{B}u.$$

Finally, allowing for $X = uv^\top$ and $Y + gh^\top$, we arrive at (9) by setting

$$\hat{A} = \tilde{A} + \beta\tilde{A}gh^\top\tilde{A}, \quad \hat{B} = \tilde{B} + \alpha\tilde{B}uv^\top,$$

$$Y_1 = (1+\alpha)gh^\top\tilde{A}, \quad X_1 = \tilde{B}uv^\top, \quad Y_2 = c\tilde{A}gh^\top, \quad X_2 = uv^\top\tilde{B},$$

which completes the proof. □

Corollary 2.1 *Under the premises of Theorem 2.1, the inverse matrix M^{-1} can be alternatively written in the form*

$$M^{-1} = I + \tilde{A}\otimes uv^\top + gh^\top\otimes\tilde{B} + \hat{Y}_1\otimes X_1 + Y_2\otimes\hat{X}_2,$$

where $\tilde{A}$ and $\tilde{B}$ are defined by (2) and (5), and the matrices X_1, Y_2 are of rank at most 1 while $\hat{Y}_1$, $\hat{X}_2$ are of rank at most 2.

Theorem 2.2 *Let $\mathbb{F} \subset \mathbb{C}$. If a matrix of the form*

$$M = C \otimes D + A \otimes X + Y \otimes B$$

is nonsigular and, besides that, $\text{rank} X = \text{rank} Y = 1$*, then*

$$\text{tRank}(M^{-1}) \le 5.$$

Proof. If the matrices C and D are nonsingular, then

$$M = (C \otimes D)\left(I + C^{-1}A \otimes D^{-1}X + C^{-1}Y \otimes D^{-1}B\right)$$

and the case reduces to direct application of Theorem 2.1, provided that the matrices $K = I + C^{-1}A \otimes D^{-1}X$ and $L = I + C^{-1}Y \otimes D^{-1}B$ are nonsingular. If the latter does not hold for the given A and B, it still does for their ε-approximations $A_\varepsilon \approx A$ and $B_\varepsilon \approx B$ for all sufficiently small $\varepsilon > 0$. Furthermore, at any rate we can find nonsingular matrices C_ε and D_ε within an ε-distance from C and D, respectively; then we fall back on the following: if a sequence of matrices converges, then the limit matrix is of tensor rank that cannot exceed the maximum of tensor ranks of the sequence terms (note that this is no longer so in the m-adic case with $m > 2$) [8,15]. □

For given p and q we can find a matrix M with $\text{tRank}(M^{-1}) = 5$ using a computer. That is what we really did for a sample of p and q. If we have such a matrix for some $p = p_0$ and $q = q_0$ than it is easy to produce a matrix $\tilde{M}$ with the same property $\text{tRank}(M^{-1}) = 5$ for any $p \ge p_0$ and $q = q_0$. In particular, we can consider

$$\tilde{M} = \begin{bmatrix} M & 0 \\ 0 & I \end{bmatrix}.$$

It is easy to recognize that $\tilde{M}$ is structured according to (7) so long as M is.

3 Another proof and inversion algorithm

Given a matrix Z, denote by

$$z = \text{VECTOR}(Z)$$

a column vector assembling the columns of Z in the natural order. In this case, let us also agree to write

$$Z = \text{MATRIX}(z),$$

no ambiguity rising if the sizes are fixed (below they are clear from the context). These *vectorization* and *matrization* operations are closely related to a Kronecker product multiplication by a vector.

Lemma 3.1 *For any matrices A and B*

$$\mathrm{VECTOR}(BZA^\top) = (A \times B)z, \quad z = \mathrm{VECTOR}(Z).$$

The claim is checked straightforwardly.

Now, let M be a matrix of order $n = pq$. Then a kth column of M can be pointed to by a unique pair of indices $(i, j) \leftrightarrow k$ such that

$$k = (i-1)q + j, \quad 1 \le i \le p, \quad 1 \le j \le q.$$

Lemma 3.2 *Assume that $A = [a_1, ..., a_p]$ is $p \times p$ and $B = [b_1, ..., b_q]$ is $q \times q$. Then M is of the Kronecker product structure*

$$M = A \otimes B$$

if and only if the columns M_k of M are of the form

$$M_k = \mathrm{VECTOR}(b_j a_i^\top), \quad k \leftrightarrow (i, j).$$

Proof. Obviously, $(A \otimes B)_k = a_i \otimes b_j$. Then, by application of Lemma 3.1 with $z = [1]$ we obtain

$$a_i \otimes b_j = \mathrm{VECTOR}(b_j a_i^\top). \quad \square$$

Let e'_j and e_i are the columns j and i of the identity matrices of order q and p, respectively. If $k \leftrightarrow (i, j)$ then the column k of the identity matrix of order $n = pq$ is of the form $e_i \otimes e'_j$. By Lemma 3.2,

$$\mathrm{MATRIX}(e_i \otimes e'_j) = e'_j e_i^\top.$$

Suppose that z is a kth column of the inverse to a matrix M of the form

$$M = I + A \otimes uv^\top + gh^\top \otimes B.$$

Then, taking into account that $\mathrm{MATRIX}(Mz) = e'_j e_i^\top$ and using Lemma 3.1, we obtain

$$Z + uv^\top ZA^\top + BZhg^\top = e'_j e_i^\top,$$

and thence

$$Z = e'_j e_i^\top - xg^\top - uy^\top, \tag{11}$$

where

$$x = BZh, \quad y = AZ^\top v. \tag{12}$$

Thus, the kth column z of M^{-1} is given by the matrix Z which is of rank at most 3 and completely defined by the vectors x and y depending on $k \leftrightarrow (i,j)$. To calculate these vectors, we derive convenient equations by excluding Z from (11) and (12). The first equation comes by multiplication of both sides of (11) by B from the left and by h from the right while the second appears from the transposed equation after premultiplication by A and postmultiplication by v. In the result,

$$\begin{aligned} x =& \; Be'_j e_i^\top h - Bxg^\top h - Buy^\top h, \\ y =& \; Ae_i (e'_j)^\top v - Agx^\top v - Ayu^\top v, \end{aligned}$$

or, in the matrix form,

$$\begin{bmatrix} I + (g^\top h)B & Buh^\top \\ Agv^\top & I + (u^\top v)A \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} Be'_j e_i^\top h \\ Ae_i (e'_j)^\top v \end{bmatrix}. \tag{13}$$

Let us assume that the diagonal blocks $I + (g^\top h)B$ and $I + (u^\top v)A$ are both nonsingular. Then (13) can be recast in a simpler form

$$\begin{bmatrix} I & \hat{B}uh^\top \\ \hat{A}gv^\top & I \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \hat{B}e'_j e_i^\top h \\ \hat{A}e_i (e'_j)^\top v \end{bmatrix},$$

where

$$\hat{B} = \Big(I + (g^\top h)B\Big)^{-1} B, \qquad \hat{A} = \Big(I + (u^\top v)A\Big)^{-1} A.$$

By the block Gaussian elimination it reduces to

$$\begin{bmatrix} I & \hat{B}uh^\top \\ 0 & I - (v^\top \hat{B}u)\hat{A}gh^\top \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} h_i \hat{B}e'_j \\ v_j \hat{A}e_i - c_j h_i \hat{A}g \end{bmatrix}$$

with

$$h_i = e_i^\top h, \quad v_j = (e'_j)^\top v, \quad c_j = (v^\top \hat{B}u)(h^\top \hat{B}e'_j).$$

According to the Sherman–Morrison–Woodbury formula,

$$(I - (v^\top \hat{B}u)\hat{A}gh^\top)^{-1} = I + c\hat{A}gh^\top$$

with some scalar c. Therefore,

$$y = \Big(I + c\hat{A}gh^\top\Big)\Big(v_j\hat{A}e_i - c_j h_i \hat{A}g\Big) = v_j \hat{A}e_i + (d_j h_i + c v_j s_i)\hat{A}g,$$

where

$$d_j = -c_j - c_j c(h^\top \hat{A}g), \quad s_i = h^\top \hat{A}e_i.$$

It now follows that

$$x = h_i \hat{B} e'_j - \hat{B} u h^\top \left(v_j \hat{A} e_i + (d_j h_i + c v_j s_i) \hat{A} g \right) = h_i \hat{B} e'_j - (\gamma v_j s_i + d_j h_i) \hat{B} u,$$

where

$$\gamma = 1 + c(h^\top \hat{A} g).$$

Finally,

$$\begin{aligned} x &= h_i \hat{B} e'_j - (d v_j s_i + d_j h_i) \hat{B} u, \\ y &= \ v_j \hat{A} e_i + (d_j h_i + c v_j s_i) \hat{A} g, \end{aligned} \tag{14}$$

and, using (11),

$$\begin{aligned} Z = e'_j e_i^\top \ - \ (v_j u)(\hat{A}(e_i + c s_i g))^\top - (\hat{B}(e'_j - d_j u))(h_i g)^\top \\ -(d_j u)(h_i \hat{A} g)^\top - (\gamma v_j \hat{B} u)(s_i g)^\top. \end{aligned} \tag{15}$$

Now we are in a position to apply Lemma 3.2. Introduce some new matrices as follows:

$$A' = \hat{A}(I + c g s^\top), \quad s^\top = [s_1 ... s_p],$$

$$B' = \hat{B}(I - u d^\top), \quad d^\top = [d_1 ... d_q].$$

Then, on the base of Lemma 3.2,

$$M^{-1} = I - A' \otimes u v^\top - g h^\top \otimes B' - \hat{A} g h^\top \otimes u d^\top - \gamma g s^\top \otimes \hat{B} u v^\top. \tag{16}$$

In the derivation we used nonsingularity of the coefficient matrix in (13). This is not an additional assumption, because it follows from the nonsingularity of M. Indeed, the coefficient matrix is nonsingular if the solutions x and y are defined uniquely. The latter is granted as Z is unique.

At the same time, we used an extra assumption, that the diagonal blocks $I + (g^\top h)B$ and $I + (u^\top v)A$ are nonsingular. Note that it is equivalent to nonsingularity of the matrices $K = I + A \otimes u v^\top$ and $L = I + g h^\top \otimes B$ (cf. Lemmas 2.1 and 2.2).

Thus, in this section we have produced another proof of the estimate

$$\text{tRank}(M^{-1}) \leq 5.$$

Moreover, it is given by constructions that show explicitly how all the matrices and vectors involved in (16) can be computed. In the case of unstructured matrices A and B the computational complexity is obviously $O(p^3 + q^3)$. Note also that for diagonal A and B the cost falls down to $O(p + q)$.

4 More results

Lemma 4.1 *Let the entries of all matrices belong to an arbitrary field* $\mathbb{F}$*. If a matrix of the form*

$$K = I + \sum_{i=1}^{r} A_i \otimes u_i v_i^\top \tag{17}$$

with $p \times p$ *matrices* A_i *and column vectors* u_i*,* v_i *with* q *entries is nonsingular, then*

$$K^{-1} = I + \sum_{i=1}^{r} \sum_{j=1}^{r} A_{ij} \otimes u_i v_j^\top \tag{18}$$

with some $p \times p$ *matrices* A_{ij}*.*

Proof. The given matrix K can be viewed as a rank pr update of the identity matrix:

$$K = I + \sum_{i=1}^{r} (A_i \otimes u_i)(I \otimes v_i^\top) = \begin{bmatrix} A_1 \otimes u_1 \ \dots \ A_r \otimes u_r \end{bmatrix} \begin{bmatrix} I \otimes v_1^\top \\ \dots \\ I \otimes v_r^\top \end{bmatrix}.$$

Then, the Sherman–Morrison–Woodbury formula yields

$$K^{-1} = I + \begin{bmatrix} A_1 \otimes u_1 \ \dots \ A_r \otimes u_r \end{bmatrix} \begin{bmatrix} W_{11} & \dots & W_{1r} \\ \dots & \dots & \dots \\ W_{r1} & \dots & W_{rr} \end{bmatrix} \begin{bmatrix} I \otimes v_1^\top \\ \dots \\ I \otimes v_r^\top \end{bmatrix},$$

where $W_{ij} \in \mathbb{F}^{p \times p}$. Along with the equalities

$$(A_i \otimes u_i) W_{ij} (I \otimes v_j^\top) = A_i W_{ij} \otimes u_i v_j^\top$$

it implies that

$$K^{-1} = I + \sum_{i=1}^{r} \sum_{j=1}^{r} (A_i \otimes u_i) W_{ij} (I \otimes v_j^\top) = \sum_{i=1}^{r} \sum_{j=1}^{r} (A_i W_{ij} \otimes u_i v_j^\top).$$

□

Lemma 4.2 *If a matrix of the form*

$$L = I + \sum_{i=1}^{r} g_i h_i^\top \otimes B_i \tag{19}$$

with $q \times q$ *matrices* B_i *and column vectors* g_i*,* h_i *with* p *entries is nonsingular,* *then*

$$L^{-1} = I + \sum_{i=1}^{r}\sum_{j=1}^{r} g_i h_j^\top \otimes B_{ij} \tag{20}$$

with some $q \times q$ matrices B_{ij}.

Proof. We adduce the argument of the proof of Lemma 2.2. $\square$

Theorem 4.1 *Let a nonsingular matrix M be of the form*

$$M = I + \sum_{i=1}^{r} \left(A_i \otimes u_i v_i^\top + g_i h_i^\top \otimes B_i \right). \tag{21}$$

Assume additionally that the corresponding matrices K and L defined by (17) and (19) are both nonsingular. Then

$$\mathrm{tRank}(M^{-1}) = O(r^{10}) \tag{22}$$

uniformly in p, q.

If $\mathbb{F} \subset \mathbb{C}$ then it holds notwithstanding the nonsingularity of K and L.

Proof. As in the proof of Theorem 2.1, observe that

$$\begin{aligned} KL - M &= \left(\sum_{i=1}^{r} A_i \otimes u_i v_i^\top\right)\left(\sum_{j=1}^{r} g_j h_j^\top \otimes B_j\right) \\ &= \sum_{i=1}^{r}\sum_{j=1}^{r} (A_i \otimes u_i v_i^\top)(g_j h_j^\top \otimes B_j) = \sum_{i=1}^{r}\sum_{j=1}^{r} (A_i g_j h_j^\top \otimes u_i v_i^\top B_j) \\ &= \sum_{i=1}^{r}\sum_{j=1}^{r} (A_i g_j \otimes u_i)(h_j^\top \otimes v_i^\top B_j). \end{aligned}$$

Thus, M is a rank r^2 update of the matrix KL. Therefore,

$$M^{-1} = L^{-1}\left(I + \sum_{i=1}^{r}\sum_{j=1}^{r}\sum_{k=1}^{r}\sum_{l=1}^{r} c_{ijkl} K^{-1}(A_i g_j \otimes u_i)(h_l^\top \otimes v_k^\top B_l) L^{-1}\right) K^{-1}$$

for some scalar coefficients $c_{ijkl} \in \mathbb{F}$. Allowing for (18), we obtain

$$
\begin{aligned}
K^{-1}(A_i g_j \otimes u_i) &= \left(I + \sum_{s=1}^{r}\sum_{t=1}^{r} A_{st} \otimes u_s v_t^\top\right)(A_i g_j \otimes u_i) \\
&= (A_i g_j \otimes u_i) + \sum_{s=1}^{r}\sum_{t=1}^{r} (v_t^\top u_i) A_{st} A_i g_j \otimes u_s \\
&= A_i g_j \otimes u_i + \sum_{s=1}^{r}\left(\sum_{t=1}^{r} (v_t^\top u_i) A_{st} A_i\right) g_j \otimes u_s \\
&= \sum_{s=1}^{r} \tilde{A}_{is} g_j \otimes u_s
\end{aligned}
$$

for properly defined matrices $\tilde{A}_{is}$. Similarly, using (20),

$$
\begin{aligned}
(h_l^\top \otimes v_k^\top B_l) L^{-1} &= (h_l^\top \otimes v_k^\top B_l)\left(I + \sum_{s=1}^{r}\sum_{t=1}^{r} g_s h_t^\top \otimes B_{st}\right) \\
&= (h_l^\top \otimes v_k^\top B_l) + \sum_{s=1}^{r}\sum_{t=1}^{r} h_t^\top \otimes (h_l^\top g_s) v_k^\top B_l B_{st} \\
&= (h_l^\top \otimes v_k^\top B_l) + \sum_{t=1}^{r} h_t^\top \otimes \left(\sum_{s=1}^{r} v_k^\top (h_l^\top g_s) B_l B_{st}\right) \\
&= \sum_{t=1}^{r} h_t^\top \otimes v_k^\top \tilde{B}_{lt}
\end{aligned}
$$

for some matrices $\tilde{B}_{lt}$. Putting these formulas together, we deduce that

$$
M^{-1} = L^{-1}\left(I + \sum_{i=1}^{r}\sum_{j=1}^{r}\sum_{k=1}^{r}\sum_{l=1}^{r}\sum_{s=1}^{r}\sum_{t=1}^{r} c_{ijkl} \tilde{A}_{is} g_j h_t^\top \otimes u_s v_k^\top \tilde{B}_{lt}\right) K^{-1},
$$

and it remains to take into account the formulas for K^{-1} and L^{-1}. $\square$

We should admit that the estimate of Theorem 4.1 is likely to be improved. A thorougher look at the algebra making the proof reveals that

$$
\mathrm{tRank}(M^{-1}) \le 1 + 2r^2 + r^4 + r^6 + 2r^8 + r^{10}.
$$

If $r = 1$ then it transforms into $\mathrm{tRank}(M^{-1}) \le 8$. This is certainly a rough estimate in this case, because Theorem 2.1 establishes a finer and sharp result $\mathrm{tRank}(M^{-1}) \le 5$.

5 Special cases

In some particular cases we manage to refine the estimate of Theorem 4.1. Pursueing the same line, we begin with a revision of Lemmas 4.1 and 4.2.

Lemma 5.1 *Let $A \in \mathbb{F}^{p\times p}$ be a diagonal matrix and $U, V \in \mathbb{F}^{q\times r}$ consist of the columns $u_1, ..., u_r$ and $v_1, ..., v_r$. If a matrix*

$$M = I + A \otimes UV^\top \tag{23}$$

is nonsingular, then for $p \geq r$

$$\mathrm{tRank}(M^{-1}) \leq 3r - 1. \tag{24}$$

More specifically,

$$M^{-1} = I + \sum_{|i-j|\leq 1,\ 1\leq i,j\leq r} \tilde{A}_{ij} \otimes u_i v_j^\top \tag{25}$$

with some $p \times p$ matrices $\tilde{A}_{ij}$.

Proof. As previously, we make use of viewing M as a rank pr update of the identity:

$$M = I + A \otimes UV^\top = I + (A \otimes U)(I \otimes V^\top).$$

Then

$$M^{-1} = I - (A \otimes U)W^{-1}(I \otimes V^\top), \tag{26}$$

where

$$W = I + (I \otimes V^\top)(A \otimes U) = I + A \otimes V^\top U.$$

Since A is diagonal, W turns out to be block diagonal with p blocks of order r. Obviously, W^{-1} inherits the same structure. Now, let us consider W^{-1} as a block matrix with $r \times r$ blocks of order p:

$$W^{-1} = \begin{bmatrix} \hat{W}_{11} & ... & \hat{W}_{1r} \\ ... & ... & ... \\ \hat{W}_{r1} & ... & \hat{W}_{rr} \end{bmatrix}.$$

If we take up indices i, j in the range $1 \leq i, j \leq r$ so that $|i - j| > 1$, then $\hat{W}_{ij} = 0$ by virtue of the condition $p \geq r$. From (26),

$$M^{-1} = I - \sum_{i=1}^{r}\sum_{j=1}^{r}(A \otimes u_i)\hat{W}_{ij}(I \otimes v_j^\top), = I - \sum_{i=1}^{r}\sum_{j=1}^{r} A\hat{W}_{ij} \otimes u_i v_j^\top,$$

and we can discard the terms with zero blocks $\hat{W}_{ij}$. $\square$

A similar couterpart of Lemma 4.2 evidently reads as follows.

Lemma 5.2 *Let $B \in \mathbb{F}^{q\times q}$ be a diagonal matrix and $G, H \in \mathbb{F}^{p\times r}$ consist of the columns $g_1, ..., g_r$ and $h_1, ..., h_r$. If a matrix*

$$M = I + GH^\top \otimes B \tag{27}$$

is nonsingular, then for $q \geq r$

$$\text{tRank}(M^{-1}) \leq 3r - 1. \tag{28}$$

More specifically,

$$M^{-1} = I + \sum_{|i-j|\leq 1,\ 1\leq i,j\leq r} g_i h_j^\top \otimes \tilde{B}_{ij} \tag{29}$$

with some $q \times q$ matrices $\tilde{B}_{ij}$.

Theorem 5.1 *Let a nonsingular matrix M be of the form*

$$M = A \otimes UV^\top + GH^\top \otimes B, \tag{30}$$

where the matrices $A \in \mathbb{F}^{p\times p}$ and $B \in \mathbb{F}^{q\times q}$ are diagonal, $U, V \in \mathbb{F}^{q\times r}$ and $G, H \in \mathbb{F}^{p\times r}$. Assume additionally that the matrices $K = I + A \otimes UV^\top$ and $L = I + GH^\top \otimes B$ are both nonsingular. If $p, q \geq r$ then

$$\text{tRank}(M^{-1}) = O(r^6) \tag{31}$$

uniformly in p, q.

If $\mathbb{F} \subset \mathbb{C}$, then the diagonality of A, B and nonsingularity of K, L can be dismissed.

Proof. As before, M is a rank r^2 modification of KL:

$$KL = M + AGH^\top \otimes UV^\top B = M + (AG \otimes U)(H^\top \otimes V^\top B) \equiv M + R.$$

Hence,

$$M^{-1} = L^{-1} \left(I - K^{-1}RL^{-1}\right)^{-1} K^{-1},$$

where, with some $r \times r$ matrices Z_{ij},

$$\begin{aligned}\left(I - K^{-1}RL^{-1}\right)^{-1} &= I + K^{-1} \left(\sum_{i=1}^{r}\sum_{j=1}^{r} (Ag_i \otimes U) Z_{ij} (h_j^\top \otimes V^\top B) \right) L^{-1} \\ &= I + K^{-1} \left(\sum_{i=1}^{r}\sum_{j=1}^{r} (Ag_i \otimes UZ_{ij}) (h_j^\top \otimes V^\top B) \right) L^{-1}\end{aligned}$$

By Lemmas 5.1 and 5.2 the tensor ranks of K^{-1} and L^{-1} behave as $O(r)$. Consequently,

$$\text{tRank}\left((I - K^{-1}RL^{-1})^{-1} \right) = O(r^4),$$

and, by the appearance of M^{-1}, its tensor rank is $O(r^6)$.

If $\mathbb{F} \subset \mathbb{C}$ then A and B become diagonalizable by arbitrarily small perturbations providing as well the nonsinguilarity of the correspondingly perturbed matrices K and L. $\square$

6 A mixed format

Let us revisit somewhat the claim and proof of Theorem 4.1. It states that a general class of nonsingular matrices of the form

$$M = I + \sum_{i=1}^{r} \left(A_i \otimes u_i v_i^\top + g_i h_i^\top \otimes B_i \right)$$

is closed under inversion, in the sense that the tensor ranks of M and M^{-1} are bounded from above independetly of size.

The estimate (22), all the same, manifests a certain increase of the tensor rank for the inverse matrix. It seems it may grow even considerably (although it might be an artefact of the proof technique). However, when seeking a numerically viable structure in the inverse matrix we need not confine ourselves to the tensor format solely. Moreover, a good format is already suggested by the proof of Theorem 4.1.

Theorem 6.1 *Under the hypotheses of Theorem 4.1 the inverse matrix can be written in the form*

$$M^{-1} = L^{-1} K^{-1} + R, \tag{32}$$

where K and L are defined by (17) and (19). The matrices K^{-1} and L^{-1} are of low tensor rank:

$$\text{tRank}(K^{-1}) \le r^2 + 1, \qquad \text{tRank}(L^{-1}) \le r^2 + 1,$$

while R is of low classical rank:

$$\text{rank} R \le r^2.$$

Proof. Its suffices to take into account that

$$L^{-1} K^{-1} - M^{-1} = (KL)^{-1} - M^{-1} = -(KL)^{-1} (KL - M) M^{-1}$$

and use the low-rank expression for $KL - M$ in the proof of Theorem 4.1. $\square$

The right-hand side of (32) involves low-tensor-rank and low-rank matrices and can be referred to as a mixed format. Anyway it provides a structured low-parametric representation for the inverse of M with as small as $O(r^2)$ parameters.

Another natural suggestion is to use $L^{-1}K^{-1}$ as an explicit preconditioner for M in iterative processes. It is clear that

$$M(L^{-1}K^{-1}) = I + \hat{R}, \qquad \text{rank}\hat{R} \leq r^2.$$

Since the rank of $\hat{R}$ is bounded uniformly in size, we should regard $L^{-1}K^{-1}$ as a so-called "superlinear precondioner".

References

[1] D. Bini and V. Pan, *Matrix and polynomial computations*, Birkhauser, Boston (1994).

[2] L. De Lathauwer, B. De Moor, J. Vandewalle, A multilinear singular value decomposition, *SIAM J. Matrix Anal. Appl.*, Vol. 21, no. 4, pp. 1253–1278 (2000).

[3] J. M. Ford, E. E. Tyrtyshnikov, Combining Kronecker product approximation with discrete wavelet transforms to solve dense, function-related systems, *SIAM J. Sci. Comp.*, Vol. 25, No. 3, 961–981, (2003).

[4] W. Hackbusch, B. N. Khoromskij and E. E. Tyrtyshnikov, Hierarchical Kronecker tensor-product approximations, *J. Numer. Math.*, **13**, 119–156 (2005).

[5] W. Hackbusch, B. N. Khoromskij, E. Tyrtyshnikov, Approximate iterations for structured matrices, *Numer. Math.*, accepted (2007).

[6] J. Kamm and J. G. Nagy, Optimal Kronecker Product Approximations of Block Toeplitz Matrices, *SIAM J. Matrix Anal. Appl.*, Vol. 22, No. 1, pp. 155-172 (2000).

[7] J. B. Kruskal, Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics, *Linear Algebra Appl.*, 18, pp. 95–138 (1977).

[8] L.-H. Lim, V. de Silva, *Tensor rank and the ill-posedness of the best low-rank approximation problem*, SCCM-06-06, Stanford, 2006.

[9] V. Olshevsky, I. Oseledets, E. Tyrtyshnikov, Tensor properties of multilevel Toeplitz and related matrices, *Linear Algebra Appl.* 412 (2006), 1–21.

[10] V. Olshevsky, I. Oseledets, E. Tyrtyshnikov, Superfast inversion of two-level Toeplitz matrices using Newton iteration and tensor-displacement structure, *Operator Theory Advances and Applications*, vol. 179, pp. 229–240 (2007).

[11] I. Oseledets, D. Savostianov, E. Tyrtyshnikov, Tucker dimensionality reduction of three-dimensional arrays in linear time, *SIMAX*, accepted (2007).

[12] I.Oseledets, E.Tyrtyshnikov, A unifying approach to construction of circulant preconditioners, *Linear Algebra Appl.* 418 (2006), 435-449.

[13] E. E. Tyrtyshnikov, Tensor approximations of matrices generated by asymptotically smooth functions, *Sbornik: Mathematics* **194**, No. 5-6, 941–954, (2003).

[14] E. E. Tyrtyshnikov, Kronecker-product approximations for some function-related matrices, *Linear Algebra Appl.*, 379, pp. 423–437, (2004).

[15] E. E. Tyrtyshnikov, Tensor ranks for inversion of tensor-product binomials, *Journal of Comput. and Appl. Math.*, submitted (2007).

[16] C. F. Van Loan, N. P. Pitsianis, Approximation with Kronecker products, *NATO Adv. Sci. Ser E Appl. Sci.* 232, Kluwer: Dordrecht, pp. 293–314 (1993).

[17] N.L.Zamarashkin, I.V.Oseledets, E.E.Tyrtyshnikov, Approximation of Toeplitz matrices by sums of circulants and small-rank matrices, *Doklady Mathematics*, Vol. 73, No. 1 (2006), 100–101.