Block Approximation of Tall Sparse Matrices and Block-Givens Rotations

Alexander Kushkuley  
(Salesforce, akushkuley@salesforce.com, kushkuley@gmail.com)  
July 26, 2017

Abstract

Estimation of top singular values is one of the widely used techniques and one of the intensively researched problems in Numerical Linear Algebra and Data Science. We consider here two general questions related to this problem:

How top singular values are affected by zeroing out a sparse rectangular block of a matrix?  
How much top singular values differ from top column norms of a tall sparse non-negative matrix?

AMS Subject classification: 15A18, 15B52, 15B10, 65F30, 65F50

1 Introduction

For a real matrix $X$ of rank $r$ its singular values are indexed in decreasing order, i.e. it is always assumed that

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$$

Everywhere below $\|X\|$ means an operator norm of a matrix $X$ and by $\sigma_i(X)$ we denote the $i$-th singular value of $X$ so that $\|X\| = \sigma_1(X)$. It is well known (cf. e.g. [1], [2]) that the optimal rank $k$ approximation $X_k$ of the a matrix $X$ is characterized by the property

$$\|X - X_k\| = \sigma_{k+1}(X)$$

Let

$$R = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

(1)

be a block partitioning of a real matrix $R$ such that $A$ is a square $k \times k$ matrix, $m \geq n$, $k < n$. Set also

$$R_0 = \begin{pmatrix} A & B \\ C & 0 \end{pmatrix}$$

(2)
Clearly \( \|R - R_0\| \leq \|D\| \). Let \( R_i \) be the optimal rank \( i \) approximation of \( R \) and let \( R_0_i \) be the optimal rank \( i \) approximation of \( R_0 \). According to a well known Weyl inequality (cf. e.g. [1], [2]), for any matrices \( X,Y \)

\[
\sigma_i(X + Y) - \sigma_i(X) \leq \sigma_i(Y), \quad i = 1,2, \ldots , n
\]

Setting here \( X \leftarrow R_0, \ Y \leftarrow D \) and using inequality (3) twice by interchanging \( R \) and \( R_0 \) we get

\[
\|R_0 - R_0\| - \|D\| \leq \|R - R_i\| \leq \|R_0 - R_0\| + \|D\|
\]

for \( i = 1,2, \ldots , n \). Noticing that

\[
\|R - R_0\| = \|R - D - R_0 + D\| \leq \|R_0 - R_0\| + \|D\|
\]

and interchanging here \( R \) and \( R_0 \) we get also the following estimate (cf. [3] or [8] Lemma 2)

\[
\|R - R_0\| - 2\|D\| \leq \|R - R_i\| \leq \|R - R_0\| + 2\|D\|
\]

In other words, the operator norm error of replacing rank \( i \) approximation of \( R \) with rank \( i \) approximation of \( R_0 \), is no greater than two operator norms of the removed \( m \times (n - k) \) block \( D \) and the estimates (4,5) do not even depend on the fact that \( D \) is a rectangular matrix block. Further, since rank of \( R_0 \) is no greater than \( 2k \) it follows from (4) that

\[
\sigma_{i+1}(R) \leq \|D\| \quad \text{if} \quad i \geq 2k
\]

and it is quite obvious that

\[
\|R_0 - R_{0k}\| \leq \min\{\|B\|, \|C\|\}
\]

\[
\|R - R_k\| \leq \min\{\|B\|, \|C\|\} + \|D\|
\]

For an arbitrary matrix \( X \), let \( Q \equiv Q(X), Q' \equiv Q'(X) \) be its left and right orthogonal SVD multipliers. In other words

\[
Q R Q' = \Sigma, \quad Q \equiv Q(R) = \begin{pmatrix} c_1 & s_1 \\ s_2 & c_2 \end{pmatrix}, \quad Q' \equiv Q'(R) = \begin{pmatrix} c_1' & s_1' \\ s_2' & c_2' \end{pmatrix}
\]

where \( \Sigma \equiv \Sigma(R) \) is a diagonal matrix of downward ordered singular values of \( R \) and it is assumed that matrix blocks \( c_i \equiv c_i(R), s_i \equiv s_i(R), c_i' \equiv c_i'(R), s_i' \equiv s_i'(R), i = 1,2 \) are dimension-compatible with the block structure of \( R \) in (1).

Let \((X,Y)\) be a partitioning of \( \Sigma(R) = Q R Q' \) into \( m \times i \) and \( m \times (n - i) \) vertical matrix bands. Let also \((X_0,Y_0)\) be a similarly shaped partitioning of \( Q R_0 Q' \).

Clearly

\[
\|R_0 - R_0\| \leq \|Y_0\| = \|Y + (Y_0 - Y)\| \leq \|Y - Y_0\| + \|Y\|
\]

where \( Y = R - R_i \) by definition of singular value decomposition. Further, \((X - X_0,Y - Y_0)\equiv Q R Q' - Q R_0 Q' \equiv Q D Q' \) and therefore \( Y - Y_0 \) is just the rightmost \( m \times (n - i) \) vertical band of the matrix

\[
Q D Q' \equiv \begin{pmatrix} s_1 D s_2' \\ c_2 D s_2' \end{pmatrix}, \quad Q \equiv \begin{pmatrix} s_1 & c_1' \\ c_2 & c_2' \end{pmatrix}, \quad D \equiv \begin{pmatrix} s_1 D c_2' \\ c_2 D c_2' \end{pmatrix}
\]

\[
\|Q D Q'\|
\]
It follows then, that
\[
\|R_0 - R_0\| \leq |R - R_i| + \|D(Q(R))[k + 1 : n, i + 1 : n]|\| \tag{12}
\]
and similar estimate is obtained by replacing column partitions with row partitions
\[
\|R_0 - R_0\| \leq |R - R_i| + \|Q(R)[i + 1 : m, k + 1 : m]|D\|
\tag{13}
\]
where "submatrix designation" notation (cf. [2] p. 27) is used on the right hand sides of (12,13). To get opposite direction inequalities, interchange \( R \) and \( R_0 \) in (13)
\[
\|R - R_i\| \leq \|R_0 - R_0\| + \|Q(R_0)[i + 1 : m, k + 1 : m]|D\|
\tag{14}
\]

Setting
\[
\mu_{i,k}(R) = \min\{\|D(Q(R))[k + 1 : n, i + 1 : n]|\|, \|Q(R)[i + 1 : m, k + 1 : m]|D\| \}
\tag{15}
\]
for \( i = 1, \ldots, n \) and
\[
c_2(R_0,i) = c_2(R_0)[i - k : m - k, 1 : m - k], c'_2(R_0,i) = c'_2(R_0)[1 : n - k, i - k : n - k]
\tag{16}
\]
\[
c_2(R,i) = c_2(R)[i - k : k, 1 : m - k], c'_2(R,i) = c'_2(R)[1 : n - k, i - k : n - k]
\tag{16}
\]
for \( i > k \) (cf. (9)), we obtain by the way of introduction

**Theorem 1.** For any \( i = 1, 2, \ldots, n \)
\[
\|\sigma_i(R) - \sigma_i(R_0)\| \leq \mu_{i,k}(R) \tag{17}
\]

In particular, if \( i > k \) then
\[
\|\sigma_i(R) - \sigma_i(R_0)\| \leq \max\{\min\{c_2(R,i)D\|, \|Dc'_2(R,i)\|\}, \min\{c_2(R,i)D\|, \|Dc'_2(R_0,i)\|\} \} \tag{18}
\]
and if \( i > 2k \) then
\[
|\sigma_i(R)| \leq \max\{\min\{c_2(R,i)D\|, \|Dc'_2(R,i)\|\}, \min\{c_2(R_0,i)D\|, \|Dc'_2(R_0,i)\|\} \} \tag{19}
\]

**Example 1.** Let
\[
r = \begin{pmatrix} c & -s \\ s & c \end{pmatrix}
\]
be a plane rotation matrix and let
\[
r_0 = \begin{pmatrix} c & -s \\ s & 0 \end{pmatrix}
\]
Then \( \sigma_2(r) = 1, \sigma_2(r_0) = \left( \frac{1 + s^2}{2} - \sqrt{\left(\frac{1 + s^2}{2}\right)^2 - s^4} \right)^{1/2}, \mu_{1,1}(r) = c \) and direct computation shows that \( \sigma_2(r) - \sigma_2(r_0) = c^2 \) in accordance with Theorem 1. Computational experiments show that max \( \{\mu_{i,k}(R), \mu_{i,k}(R_0)\} \) in (15) cannot be replaced with either \( \mu_{i,k}(R) \) or \( \mu_{i,k}(R_0) \).
Example 2. Let $M$ be a symmetric $p \times p$ matrix and let $M'$ be the same matrix with the corner element $M_{p,p} \neq 0$ zeroed out. Let $Q, Q'$ be orthogonal matrices that diagonalize $M$ and $M'$ respectively. Then by Theorem 1

$$|\sigma_p(M) - \sigma_p(M')| \leq |M_{p,p}| \max\{|Q_{p,p}, |, |Q'_{p,p}|\}$$

and therefore, an upper bound equality $|\sigma_p(M) - \sigma_p(M')| = |M_{p,p}|$ is reached if and only if $\sigma_p(M) = |M_{p,p}|$, $\sigma_p(M') = 0$.

In practice, estimates (17,18) are no more useful than (4) since computing numbers $\mu_{i,k}$ seems to be no easier than diagonalizing original matrix $R$. We will explore possibilities of improving estimates (7-8) and (17-19) in the next section. The idea is to block-diagonalize matrix $R$ using block-Givens rotations that can be written down in closed form. A case when the column norms of the matrix $R$ are rapidly decreasing is considered in section 3. We end the introduction with a few auxiliary lemmas.

1.1 Preliminaries

Lemma 1. Let $f$ be a reasonable non-negative monotone function defined on $[0, \infty)$ and let $X$ be a positive semi-definite symmetric $n \times n$ matrix of rank $r$

(i) If $f$ is non-decreasing then $\sigma_i(f(X)) = f(\sigma_i(X))$ for $i = 1, 2, \cdots, n$. In particular $\|f(X)\| = f(\|X\|)$

(ii) If $f$ is non-increasing then $\sigma_i(f(X)) = f(\sigma_{n-i+1}(X))$, $i = 1, 2, \cdots, n$. In particular, if $X$ is of maximal rank $(n = r)$ then $\|f(X)\| = f(\sigma_r(X))$

Lemma 2. Take two matrices $X$ and $Y$ of dimensions $(n-k) \times k$ and $(n-k) \times (n-k)$ and let $Q_1$ and $Q_2$ be matrices of dimensions $k \times (n-k)$ and $(n-k) \times (n-k)$. Then the norm of the "horizontally-stacked" $(n-k) \times (n-k)$ matrix $(XQ_1, YQ_2)$ is no greater than $\sqrt{\|Q_1Q_1^T + Q_2Q_2^T\| \max\{|X|, \|Y\|\}}$.

Proof. Let $\lambda = \max\{|X|, \|Y\|\}$. For any $x \in \mathbb{R}^{(n-k)}$ we have

$$< (XQ_1Q_1^TX + YQ_2Q_2^TY^T)x, x > = < Q_1Q_1^TX^T, x, x > + < Q_2Q_2^TY^T, x, x > \leq \lambda^2 < (Q_1Q_1^T + Q_2Q_2^T)x, x > \leq \lambda^2 \|Q_1Q_1^T + Q_2Q_2^T\|$$

Corollary 1. If under conditions of Lemma 2 the matrices $Q_1$ and $Q_2$ are such that $Q_1^TQ_2 = 0$ then

$$\|(XQ_1, XQ_2)\| \leq \sqrt{\max\{|Q_1Q_1^T\|, \|Q_2Q_2^T\|\} \max\{|X|, \|Y\|\}}$$

Indeed, in this case matrices $Q_1Q_1^T$ and $Q_2Q_2^T$ commute with each other and can be simultaneously diagonalized. Moreover, the resulting diagonal matrices will have no common non-zero elements.

Lemma 3. Under conditions of Lemma 3 chose non-negative numbers $j \leq k \leq i \leq n - k$ and let $P$ be a $(n-k) \times i$ matrix with its first $j$ columns being the same as the first $j$ columns of $X$ and the rest $i-j$ of its columns being the same as the columns indexed $k+j+1, \cdots, i$ of $Y$. Then $\|P\| \leq \max\{|X|, \|Y\|\}$. 

4
The lemma trivially follows from Corollary 1. Chose $Q_1$ to be a matrix with ones on its main diagonal in the first $j$ columns and zeros everywhere else. Select $Q_2$ in the same way so that its only nonzero elements are located on the main diagonal in the columns indexed by $k + 1, \cdots, i$.

For a symmetric matrix $S$ denote its $i$-th eigenvalue by $\lambda_i(S)$.

**Lemma 4.** (see [4] Theorem 2) If $S_1$ and $S_2$ are symmetric $p \times p$ matrices then $\lambda_i(S_1 + S_2) \geq \lambda_i(S_1) + \lambda_i(S_2)$. In particular, if $S_1$ and $S_2$ are positive semidefinite then $\sigma_i(S_1 + S_2) \geq \sigma_i(S_1) + \sigma_i(S_2)$

**Lemma 5.** For any square $n \times n$ matrix $M$ of rank less than $n$, there is a $QR$-decomposition $QX = M$ such that the last row of the matrix $X$ is zero.

Proof. Houholder QR algorithm (cf. e.g. [2]) applied to $M$ will encounter a zero diagonal entry at some point because $M$ is singular. Thereafter, the remaining Houholder transformations can be chosen in such a way that further diagonal entries will be eliminated.

Below we will use a "block decomposition" of a block partitioned orthogonal matrix. We will write $X = X_1 \oplus X_2 \oplus \cdots \oplus X_s$ if all the non-zero elements of a matrix $X$ are located in diagonal blocks $X_1, X_2, \ldots, X_s$. A unity $r \times r$ matrix will be denoted by $I_r$. As is well known, any $n \times n$ orthogonal matrix is orthogonally equivalent to a matrix that looks like $I_a \oplus -I_b \oplus \phi_1 \oplus \phi_2 \oplus \cdots \oplus \phi_{(n-a-b)/2}$ where $\phi_i, i = 1, 2, \cdots$ are plane rotations.

**Lemma 6.** (Block-Rotation Decomposition). Any $(n \times n)$ orthogonal matrix $Q$ block partitioned as in (9) can be factored as

$$\begin{pmatrix} q_1 & 0 \\ 0 & q_2 \end{pmatrix} \begin{pmatrix} I_r \oplus \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \oplus I_{n-l} \end{pmatrix} \begin{pmatrix} q_1' & 0 \\ 0 & q_2' \end{pmatrix}$$ (20)

where $q_1, q_1'$ are orthogonal $k \times k$ matrices, $q_2, q_2'$ are orthogonal $(n-k) \times (n-k)$ matrices, $r \leq k$, $l \leq n-k$, $k-r = n-k-l$ and $c$ and $s$ are $(k-r) \times (k-r)$ diagonal matrices, $\|c\| < 1$, $0 < \|s\| \leq 1$. Obviously, $c^2 + s^2$ must be a unity $(k-r) \times (k-r)$ matrix.

An almost obvious proof is sketched here for completeness. Diagonalize blocks $s_1, s_2$ in (9) using singular value decomposition. This decomposition will almost diagonalize blocks $c_1$ and $c_2$ as well with possible exception of scalar multiples of some square orthogonal sub-blocks along the main diagonal. It is easy to see that these sub-blocks must be symmetric and hence can be diagonalized by block-structure preserving orthogonal transformations as well. Finally, the signs of the diagonal entries of $c$ and $s$ can be adjusted by appropriately changing matrices $q_i, i = 1, 2$.

**Remark 1.** It is easy to see, that for a fixed block structure (9) the block-rotation decomposition (20) is unique in the same sense the singular value decomposition is. In particular, absolute values of the entries of diagonal blocks $c$ and $s$ are unique up-to a permutation.

Using factorization (20) and Remark 1 we can define a block partition "weight" $\omega(Q) \leq 1$ of a block structure of $Q$ in (9) as

$$\omega(Q) = \max\{\|c\|, \|s\|\}$$
Lemma 7. Let \( f \) be a continuous function such that \( f(0) = 1 \) and let \( X \) be a symmetric matrix. Let \( \nu = |f(X)|_{\text{Im},X} \). Then \( ||Yf(X)|| \leq \nu||Y|| + (1 - \nu)||Y|_{\text{ker},X}|| \) for any (appropriately sized) matrix \( Y \).

Proof. By diagonalizing \( X \) (and hence \( f(X) \)) we find a symmetric matrix \( M \) and an orthogonal projector matrix \( P \) such that \( MP = PM = 0 \) and \( f(X) = M + P \). Hence \( ||M|| = ||\nu|| \) and \( ||Yf(X)|| = ||Y(M + \nu P) + (1 - \nu)YP|| \leq \nu||Y|| + (1 - \nu)||Y|_{\text{ker},X}|| \).

Recall (cf. e.g. [1]) that column (\( ||\cdot||_1 \)) and row (\( ||\cdot||_{\infty} \)) matrix norms are defined as

\[
\max\{\text{column-wise sums of absolute values of matrix entries}\} \\
\max\{\text{row-wise sums of absolute values of matrix entries}\}
\]

The following fact that is sometimes called "Schur Test Inequality" (see [1], [5]) is well known

Lemma 8. \( ||X|| \leq \sqrt{||X||_{\text{\infty}}||X||_1} \) for any matrix \( X \). In particular, if \( X \) is symmetric then \( ||X|| \leq ||X||_{\text{\infty}} = ||X||_1 \).

2 Block-Givens Rotations

Notational conventions that were introduced in section 1 will be retained till the end of this section. It is assumed in addition that the block \( A \) of the matrix \( R \) in (1) is invertible. Let

\[
Q = \begin{pmatrix} c_1 & s_1 \\ s_2 & c_2 \end{pmatrix}
\]

be an orthogonal \( n \times n \) block matrix with \( k \times k \) and \( (n - k) \times (n - k) \) diagonal blocks \( c_1, c_2 \) and \( k \times (n - k), (n - k) \times k \) off-diagonal blocks \( s_1, s_2 \). We have

\[
s_1^T s_1 + c_2^T c_2 = I_{n-k}
\]

and

\[
RQ = \begin{pmatrix} Ac_1 + Bs_2 & As_1 + Bc_2 \\ Cc_1 + Ds_2 & Cs_1 + Dc_2 \end{pmatrix}
\]

It’s easy to see that \( Q \) can be chosen in such a way that

\[
As_1 + Bc_2 = 0
\]

Lemma 9. If \( As_1 + Bc_2 = 0 \) then \( c_2 \) is invertible.

Indeed, assuming that \( c_2 \) is degenerate, take its QR decomposition \( c_2 = xq \) that satisfies conditions of Lemma 5, so that \( x \) is a matrix with zero last column. Multiplying (23) by \( q^T \) on the right, we see that the last column of \( As_1q^T \) must be a zero column, hence last column of \( s_1q^T \) is a zero column as well because \( A \) is invertible. Therefore the same is true for the last column of the matrix

\[
q_1 = \begin{pmatrix} s_1 \\ c_2 \end{pmatrix}
\]

and that is, impossible because \( q_1 \) is orthogonal.
Lemma 10. If \( As_1 + Bc_2 = 0 \) then there is an orthogonal matrix \( q \) such that

\[
c_2 = (1 + B^T(AA^T)^{-1}B)^{-\frac{1}{2}} q
\]

Proof. Solving (23) for \( s_1 \) ( \( s_1 = -A^{-1}Bc_2 \) ) we get

\[
s_1^T s_1 = c_2^T(A^{-1}B)^T(A^{-1}B)c_2
\]

It follows then from (22) that

\[
c_2^Tc_2 = I_{n-k} - c_2^T(A^{-1}B)^T(A^{-1}B)c_2
\]

Using Lemma 9, multiply by \( c_2^{-1} \) on the right and by \((c_2^T)^{-1}\) on the left, getting

\[
c_2c_2^T = (I_{n-k} + (A^{-1}B)^T(A^{-1}B))^{-1}
\]

Applying Lemma 9 once again, write polar decomposition of \( c_2 \) as \((c_2c_2^T)^{\frac{1}{2}}q\) for some orthogonal \( q \) thus obtaining (24).

2.1 Block-partitioned Orthogonal Matrices

With a little bit of extra work we can write down a generalized Givens rotation matrix of the shape (21) in almost closed form. Let \( b_0q_B = A^{-1}B \) be a polar decomposition of \( A^{-1}B \) so that \( b_0 \) is a \( k \times k \) positive semi-definite matrix and \( q_B \) is an orthogonal \((n-k) \times (n-k)\) matrix such that \( q_B^T b_0^2 q_B = B^T(AA^T)^{-1}B \). Similarly, let \( q_C^T c_0 = CA^{-1} \) be a polar decomposition of \( CA^{-1} \). The following corollary is obtained by straightforward calculations.

Corollary 2. The matrix

\[
\begin{pmatrix}
(1 + b_0^2)^{-\frac{1}{2}} & -A^{-1}B (1 + B^T(AA^T)^{-1}B)^{-\frac{1}{2}} \\
(1 + B^T(AA^T)^{-1}B)^{-\frac{1}{2}} (A^{-1}B)^T & (1 + B^T(AA^T)^{-1}B)^{-\frac{1}{2}}
\end{pmatrix}
\]

is orthogonal and when applied on the right, annihilates top right corner of \( R \).

The matrix in Corollary 2 is a direct analogue of a two-dimensional Givens rotation (cf. e.g [2]). Since it depends on \( A^{-1}B \) only, we can set by definition

\[
b_0q_B = A^{-1}B \\
\cos(A, B) := (1 + b_0^2)^{-\frac{1}{2}}
\]

\[
\cos(B, A) := (1 + B^T(AA^T)^{-1}B)^{-\frac{1}{2}} \equiv q_B^T (1 + b_0^2)^{-\frac{1}{2}} q_B
\]

\[
\sin(A, B) := A^{-1}B (1 + B^T(AA^T)^{-1}B)^{-\frac{1}{2}} \equiv b_0 (1 + b_0^2)^{-\frac{1}{2}} q_B
\]

and similarly

\[
q_C^T c_0 = CA^{-1}
\]

\[
\cos(A, C) := (1 + c_0^2)^{-\frac{1}{2}}
\]

\[
\cos(C, A) := (1 + C(A^T A)^{-1}C^T)^{-\frac{1}{2}} \equiv q_C^T (1 + c_0^2)^{-\frac{1}{2}} q_C
\]

\[
\sin(C, A) := (CA^{-1})^T (1 + C(A^T A)^{-1}C^T)^{-\frac{1}{2}} \equiv c_0 (1 + c_0^2)^{-\frac{1}{2}} q_C
\]
It is easy to check that
\[
-A \sin(A, B) + B \cos(B, A) = 0 \\
-\sin(C, A)^T A + \cos(C, A) C = 0
\] (25)
and therefore, we have

**Corollary 3. Matrices**

\[
G_R := \begin{pmatrix}
\cos(A, B) & -\sin(A, B) \\
\sin(A, B)^T & \cos(B, A)
\end{pmatrix}
\]

\[
G_L := \begin{pmatrix}
\cos(A, C) & \sin(C, A) \\
-\sin(C, A)^T & \cos(C, A)
\end{pmatrix}
\]

are orthogonal and satisfy equations (25). More precisely, there are orthogonal \(k \times k\) matrices \(q_1, q_2\) such that

\[
G_L R = \begin{pmatrix}
q_1 \sqrt{A^T A + C^T C} & \cos(A, C) B + \sin(C, A) D \\
0 & -\sin(C, A)^T B + \cos(C, A) D
\end{pmatrix}
\]

\[
R G_R = \begin{pmatrix}
\sqrt{A A^T + C C^T} q_2 & 0 \\
C \cos(A, B) + D \sin(A, B)^T & -C \sin(A, B) + D \cos(B, A)
\end{pmatrix}
\]

For block-rotation decompositions (cf. (20)) of block-Givens rotations \(G_R\) and \(G_L\) one has immediately

**Corollary 4. If \(A\) is invertible and \(B \neq 0, C \neq 0\) then**

\[
\omega(G_R) = \max \left\{ \frac{1}{\sqrt{1 + \sigma_r(A^{-1} B)^2}}, \frac{\sigma_1(A^{-1} B)}{\sqrt{1 + \sigma_1(A^{-1} B)^2}} \right\} < 1
\]

\[
\omega(G_L) = \max \left\{ \frac{1}{\sqrt{1 + \sigma_l(C A^{-1})^2}}, \frac{\sigma_1(C A^{-1})}{\sqrt{1 + \sigma_1(C A^{-1})^2}} \right\} < 1
\]

where \(r = \text{rank}(B), \ l = \text{rank}(C)\).

**Example 3.** A Housholder transformation (cf. e.g. [2]) can be viewed as a particular case of a block-Givens rotation. Let \(v\) be a column vector and let \(a \in \mathbb{R}, a \neq 0\). It follows from Corollary 3 that a closed form for a Housholder matrix that "rotates" vector \(\begin{pmatrix} a \\ v \end{pmatrix}\) into \((\sqrt{a^2 + \|v\|^2}, 0, \ldots, 0)^T\) is

\[
\begin{pmatrix}
(1 + v^T v/a^2)^{-\frac{1}{2}} & (v/a)^T (1 + (v/a)(v/a)^T)^{-\frac{1}{2}} \\
- (1 + (v/a)(v/a)^T)^{-\frac{1}{2}} (v/a) & (1 + (v/a)(v/a)^T)^{-\frac{1}{2}}
\end{pmatrix}
\]

To verify this directly, note that quadratic form can be evaluated in any orthogonal coordinates, and in this case one can choose coordinate system in which \(v = (\|v\|, 0, \ldots, 0)^T\).
Set
\[
\nu_1 = \frac{1}{\sqrt{1 + \sigma_r(A^{-1}B)^2}} \\
\nu_2 = \frac{1}{\sqrt{1 + \sigma_l(CA^{-1})^2}} \\
\rho_1 = \nu_1 \|D\| + (1 - \nu_1)\|D\|_{\ker B} \\
\rho_2 = \nu_2 \|D\| + (1 - \nu_2)\|D\|_{\ker C^T}
\]

Applying block-Givens rotations of Corollary 3 to matrices \(R_0\), \(R\) and using lemmas 1 and 7 we improve estimates (7-8) as follows

**Theorem 2.** If matrix \(A\) in (1) is invertible, then
\[
\sigma_{k+1}(R_0) \leq \min \left\{ \frac{\|CA^{-1}\|}{\sqrt{1 + \|CA^{-1}\|^2}} \|B\|, \frac{\|A^{-1}B\|}{\sqrt{1 + \|A^{-1}B\|^2}} \|C\| \right\} \leq \frac{\|B\|\|C\|}{\sqrt{(\sigma_k(A)^2 + \max\{\|B\|, \|C\|\})^2}} \\
\sigma_{k+1}(R) \leq \min \left\{ \frac{\|CA^{-1}\|}{\sqrt{1 + \|CA^{-1}\|^2}} \|B\| + \rho_1, \frac{\|A^{-1}B\|}{\sqrt{1 + \|A^{-1}B\|^2}} \|C\| + \rho_2 \right\}
\]

**Corollary 5.** If \(n = m = 2k\) and matrices \(A\), \(B\) and \(C\) are invertible, then
\[
\sigma_{k+1}(R) \leq \min \left\{ \frac{\|B\|\|C\|}{\sqrt{(\sigma_k(A)^2 + \|B\|^2)}}, \frac{\|A\|\|D\|}{\sqrt{(\|A\|^2 + \sigma_k(B)^2)}}, \frac{\|B\|\|C\|}{\sqrt{(\sigma_k(A)^2 + \|C\|^2)}}, \frac{\|A\|\|D\|}{\sqrt{\|A\|^2 + \sigma_k(C)^2}} \right\}
\]

### 2.2 Block Diagonalization

Another straightforward application of Corollary 3 is a block-diagonalization algorithm

**Algorithm 1.** A matrix \(R\), block-partitioned as in (1) can be block diagonalized by a following iterative procedure.

1) Eliminate bottom left \(C\)-block of \(R\) by computing \(R_1 = G_L R\)

2) Compute matrix \(R_2\) by applying right block-Givens rotation that eliminates top-right \(B\)-block of \(R_1\)

3) Repeat the above steps for \(t = 2, 3, \cdots\) alternatively multiplying \(R_t\) by a left block-Givens rotation when \(t\) is even and multiplying \(R_t\) by a right block-Givens rotation when \(t\) is odd

By design
\[
R_t = \begin{pmatrix} A_t & B_t \\ 0 & D_t \end{pmatrix}, \text{ for odd } t \text{ and } R_t = \begin{pmatrix} A_t & 0 \\ C_t & D_t \end{pmatrix}, \text{ for even } t, \ t = 1, 2, \cdots
\]

Next lemma summarizes some useful properties of the block-diagonalization algorithm.

**Lemma 11.**

(i) \(\sigma_i(A_{t+1}) \geq \sigma_i(A_t)\) for all \(i = 1, 2, \cdots, k\) and \(t = 0, 1, \cdots\)
Proof. Statement (i) follows from Lemma 4, statements (ii) and (iii) and (vi) are obvious, statement (iv) can be verified by a direct computation and (v) follows from (iii), (iv) and Lemma 4.

Corollary 6. A matrix (or a vector) will be called non-negative if all its entries are non-negative. We will consider large sparse non-negative random matrices with not too different non-zero entries. The notions involved will be introduced step by step.

Example 4. It is well known that for any $p \times q$ matrix $Y$ with columns $v_1, \ldots, v_q$ and for any $i$, $1 \leq i \leq q$

$$\sum_{j=1}^{i} \sigma_j^2(Y) \geq \sum_{j=1}^{i} ||v_j||^2, \quad \sum_{j=i+1}^{q} \sigma_j^2(Y) \leq \sum_{j=i+1}^{q} ||v_j||^2$$

and one of the ways to see that is to use Algorithm 1

As a direct corollary of Lemma 11 we have

**Corollary 6.** If $\sigma_i(R[1 : m, 1 : k]) \geq ||R[1 : m, k + 1 : n]||$ for some $1 \leq i \leq k$ then first $i$ singular values of the matrix $\lim_t A_t$ coincide with first $i$ singular values of $R$

**Remark 2.** Suppose that block-diagonalization performed by Algorithm 1 is followed by a singular value decomposition of remaining diagonal blocks. Using the same notation as in Theorem 1 for the resulting orthogonal multipliers and their "c, s"-blocks we can state

**Theorem 1'.** For $i > k$ there are horizontal $(m - i + 1) \times (m - k)$ slices $S_i, S_0$ of $(s_1(R), c_1(R))$, $(s_1(R_0), c_2(R_0))$ and vertical $(n - k) \times (n - i + 1)$ slices $S_i', S_0'$ of $(s_2^{'}(R), c_2^{'}(R))$, $(s_2^{'}(R_0), c_2^{'}(R_0))$ such that

$$|\sigma_i(R) - \sigma_i(R_0)| \leq \max\{\min(||SD||, ||DS'||), \min(||S_0D||, ||DS_0'||)\}$$

**Sketch of the proof.** Repeat the proof of Theorem 1 and use Lemma 3.

3. **Singular Values of Sparse non-negative Random Matrices**

A matrix (or a vector) will be called non-negative if all its entries are non-negative. We will consider large sparse non-negative random matrices with not too different non-zero entries. The notions involved will be introduced step by step.
Matrix Density. Let $X$ be $m \times k$ non-negative matrix. For any row index subset $I' \subset \{1, 2, \ldots, m\}$ and any column index subset $J' \subset \{1, 2, \ldots, k\}$ the density of $|I'| \times |J'|$ sub-matrix $X[I', J']$ that is formed by intersection of rows indexed by $I'$ and columns indexed by $J'$ can be defined (cf. [11]) as

$$
\delta_{I', J'} = \delta(X[I', J']) = \frac{\sum_{i \in I', j \in J'} X_{ij}}{|I'||J'|}.
$$

(26)

We define the size $|x|$ of a non-negative vector $x = (x_1, \ldots, x_m)$, as a sum of its coordinates. For $I \equiv I_m = \{1, 2, \ldots, m\}$ and $J \equiv J_k = \{1, 2, \ldots, k\}$ we will use a shorthand $\delta = \delta(X[I_m, J_k]) = \delta_{I_m, J_k}(X) = \delta(X)$

Invariant Random Vectors. Let $G$ be a compact subgroup of orthogonal group $O(p)$. A random vector $x$ in $\mathbb{R}^p$ will be called $G$-invariant if for any $v \in \mathbb{R}^p$ and $g \in G$, $E(<gv, x>) = E(<v, x>)$. For example a uniformly distributed (on a sphere) random vector is $O(p)$ invariant.

Lemma 12. Let $V = \mathbb{R}^p$ be a real $p$-dimensional vector space with the standard Euclidean norm and let $G$ be a compact (e.g. finite) Lie-subgroup of the group of isometries $O(V)$. Suppose that natural representation of $G$ in $V$ does not have fixed points. Then

(i) If $z$ is a $G$-invariant random vector in $V$ then $E(L(z)) = 0$ for any linear function $L$ on $V$

(ii) $E(<x, y>) = 0$ for any pair of independent random $G$-invariant vectors $x, y$ in $V$

(iii) (cf. [12]) If in addition $G$-action on $V$ is absolutely irreducible, then for independent $G$-invariant random vectors $x, y$ on a unit sphere in $V$ we have

$$
E(<x, y>) = \frac{1}{p}
$$

Proof. The expectation vector $E(z) = (E(z_1), \ldots, E(z_n))$ must be $G$-invariant and (i) follows from the linearity of expectation: $E(L(z)) = L(E(z)) = 0$. Moreover, by independence of $x$ and $y$, $E(<x, y>) = E(<x(x), E(y)> = 0$ (e.g. by (i)).

We start the proof of (iii) by noticing that conditions and conclusions of the lemma do no depend on the choice of coordinates in $V$. Take an orthogonal basis $e_1, \ldots, e_p \in V$. An orbit of $e_1$ contains exactly $p$ linearly independent vectors $f_1 = e_1 = g_1 e_1, f_2 = g_2 e_1, \ldots, f_p = g_p e_1$ for some $g_1 = 1, g_2, \ldots, g_p \in G$ because $G$-action on $V$ is irreducible. Applying polar decomposition to a non-degenerate matrix $<e_i, f_j>$ if necessary, we can choose an orthogonal basis $e_1', e_2', \ldots, e_p'$ in such a way that $(p-1) \times (p-1)$ matrix $M = (<e_i', f_j>, i = 2, \ldots, p)$ is symmetric positive definite and therefore by Schur product theorem (see e.g [11]), its Hadamard square $M \circ M = (<e_i', f_j >^2, i, j = 2, \ldots, p)$ is non-degenerate. Therefore, we can fix the orthogonal basis $e_1, \ldots, e_p \in V$ in such a way that the matrix $<e_i, f_j >^2, i = 2, \ldots, p$ has rank $p-1$. 


Let \( x = \sum x_i e_i, \quad y = \sum y_i e_i \) so that

\[
<x, y>^2 = \sum_{i=1}^{p} x_i^2 y_i^2 + 2 \sum_{i \neq j}^{p} x_i x_j \sum_{i \neq j}^{p} y_i y_j
\]  

(27)

Since action of \( G \) on \( V \) is absolutely irreducible, the symmetric square \( S^2(V) \) of \( V \) splits into an orthogonal sum

\[
S^2(V) = V' \oplus W
\]

of one-dimensional trivial \( G \)-representation \( V' \equiv \mathbb{R} \cdot \sum e_i \otimes e_i \) and a fixed point free \( G \)-representation \( W \) (cf. e.g. [7]). It is easy to see that \( p(p-1)/2 \)-dimensional vector \((x_i x_j, i \neq j)\) can be viewed as a random \( G \)-invariant vector on \( W \) and therefore it follows from (ii) that expectation of the second summand in (38) is zero. As was already mentioned, the \( G \)-orbit of \( e_1 \) contains exactly \( p \) linearly independent vectors \( f_1 = e_1 = g_1 e_1, \quad f_2 = g_2 e_1, \quad \cdots, \quad f_p = g_p e_1 \), for some \( g_1, g_2, \cdots, g_p \in G \), and

\[
E(<x, e_1>^2) \equiv E(<x, f_1>^2) = E(<x, f_i>^2), \quad i = 2, \ldots, p
\]  

(28)

because random vector \( x \) is \( G \)-invariant. Opening brackets in (28) we get

\[
E(<x, f_j>^2) \equiv E(<x, f_j, f_j>^2) = \sum_{i=1}^{p} <e_i, f_j>^2 E(x_i^2) + \sum_{i \neq k}^{p} <e_i, f_j><e_k, f_j> E(x_i x_k), \quad j = 1, \cdots, p
\]  

(29)

The second term on the right hand side of (29) is equal to zero (by (i)) and hence \( E(x_i^2) \), \( i = 1, \cdots, p \) must satisfy a system of linear equation

\[
\sum_{i=1}^{p} <e_i, f_j>^2 E(x_i^2) = \lambda ( \equiv E(<x, f_j>^2)), \quad j = 1, \cdots, p
\]  

(30)

for some \( \lambda \in \mathbb{R} \). In particular, since \( \sum_{i=1}^{p} <e_i, f_j>^2 = 1 \), the system of \( p \) linear equations

\[
\sum_{i=1}^{p} <e_i, f_j>^2 E(x_i^2) = \frac{1}{p}, \quad j = 1, \cdots, p
\]  

(31)

has a solution

\[
E(x_i^2) = 1/p, \quad i = 1, \cdots, p
\]  

(32)

As was explained above, the rank of the system (31) is \( p - 1 \) and therefore ”the choice of expectations” (32) must be unique. It follows that the expectation of any squared coordinate of \( x \) (and of \( y \)) is equal to \( 1/p \) and using independence of \( x \) and \( y \) we get (cf. [12]):

\[
E(\sum_{i=1}^{n} x_i^2 y_i^2) = p \cdot \frac{1}{p} \cdot \frac{1}{p} = \frac{1}{p}
\]
Remark 3. The main idea behind Lemma 12 comes from [12] where statement (iii) is established for the case $G = O(V)$.

Remark 4. Obviously, the statement (iii) of lemma 12 can be "norm-scaled" for a pair of fixed norm random vectors $x, y$, i.e. $E(<x, y>) = \frac{1}{p}||x||^2||y||^2$. It is less obvious, but still easy to see that Lemma 12 can be generalized in two ways. First, a range of the random vector $x$ can be assumed to be a $G$-invariant subset of $V$ (or a subset of a (unit) sphere in $V$) since the proof presented above depends only on the properties of absolutely irreducible representations and invariance of random vectors. Second, the condition $||x|| = ||y|| = 1$ of Lemma 12 (iii) can be "replaced by Fubini’s theorem", so that

$$E(<x, y>) = \frac{1}{p}E(||x||^2)E(||y||^2)$$

for random (invariant) vectors with varying norms.

Non-negative vectors with fixed norms and sizes. An "end-point" of a uniformly distributed random nonnegative vector $x = (x_1, \cdots, x_m) \in \mathbb{R}^m$ with a fixed size $|x| = a$ belongs to a simplex $\sigma_x$ that is cut-off from the positive orthant by a hyperplane $H_x$ defined by an equation $\sum_{i=1}^m x_i = |x| \equiv a$. Take a normal to $H_x$ vector $\eta$ with coordinates $(1, \cdots, 1)$. We have $E(x) \equiv c_x = (|x|/m)\eta \equiv (a/m)\eta$. Vector $c_x$ is orthogonal to $H_x$, so that $x = c_x + r_x$ where $r_x \perp c_x$. Let’s fix the length (Euclidean norm) of $x$ as well, requiring that the endpoint of $x$ belongs to a sphere $S_x$ given by an equation $||x|| = b$, thus confining the range of $x$ to an intersection $S' = S_x \cap \sigma_x$.

Remark 5. It is easy to see that $S'$ is a sphere if and only if

$$||r_x|| = \sqrt{||x||^2 - |x|^2/m} \leq \sqrt{\frac{1}{m(m-1)}} = \text{the radius of a sphere inscribed in } \sigma_x$$

and in general, the set $S' \subset \sigma_x$ is non-empty as long as $||x|| \leq a = |x|$

Take a subgroup $O_\eta$ of orthogonal group $O(m)$ that leaves the vector $\eta$ fixed. The group $G$ of symmetries of the simplex ($\sigma_x$) as a subgroup of $O_\eta$ is isomorphic to a full symmetric group of coordinate permutations $\Gamma_m$. According to the remark above, although the set $S'$ is not necessarily $O_\eta$-invariant, it is nevertheless always $\Gamma_m$-invariant and hence restricting (truncating in a sense of multivariate distributions) a random vector $x$ uniformly distributed on a sphere $\{|x| = a, \|x\| = b\}$ to the intersection set $S'$ results in a $\Gamma_m$-invariant random vector in $S'$.

Lemma 13. Suppose that random vectors $x$ and $y$ are independent and $\Gamma_m$-invariant. If the sizes and norms of random vectors $x, y$ are fixed then

(a) $$E(<r_x, r_y>) = 0$$
$$E(<x, y>) = |x||y|/m$$

(33)
Proof. Statement (a) follows from the orthogonality of $c_x$ and $r_x$, Lemma 12 (i) and Remark 4. In more details

$$<x, y> = <c_x + r_x, c_y + r_y> = <c_x, c_y> + <r_x, r_y>$$

and therefore

$$E(<x, y>) = E(<c_x, c_y>) + E(<r_x, r_y>) = |x||y|/m$$

It is well known that the action of the symmetric group $\Gamma_m$ on the subspace orthogonal to $\eta$ is absolutely irreducible. Hence, the Statement (b) follows from Lemma 12 (ii-iii) and Remark 4:

$$\text{Var}(<x, y>) = E(<x, y>)^2 - E(<x, y>)^2 = <c_x, c_y>^2 +$$

$$2 <c_x, c_y> E(<r_x, r_y>) + E(<r_x, r_y>)^2 - <c_x, c_y>^2 =$$

$$E(<r_x, r_y>)^2 = \frac{1}{m-1} \left( \|x\|^2 - \frac{|x|^2}{m} \right) \left( \|y\|^2 - \frac{|y|^2}{m} \right)$$

Remark 6.

(i) As was already mentioned in Remark 4, Lemma 13 remains generally valid even if the sizes and norms of independent random vectors $x$, $y$ are not fixed. Sizes and norms on the right hand side of (a) and (b) can be replaced by expectations, for example, assuming only sizes of $x$, $y$ are fixed, we get under reasonable conditions

$$E(<r_x, r_y>)^2 = \frac{1}{m-1} \left( E(\|x\|^2) - \frac{|x|^2}{m} \right) \left( E(\|y\|^2) - \frac{|y|^2}{m} \right)$$

(ii) The statement (a) is essentially borrowed from [6]. A zero-one vector $u$ can be viewed as subset of a set its indexes. If $v$ is another zero-one vector of the same dimension and if $u$ and $v$ represent random independently selected subsets of the coordinate index set then (33) turns into (see [6])

$$E(<u, v>) = E(|u \cap v|) = \frac{|u||v|}{m}$$

Random Matrix with fixed column norms and sizes. With Lemma 13 in mind we now fix sizes $|u_1|, |u_2|, \cdots, |u_k|$ and norms $\|u_1\|, \|u_2\|, \cdots, \|u_k\|$ of all columns of a non-negative matrix $X$. In particular, the matrix density $\delta = (1/k) \sum_{i=1}^k (|u_i|/m)$ is thereby also fixed. Set

$$E = E_k = (\xi_1, \cdots, \xi_k)^T, \ \xi_i = |u_i|/\|u_i\|^2, \ i = 1, \cdots, k$$

$$U = U_k = (|u_1|, \cdots, |u_k|), \ \ D = D_k = \text{diag}(\|u_1\|^2, \cdots, \|u_k\|^2)$$

$$H = H_k = \text{diag} \left( 1 - \frac{|u_1|^2}{m|u_1|^2}, \cdots, 1 - \frac{|u_k|^2}{m|u_k|^2} \right) \quad (34)$$
Moment Ratio. Coefficient of variation of a random variable is defined as a ratio of its standard deviation to its mean. A ratio of the square root of the second moment of a random variable to its first moment will be somewhat loosely called below a moment ratio. The moment ratio \( \rho \) is related to the coefficient of variation \( \nu \) by a simple rule

\[
\rho^2 = \nu^2 + 1.
\]

Similarly, we can define a sample moment ratio, so that if \( \psi \) is a sample coefficient of variation of a positive sequence \( a = a_1, \ldots, a_t \), then

\[
\rho(a) = \frac{\sqrt{\frac{1}{t} \sum_{i=1}^{t} a_i^2}}{\sqrt{\frac{1}{t} \sum_{i=1}^{t} a_i}} = \sqrt{1 + \psi^2}, \quad \psi^2 = \rho^2 - 1
\]

(35)

Remark 7. We consider here only positive random variables and/or positive sequences that significantly differ from 0 (see below). Such a sequence has a small sample coefficient of variation if and only if a ratio of square root of its second moment to its first moment is close to 1.

So far we did not make any assumptions on sparsity or tallness of the matrix \( X \). Below we will assume that matrix \( X \) is "sparse enough" and that "size-to-square-of-the-norm" ratio vector \( E_k \) has small coefficients of variation. Set

\[
C = \max_{1 \leq i \leq k} |u_i|
\]

\[
L = \max_{1 \leq i \leq k} l_i \quad \text{where} \quad l_i > 0 \quad \text{is a number of non-zero entries in} \quad i\text{-th column}
\]

We assume the following set of conditions:

\((S_1)\)  
(i) \( \|u_i\|^2 \geq |u_i|, \; i = 1, \ldots, k \). In other words, as vectors in \( \mathbb{R}^m \) all the columns of \( X \) are located outside of the interior of a ball of radius \( \sqrt{m}/2 \) centered at a point \( (1/2, \cdots, 1/2) \)

(ii) \( C \leq m \)

(iii) The sample coefficient of variation of \( E_k \) is bounded by \( \sqrt{(1 + \frac{m}{Ck})^2 - 1} \) (cf. (35) above) or equivalently, the moment ratio of \( E_k \) is bounded by \( 1 + \frac{m}{Ck} \). In other words, consider the first moment and the square root of the second moment of the sequence \( \xi_i, \; i = 1, \ldots, k \)

\[
\Xi_1 \equiv \frac{1}{k} \sum_{i=1}^{k} \xi_i, \quad \Xi_2 \equiv \sqrt{\frac{1}{k} \sum_{i=1}^{k} \xi_i^2} \equiv \frac{1}{\sqrt{k}} \|E_k\|
\]

and require that

\[
\rho(E_k) = \Xi_2 / \Xi_1 \leq 1 + \frac{m}{Ck}
\]

(36)

Remark 8.

(a) All conditions in \((S_1)\) obviously hold for zero-one matrices (with \( C = L \)). In particular, for zero-one matrices \( \Xi_1 = \Xi_2 = \rho(E_k) = 1 \).

(b) A rough estimate of the right hand side of (36) is \( \frac{m}{Ck} < \frac{1}{k^2} \) where \( \delta \equiv \delta_k \) is the density of \( X \)
The condition (i) is satisfied if, for example, $\|u_i\|^2/l_i \geq 1$, $i = 1, \ldots, k$. Indeed,
\[ \xi_i = \frac{|u_i|}{\|u_i\|^2} = \frac{|u_i| \sqrt{l_i}}{l_i \|u_i\|} \]

(d) It follows from the condition (i) that at least one of the coordinates of $u_i$ is greater than 1.

Some immediate consequences of conditions $(S_1)$ are listed below

**Lemma 14.**

(a) $|u_i| \geq 1$, $\xi_i \leq 1$, $i = 1, \ldots, k$. In particular, $\|E_k\| \leq \sqrt{k}$ and $\Xi_1, \Xi_2 \leq 1$.

(b) $\frac{|u_i|^2}{m\|u_i\|^2} \leq L/m$, $i = 1, \ldots, k$.

(c) $\sum_{i=1}^{k} (\Xi_2 - \xi_i)|u_i| \leq m$

The statement (a) is obvious (see Remark 8 (d)). Inequality (b) follows from a standard inequality between arithmetic and quadratic means
\[ \frac{|u_i|^2}{m\|u_i\|^2} = \frac{|u_i|^2}{L^2} \frac{l_i}{\|u_i\|^2} m \]

Using (a) and condition $(S_1)$ (iii) we verify (c) as follows
\[ \sum_{i=1}^{k} (\Xi_2 - \xi_i)|u_i| \leq \sum_{i=1}^{k} (\Xi_2 - \xi_i)C \leq \frac{k\Xi_2 - \sum_{i=1}^{k} \xi_i}{\sum_{i=1}^{k} \xi_i} Ck \leq (1 + \frac{m}{Ck} - 1)Ck = m \]

Condition $(S_1)$ imposes restrictions on sizes and norms of columns $(u_1, \ldots, u_k)$ that constitute the matrix $X$. An additional, randomness condition imposed on $X$ is stated as follows

$(S_2)$ (i) The sizes and norms of the columns $(u_1, \ldots, u_k)$ of $X$ are fixed
\[ |u_i| = s_i, \quad \|u_i\| = n_i, \quad i = 1, \ldots, k \]

(ii) Columns of $X$ are non-negative independent random vectors invariant with respect to the group of permutations of coordinates $\Gamma_m$ (cf. Lemma 13).

**Remark 9.** It is easy to see that Lemma 13 is valid for random vectors that satisfy conditions $(S_1)$ and $(S_2)$. The condition $(S_2)$ (ii) is satisfied by uniformly distributed random vector, restricted (truncated as a multivariate random variable) to a simplex $|x| = \text{const}$ as was explained above.

### 3.1 Singular Values of the Expected Gram Matrix

Let now
\[ \rho \equiv \rho(X) = \frac{\sqrt{k} \sum_{i=1}^{k} |u_i|^2}{\frac{m}{k} \sum_{i=1}^{k} |u_i|^2} \]

be a sample moment ratio of the sequence of sizes of the columns of $X$ and let $G = E(X^T X)$ be an expected Gram matrix of $X$. 

Theorem 3. Suppose that random matrix $X$ satisfies conditions $(S_1)$ and $(S_2)$. Let $\tau$ be a transposition of the index set $\{1, \cdots, k\}$ that sorts the sequence of norms $||u_i||$ in descending order. Then for all $i = 1, \cdots, k$

\[(1 + k\delta \rho)||u_{\tau(i)}||^2 \geq \sigma_i(G) \geq (1 + \rho + O(L/m))^{-1}||u_{\tau(i)}||^2 \quad (38)\]

Proof. Let $M$ be a matrix with zero diagonal and with off-diagonal elements defined by (33), that is let $M_{i,j} = |u_i||u_j|/m$, $i, j = 1, 2, \cdots, k; i \neq j$. Using condition $(S_2)$ and definition (34), evaluate the expectation of $XX^T$ as

\[G = D(1 + D^{-1}M) = D \left( H + E_k \otimes \frac{U_k}{m} \right) \quad (39)\]

Let

\[Z = \left( H + E_k \otimes \frac{U_k}{m} \right)\]

The following two Lemmas essentially follow from conditions $(S_1)$.

Lemma 15.

\[\|Z^{-1}\| \leq 1 + \rho + O(L/m)\]

Proof. By Sherman-Morrison formula (see e.g. [14])

\[Z^{-1} = H^{-1} - \frac{(H^{-1}E_k \otimes \frac{U_k}{m}H^{-1})}{1 + \langle \frac{U_k}{m}, H^{-1}E_k \rangle}\]

We start by noticing that it follows from Lemma 14 (c) that

\[\Xi_k \sum_{i=1}^k \frac{|u_i|}{m} \leq 1 + \langle E_k, U_k/m \rangle \quad (40)\]

Computing up to the order of $O(L/m)$, using condition $(S_1)(ii)$, Lemma 14 (a),(b) and the fact that the matrices involved are of rank one, we get

\[\|Z^{-1}\| \leq 1 + \frac{\|E_k\|\|U_k/m\|}{1 + \langle E_k, U_k/m \rangle} + O(L/m) \leq 1 + \frac{\sqrt{E_k} \|U_k/m\|}{\sum_{i=1}^k |u_i|/m} + O(L/m)\]

where the last inequality follows from (40). Taking into account the definition of the moment ratio (37) we obtain the desired estimate

\[\|Z^{-1}\| \leq 1 + \frac{\sqrt{E_k} \sum_{i=1}^k |u_i|^2/m}{\sum_{i=1}^k |u_i|/m} + O(L/m) = 1 + \rho + O(L/m) \quad (41)\]

By the same token we have

Lemma 16.

\[\|Z\| \leq 1 + k\delta \rho\]
Indeed, according to (37) and Lemma 14 (a)
\[
\|Z\| \leq 1 + \frac{\sqrt{k} \|E_k\|}{m} \sqrt{\frac{1}{k} \sum_{i=1}^{k} |u_i|^2} = 1 + \frac{\sqrt{k} \|E_k\|}{m} \frac{\rho}{k} \sum_{i=1}^{k} |u_i| \leq 1 + \sqrt{k} \|E_k\| \delta \rho \leq 1 + k \delta \rho
\]

The proof of Theorem 3 is now a simple matter. To verify the second inequality, solve (39) for \(D\), apply Lemma 15 and use the well known multiplicative inequality for singular values \(\sigma_i(X_1 X_2) \leq \sigma_i(X_1) \sigma_1(X_2)\) (cf. e.g. [1]). The proof of the first inequality is almost the same with Lemma 16 being used instead of Lemma 15.

Remark 10. One can say that the expected Gramm matrix \(G\) satisfies a Restricted Isometry Condition in terminology of [9], [10].

Remark 11. To get some idea about the order of magnitude of the bounds (38), note that in most "practical" cases, \(L\) is much smaller than \(m\) and the moment ratio coefficient \(\rho\) is not significantly larger than 1 because column sizes of \(X\) are rapidly decreasing. For example, suppose that \(|u_i| = i, i = 1, \cdots, k\). Then \(\rho = 2/\sqrt{3} + O(1/k)\). On the other hand, Gershgorin radii of \(G\) can be roughly estimated as \(k \delta |u_i|, i = 1, \cdots, k\) where the average row size \(k \delta\) is usually significantly greater than 1.

3.2 Some Corollaries

In view of Example 4, the following statement is not surprising.

Corollary 7. In notation of Theorem 3
\[
\sum_{j=1}^{i} E(\sigma_j^2(X)) \geq \sum_{j=1}^{i} \sigma_j(G)
\]
\[
\sum_{j=i+1}^{k} E(\sigma_j(X))^2 \leq (1 + \rho + O(L/m)) \sum_{j=i+1}^{k} \sigma_j(G)
\]

for all \(i = 1, \cdots, k\).

Proof. The first statement does not depend on Theorem 3. Sum of the first \(i\) singular values (known as Ky Fan norm, cf. e.g. [1]) is a convex function, hence using Jensen’s inequality we get
\[
\sum_{j=1}^{i} E(\sigma_j^2(X)) = E \left( \sum_{j=1}^{i} \sigma_j(X^T X) \right) \geq \sum_{j=1}^{i} \sigma_j(E(X^T X))
\]

On the other hand, by Jensen’s inequality and Theorem 3 (cf. also Example 4)
\[
\sum_{j=i+1}^{k} E(\sigma_j(X))^2 \leq \sum_{j=i+1}^{k} E(\sigma_j^2(X)) \leq \sum_{j=i+1}^{k} \|u_T(j)\|^2 \leq (1 + \rho + O(L/m)) \sum_{j=i+1}^{k} \sigma_j(G)
\]
We will use standard notation \( \| \cdot \|_F \) for the Frobenius norm. Retaining conditions of Theorem 3 and notations preceding it, set
\[
\mathfrak{N} = \sum_{p=1}^{k} \| r_p \| \| X'_p \|_F
\]  
(42)
where \( X'_p \) is a \( m \times (k - 1) \) matrix with columns \( r_i \), \( 1 \leq i \leq k, i \neq p \). As was mentioned above (cf. Lemma 13), the norms \( \| r_i \|_2 = \| u_i \|_2 - \frac{|u_i|}{m}, i = 1, \ldots, k \) are fixed by conditions of Theorem 3.

**Corollary 8.** Let \( r_0 = \max_{1 \leq i \leq k} \| r_i \| \). For all \( i = 1, \ldots, k \)

(i) \[
\sigma_i(G) - \sqrt{\frac{k-1}{m-1}} \mathfrak{N} \leq E(\sigma_i^2(X)) \leq \sigma_i(G) + \sqrt{\frac{k-1}{m-1}} \mathfrak{N}
\]

(ii) \[
\sigma_i(G) - \frac{k-1}{\sqrt{m-1}} r_0^2 \leq E(\sigma_i^2(X)) \leq \sigma_i(G) + \frac{k-1}{\sqrt{m-1}} r_0^2
\]

(iii) (cf. [10]). In addition to conditions of Theorem 3, suppose that \( r_i, i = 1, \ldots, k \) are restrictions of uniformly distributed random vectors (see Remark 9). Then there is a constant \( c \) that does not depend on \( m, k \) and \( X \) such that with high probability
\[
\sigma_i(G) - c k r_0^2 \leq E(\sigma_i^2(X)) \leq \sigma_i(G) + c k r_0^2
\]

Proof.

(i) Write \( X^T X = G + M \) where \( M_{i,j} = < r_i, r_j > \) if \( i \neq j \) and \( M_{i,i} = 0 \) for \( i, j = 1, \ldots, k \) as in Lemma 13 and let \( t_p = \sum_{i \neq p} \| r_i \| \), \( p = 1, \ldots, k \). By (3) and Lemma 8
\[
E(\sigma_i^2(X)) = E(\sigma_i(X^T X)) \leq \sigma_i(G) + E(\| M \|) \leq E(\max_p t_p) \leq \sum_p E(t_p)
\]
By Lemma 13
\[
E(t_p) \leq \| r_p \| \sqrt{\frac{k-1}{m-1}} \left( \frac{1}{k-1} \sum_{i \neq p} \| r_i \|^2 \right) = \sqrt{\frac{k-1}{m-1}} \| r_p \| \sqrt{\sum_{i \neq p} \| r_i \|^2}
\]
and (i) follows now for example, from Lemma 4.

(ii) Let \( D' \) be a diagonal matrix with diagonal entries \( \| r_1 \|, \ldots, \| r_k \| \) and let \( M = D'M'D' \). As in the proof of (i) (and by Lemma 13)
\[
E(\sigma_i^2(X)) \leq \sigma_i(G) + \| D' \| E(\| M' \|) \leq \sqrt{\frac{1}{m-1}} E(\max_p t'_p) r_0^2
\]
where now \( \| t'_p \| \leq k - 1 \).
(iii) We follow [10] (Corollary 4.4.8). Let $T$ be a lower-triangular $(k-1)\times(k-1)$ sub-matrix of $M$, that is let $T_{i,j} = M_{i,j}$ if $i > j$ and $T_{i,j} = 0$ if $i \leq j$. Write again $T = D'T'D'$. It is easy to see that matrix entries $T'_{ij}$ of $T'$ are independent zero mean random variables with norms bounded by $\frac{k-1}{m-1}$ (Lemma 13). By our assumptions, matrix elements $T'_{ij}$ are distributed as truncated sub-gaussian random variables (see [10]) and hence are themselves sub-gaussian. Therefore, (cf. [10], Corollary 4.4.8 for details), we have

$$\sigma_i(G) + E(\|M\|) = \sigma_i(G) + 2E(T) \leq \sigma_i(G) + \sqrt{k-1} \frac{1}{m-1} \epsilon r_0^2$$

Finally, combining Corollary 8 with Theorem 3 (iii) we state

**Corollary 9.** Under conditions of Corollary 8 (iii)

$$(1 + k \delta + O(L/m)) \|u_{\tau(i)}\|^2 + \epsilon \sqrt{\frac{k-1}{m-1} r_0^2} \geq E(\sigma_i^2(X)) \geq (1 + \rho + O(L/m))^{-1} \|u_{\tau(i)}\|^2 - \epsilon \sqrt{\frac{k-1}{m-1} r_0^2}$$

### 3.3 A Generalization

Let’s relax the condition $(S_2)(i)$ by allowing column norms of $X_k$ to vary, retaining, however, the requirement for column sizes to be fixed. In other word, suppose that the following modification of the condition $(S_2)$ is satisfied

$(S'_2)$

(i) The sizes $(u_1, \ldots, u_k)$ of $X$ are fixed

$$|u_i| = s_i, \ i = 1, \ldots, k$$

(ii) Columns of $X$ are non-negative independent random vectors invariant with respect to the group of permutations of coordinates $\Gamma_m$ (cf. Lemma 13).

This change in condition $(S'_2)$ leads naturally to a change in a definition of the numbers $\xi_i$ so that definitions (34) are replaced by

$$\xi_i = |u_i|/E(|u_i|^2), \ i = 1, \ldots, k; \ E_k = (\xi_1, \ldots, \xi_k)^T$$

$$D \equiv D_k = \text{diag}(E(|u_1|^2), \ldots, E(|u_k|^2))$$

$$H \equiv H_k = \text{diag}(1 - \frac{|u_1|^2}{mE(|u_1|^2)}, \ldots, 1 - \frac{|u_k|^2}{mE(|u_k|^2)})$$

(43)

Note that condition $(S_1)$ does not require any changes, and therefore nothing happens to the density $\delta$ and the moment ratio $\rho$ of the sequence of column sizes of $X$. Moreover, it follows from $(S_1)$ (i) that $|u_i|/E(|u_i|^2) \leq 1, \ i = 1, \ldots, k$. Therefore, with these adjustments, Lemma 14 remains valid and the expression (39) for the expected Gram matrix $G$ of $X$ does not change. Hence, the proof of Theorem 3 can be repeated verbatim, leading to **Theorem 3’.** Suppose that random matrix $X$ satisfies conditions $(S_1)$ and $(S'_2)$ with $\xi_i, \ i = 1, \ldots, k$ defined by (43). Let $\tau$ be a transposition of the index set $\{1, \ldots, k\}$ that sorts the sequence of squared column norm expectations $w_i = E(|u_i|^2)$ in descending order. Then for all $i = 1, \ldots, k$

$$(1 + k\delta + O(L/m)) w_{\tau(i)} \geq \sigma_i(G) \geq (1 + \rho + O(L/m))^{-1} w_{\tau(i)}$$
3.4 Gamma Distributed Column Sizes

Finally, we will briefly touch upon a case when column sizes $|u_i|$, $i = 1, 2, \ldots, k$ are sampled from a known distribution for which an estimate of the sample moment ratio $\rho$ is available. For gamma distribution, the result of this kind was obtained in [13], we state it here in a form convenient for what follows.

**Lemma 17.** (see [13]) Let the sample $Y = (y_1, \cdots, y_k)$ be drawn from a population with gamma density

$$g(t, \alpha, \beta) \equiv \frac{\beta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\beta t}, \quad \alpha \geq 1, \beta, t > 0 \quad (44)$$

and let $A_k$, $S_k^2$ and $\Theta^2$ be respectfully its sample mean, sample variance and sample second moment. Then

$$E\left( \frac{S_k^2}{A_k^2} \right) = \frac{E(S_k^2)}{E(A_k^2)} = \alpha^{-1} + O\left( \frac{1}{k} \right)$$

and therefore, by (35)

$$E\left( \frac{\Theta^2}{A_k^2} \right) = \frac{E(\Theta^2)}{E(A_k^2)} = 1 + \alpha^{-1} + O\left( \frac{1}{k} \right)$$

and by Jenssen’s inequality

$$E(\rho(Y)) = E\left( \sqrt{\frac{\Theta^2}{A_k^2}} \right) \leq \sqrt{1 + \alpha^{-1}} + O\left( \frac{1}{k} \right)$$

A minor difficulty in combining Lemma 17 with Theorem 3’ is that to satisfy condition $(S_1)$ we have to consider a case of a left-truncated gamma distribution $g_{\leq a}(t, \alpha, \beta)$, $a \geq 1$, defined as

$$g_{\geq a}(t, \alpha, \beta) = 0 \text{ if } t \in (a, \infty) \quad (45)$$

where

$$F(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^x t^{\alpha-1} e^{-\beta t} dt$$

It is not hard to estimate sample coefficient of variation for a random variable that is distributed according to (45). First, note that a (gamma) distribution is a mixture of its left-truncated and right-truncated distributions

$$g(t, \alpha, \beta) = F(a) g_{\leq a}(t, \alpha, \beta) + (1 - F(a)) g_{\geq a}(t, \alpha, \beta) \quad (46)$$

where

$$g_{\leq a}(t, \alpha, \beta) = 0 \text{ if } t \in (a, \infty) \quad \text{and} \quad g_{\leq a}(t, \alpha, \beta) = \frac{g(t, \alpha, \beta)}{F(a)} \text{ if } t \in [0, a] \quad (47)$$

**Lemma 18.** If $a = 1$, $\alpha \geq 1$ and $\beta < 1$ then the moment ratio $\rho$ of a random variable that follow left-truncated gamma distribution (45) can be estimated as

$$\rho = \sqrt{1 + \frac{1}{\alpha}} + O(\beta^2) \quad (48)$$
Indeed, if $\alpha \geq 1$, then
\[
F(1) = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^1 t^{\alpha-1} e^{-\beta t} dt \leq \beta \int_0^1 e^{-\beta t} dt = \beta + O(\beta^2) \quad (49)
\]
and a similar estimate is valid for moments of a random variable with a right-truncated distribution (47). Now from (46), we have
\[
F(1) \mu_1 + (1 - F(1)) \mu_2 = \frac{\alpha}{\beta}, \quad F(1) s_1 + (1 - F(1)) s_2 = \frac{\alpha + \alpha^2}{\beta}
\]
where $\mu_1, s_1, \mu_2, s_2$ are first two moments of left and right truncated distributions (45) and (47) with $a = 1$. Dividing the second equation by the first and taking into account (49) we get (48).

**Corollary 10.** Under conditions (and in notation) of Theorem 3’, suppose in addition that
\[
(a) \quad \frac{k}{m} \leq \frac{s}{m}
\]
\[
(b) \quad The column sizes $|u_i|, i = 1, \ldots, k$ are sampled from left-truncated gamma distribution (45) with $a = 1, \alpha \geq 1$ and $\beta \leq 1/\sqrt{k}$
\]
Then with high probability
\[
\left(1 + \frac{k}{m\beta} \sqrt{\alpha(\alpha+1)} + O(L/m)\right) w_{\tau(i)} \geq \sigma_i(G) \geq \left(1 + \sqrt{1+1/\alpha} + O(L/m)\right)^{-1} w_{\tau(i)}
\]

Proof. Start by estimating the matrix density (cf. [13]):
\[
E(\delta) = E \left( \frac{1}{k} \sum_i \frac{|u_i|}{m} \right) = \frac{1}{m} E \left( \frac{1}{k} \sum_i |u_i| \right) = \frac{1}{m} \frac{\alpha}{\beta}
\]
It follows from Lemma (18) that we can use the estimate of Lemma 17 for the sample moment ratio of the sequence of column sizes when condition (b) is satisfied. Now, using condition (a) replace the sample moment ratio $\rho$ in Lemma 15 (cf. (41)) by the estimate obtained in Lemma 17. The same substitution can be done in Lemma 16, where just obtained estimate for the matrix density can be used as well.

**Remark 12.** The paper [13] contains some information on estimation of confidence intervals for the sample coefficient of variation. These results can be used to specify more precisely what is meant by ”high probability” in Corollary 10

## 4 Tall Sparse Matrix with rapidly decreasing column norms

We now return to the subject of first two sections. Let $R_{\text{left}} \equiv R[1 : m, 1 : k]$ be a left $m \times k$ sub-matrix of $R$ in (1) and let $R_{\text{right}} \equiv R[1 : m, k+1 : n]$ be the right $m \times (n-k)$ sub-matrix of $R$. In other words, $R_{\text{left}}$ is a matrix formed by the first $k$ columns $u_1, \ldots, u_k$ of $R$. It follows from Corollary 6 that the Algorithm 1 will produce top $i, i \leq k$ singular values of $R$ as long as $\sigma_i(R_{\text{left}}) \geq \|R_{\text{right}}\|$. Is there a way to establish this estimate beforehand, without computing the singular value decomposition? Suppose that the columns of $R$ are sorted in descending order of their norms and that
the sizes of the columns of \( R \) follow a truncated gamma distribution \((45)\) with the shape parameter \( \alpha \).

\( R \) is sparse enough in a sense that constant \( L/m \) of the Corollary 10 is small.

the ratios \( |u_i|/\|u_i\|^2, \ i = 1, \cdots, n \) are roughly equal

Although Corollary 10 is not applicable to a \( R \), we nevertheless, can adopt an estimate

\[ \sigma_i(R_{left}) \geq (1 + \sqrt{1 + 1/\alpha})^{-1/2}\|u_i\|, \ i < k \]

as a heuristic guidance. Note that the norm or the matrix \( R_{right} \) is bound from above by \( \sqrt{|u_{k+1}|\|R_{right}\|_\infty} \) (cf. Lemma (8)). Hence we can assume that the Algorithm 1 will find top \( i \) singular values of \( R \) if there is an index \( i < k \) such that

\[ \|u_i\| \geq \left(1 + \sqrt{1 + 1/\alpha}\right)^{1/2}\sqrt{|u_{k+1}|} \times \text{(the maximal size of rows of } R_{right}) \]

**Example 5.** Take a \( 10^7 \times 10^5 \) sparse (e.g. with density \( \approx 10^{-4} \)) non-negative matrix \( R \), partitioned as in \((1)\) with \( k = 10000 \). It is reasonable to assume that the sizes or the rows are on average a hundred times smaller than the sizes of the columns. Suppose that distribution of sizes of the first 10000 column norms of \( R \) is exponential \((1 + 1/\alpha = 2)\) and let \( i = 1000 \). Assuming that non-zero matrix entries in the first 10000 columns are spread-up evenly, the Algorithm 1 has a chance to recover at least top 1000 singular values of \( R \) if the norm of the 1000-th column of \( R \) is no less than \( 0.1554 \times \text{the size of the } 10001\text{-th column} \)

Consider now a following low rank approximation procedure.

**Algorithm 2.** Given a large sparse \( m \times n \) matrix \( R \)

1) Find block partitioning \((1)\) with relatively dense invertible \( k \times k \) matrix \( A \) and sparse small-norm matrix \( D \)

2) Throw away the bottom right block \( D \) and use Algorithm 1 to find a low rank approximation of the remaining matrix \( R_0 \).

We summarize the heuristics behind this algorithm as follows

**Remark 13.** The Algorithm 2 will approximate a few top singular values of a tall sparse non-negative matrix with an approximation error of less than two operator-norms of the removed bottom right block

5 Acknowledgments

The author would like to thank Rama Ramakrishnan for useful comments and stimulating discussions.
References

[1] Roger A. Horn, Charles R. Johnson, Matrix Analysis, Cambridge Univ. Press, 2012

[2] Gene H. Golub, Charles F. Van Loan, Matrix Computations, John Hopkins Univ. Press, 1991

[3] P. P. Drineas, A. Frieze, R. Kannan, S. Vempala, V. Vinay, Clustering Large Graphs via the Singular Value Decomposition, Machine Learning, 56, 9–33, 2004

[4] William Fulton, Eigenvalues, Invariant Factors, Highest Weights, and Schubert Calculus, [arXiv:math/9908012v3], March 2000

[5] Roy Mathias, The Spectral Norm of a Nonnegative Matrix, Linear Algebra Appl., vol. 139, 2001, 269-284

[6] Larry Goldstein, Gesine Reinert, Total Variation Distance for Poisson Subset Numbers, Annals of Combinatorics, Volume 10, Issue 3, pp 333-341, 2006

[7] Joseph A. Wolf. Spaces of Constant Curvature, University of California, Berkeley, 1972

[8] Dimitris Achlioptas, Frank McSherry, Fast Computation of Low Rank Matrix Approximations, Journal of the ACM (JACM) Volume 54 Issue 2, April 2007

[9] Mark Rudelson, Roman Vershynin, Non-asymptotic theory of random matrices: extreme singular values, Proceedings of the International Congress of Mathematicians, Vol. III, 1576-1602, New Delhi, 2010

[10] Roman Vershynin, Introduction to High Dimensional Probability, 2016, [https://www.math.uci.edu/~rvershyn/papers/HDP-book/HDP-book.html#]

[11] N. Alon, T. Lee, A. Shraibman and S. Vempala, The approximate rank of a matrix and its algorithmic applications, Proc. STOC 2013, 675-684.

[12] Noga Alon and Assaf Naor, Approximating the Cut-Norm via Grothendieck’s Inequality, SIAM J. Comput., 35(4), 787–803, 2006

[13] Tea-Yan Hwang and Ping-Huang Huang, On new Moment Estimation of Parameters of the Gamma Distribution using its Characterization, Ann. Inst. Statist. Math, Vol. 54, No 4, 840-847, 2002

[14] William H. Press, Saul A. Teukolsky, William T. Vetterling, Brian P. Flannery, Numerical Recipes, Third Edition, Cambridge Univ. Press, 2007