Finding Maximum Volume Sub-matrices of a Matrix

Ali Çivril and Malik Magdon-Ismail

Computer Science Department, RPI, 110 8th Street, Troy, NY 12180.
Email: {civria,magdon}@cs.rpi.edu
May 31, 2007

Abstract

Given a matrix $A \in \mathbb{R}^{m \times n}$ (n vectors in m dimensions), we consider the problem of selecting a submatrix (subset of the columns) with maximum volume. The motivation to study such a problem is that if A can be approximately reconstructed from a small number k of its columns (A has "numerical" rank k), then any set of k independent columns of A should suffice to reconstruct A. However, numerical stability results only if the chosen k have large volume. We thus define an appropriate algorithmic problem **Max-Vol(k)**, which asks for the k columns with maximum volume. We show that Max-Vol is NP-hard, and in fact does not admit any PTAS. In particular, it is NP-hard to approximate **Max-Vol** within $\frac{2\sqrt{2}}{3} + \epsilon$. We study a natural greedy heuristic for **Max-Vol** and show that it has approximation ratio $2^{-O(k \log k)}$. We show that our analysis of the greedy heuristic is tight to within a logarithmic factor in the exponent by giving an instance of **Max-Vol** for which the greedy heuristic is $2^{-\Omega(k)}$ from optimal. When A has unit norm columns, a related problem is to select the maximum number of vectors with a given volume (this pre-specified volume could be the volume required on grounds of numerical stability for the reconstruction). We show that if the optimal solution selects k columns, then greedy will select $\Omega(\frac{k}{\log k})$ columns, providing a $\log k$-approximation.

1 Introduction

To motivate the discussion, consider the set of three vectors

$$\left\{ e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, u = \begin{bmatrix} \sqrt{1-\epsilon^2} \\ \epsilon \end{bmatrix} \right\},$$

which are clearly dependent, and any two of which are a basis. Thus any pair can serve to reconstruct all vectors. Suppose we choose e_1, u as the basis, then $e_2 = \frac{1}{\epsilon}u - \frac{\sqrt{1-\epsilon^2}}{\epsilon}e_1$, and we have a numerical instability in this representation as $\epsilon \to 0$. Such problems get more severe as the dimensionality of the space gets large (curse of dimensionality), and it is natural to ask for the representatives to be "as far away from each other as possible". A natural formalization of this problem is to find the representatives which span the largest volume, since the volume is a quantification of how far the vectors are from each other. Thus, given a set of n vectors in $\mathbb{R}^m$ represented as a matrix $A \in \mathbb{R}^{m \times n}$ and a positive integer k, we ask for a subset of size k with maximum volume.

Given a non-empty set of vectors $S = \{v_1, v_2, \cdots v_k\}$ all in $\mathbb{R}^m$, let $Vol(S)$ be the volume defined by the vectors in S. $Vol(S)$ can be recursively defined as follows: if S contains one element, i.e. $S = \{v\}$, then $Vol(S) = \|v\|$, where $\| \cdot \|$ is the Euclidean norm. If S has more than one element,

$Vol(S) = \|v - \pi_{(S-\{v\})}(v)\| \cdot Vol(S - \{v\})$ for any $v \in S$, where $\pi_S(v)$ is the projection of v onto the space spanned by the vectors in S. It is well known that $\pi_{(S-\{v\})}(v) = AA^+v$, where A is the matrix whose columns are the vectors in $S - \{v\}$, and A^+ is the pseudo-inverse of A (see for example [9]). Using this recursive expression, we have

$$Vol(S) = \|v_1\| \cdot \prod_{i=1}^{k-1} \|v_{i+1} - A_i {A_i}^+ v_{i+1}\|$$

where $A_i = [v_1 \cdots v_i]$.

The notion of volume has already received some interest in the algorithmic aspects of linear algebra. In the past decade, the problem of matrix reconstruction and finding low-rank approximations to matrices using a small sample of columns has received much attention (See for example [3, 5, 6, 7]). Ideally, one has to choose the columns to be as independent as possible when trying to reconstruct a matrix using few columns. Along these lines, in [3], the authors introduce 'volume sampling' to find low-rank approximation to a matrix where one picks a subset of columns with probability proportional to their volume squared. Improving the existence results in [3], [4] also provides an adaptive randomized algorithm which includes repetitively choosing a small number of columns in a matrix to find a low-rank approximation. The authors show that if one samples columns proportional to the volume squared, then one obtains a provably good matrix reconstruction (randomized). Thus, sampling larger volume columns is good. A natural question is to ask what happens when one uses the columns with largest volume (deterministic). The problem we address here is the algorithmic problem of obtaining the columns with largest volume and we rely on [4] as the qualitative intuition behind why obtaining the maximum volume submatrix should play an important role.

Another important line of research in linear algebra community is finding rank revealing factorizations of matrices. This problem is closely related to *subset selection* [9] in which one tries to find a subset C of columns of a matrix such that C is as non-singular as possible. Among several approaches to solve the problem, the notion of volume plays an important role [2, 10, 11, 14] where a pivoting strategy based on computing the volumes of subsets differing by one column is considered.

Our Contributions. We prove that an appropriately defined decision version of volume maximization is NP-hard. In fact we prove that no PTAS for volume maximization exists by showing that the problem is inapproximable to within $\frac{2\sqrt{2}}{3} + \epsilon$. Next, we consider a simple (deterministic) greedy algorithm and show that it achieves a $1/k!$ approximation to the optimal volume when selecting k columns. We also construct an explicit example for which the greedy algorithm gives no better than a $1/2^{k-1}$ approximation ratio, thus proving that our analysis of the greedy algorithm is almost tight (to within a logarithmic factor in the exponent).

We then consider the related problem of choosing the maximum number of vectors with a given volume, in the case when all columns in A have unit norm. If the optimal algorithm loses a constant factor with every additional vector selected (which is a reasonable situation), then the optimal volume will be $2^{-\Omega(k)}$. When the optimal volume for k vectors is $2^{-\Omega(k)}$ as motivated above, we prove that the greedy algorithm chooses $\Omega(\frac{k}{\log k})$ columns having at least that much volume. Thus, the greedy algorithm is within a $\log k$-factor of the maximum number of vectors which can be selected given a target volume. The remainder of the paper is structured as follows: In Section 2, we provide hardness results for Max-Vol. The approximation ratio of a natural greedy algorithm is analyzed in Section 3 where we also present a lower bound for the greedy algorithm. Finally, some open questions and comments are outlined in Section 4.

2 Hardness of Volume Maximization

We show NP-hardness even under the restriction that the columns of A have unit norm, followed by a hardness of approximation result. We are interested in choosing a subset of the columns in A whose volume is maximum. Hence, we formulate the following decision problem.

Problem: **Max-Vol**

Instance: A matrix $A \in \mathbb{R}^{m\times n}$ with normalized columns, a real number $V \in [0,1]$, and a positive integer $k \leq \min\{m,n\}$.

Question: Does there exist a subset A' of at least k columns of A, such that $Vol(A') \geq V$?

Theorem 1. *Max-Vol is NP-Hard.*

Proof. We give a reduction from 'exact cover by 3-sets', which is known to be NP-complete (See for example [8, 12]).

Problem: **Exact cover by 3-sets (X3C)**

Instance: A set Q and a collection C of 3-element subsets of Q.

Question: Does there exist an exact cover for Q, i.e. a sub-collection $C' \subseteq C$ such that every element in Q appears exactly once in C'?

We use the following reduction from X3C to Max-Vol: let $Q = \{q_1, q_2, \dots q_m\}$ and $C = \{c_1, c_2, \dots c_n\}$ be given as an instance of X3C. We construct the matrix $A \in \mathbb{R}^{m\times n}$, in which each column $A^{(j)}$ corresponds to the 3-element set c_j. The non-zero entries in $A^{(j)}$ correspond to the elements in c_j. Specifically, set

$$A_{ij} = \begin{cases} 1/\sqrt{3} & \text{if } q_i \in c_j \\ 0 & \text{otherwise} \end{cases}$$

(Note that every $A^{(j)}$ has exactly 3 non-zero entries and has unit norm.) For the instance of Max-Vol, we set $V = 1$ and $k = m/3$.

It is clear that the reduction is polynomial time. All that remains is to show that the instance of X3C is true if and only if the corresponding instance of Max-Vol is true.

Suppose the instance of X3C is true. Then, there is a collection $C' = \{c_{i_1}, c_{i_2}, \dots, c_{i_{m/3}}\}$ of cardinality $m/3$, which exactly covers Q. (Note that, m should be a multiple of 3, otherwise no solution exists.) Consider the columns of A corresponding to the 3-element sets in C'. Since the cover is exact, $c_{i_j} \cap c_{i_k} = \emptyset \ \forall j,k \in \{1,\dots,m/3\}$ where $j \neq k$, which means that $A^{(i_j)} \cdot A^{(i_k)} = 0$. Hence, the columns in $A' = \{A^{(i_1)}, A^{(i_2)}, \dots, A^{(i_{m/3})}\}$ are pair-wise orthonormal. Thus, $Vol(A') = 1$ and the instance of Max-Vol is true.

Conversely, suppose the instance of Max-Vol is true. Let A' be a set of $m/3$ columns of A with $Vol(A') = 1$, which means that the columns in A' are pair-wise orthonormal. Let u, v be two columns in A'; we have $u \cdot v = 0$. Since the entries in A' are all non-negative, $u_i \cdot v_i = 0 \; \forall i \in [1, m]$, i.e. u and v correspond to 3-element sets which are disjoint. Hence, the columns in A' correspond to a sub-collection C' of 3-element sets, which are pair-wise disjoint. Therefore, every element of Q appears at most once in C'. C' contains m elements corresponding to the m non-zero entries in A', it follows that every element of Q appears exactly once in C', concluding the proof. ■

Having shown that the decision problem Max-Vol is NP-hard, it has two natural interpretations as an optimization problem for a given matrix A:

1. *Max-Vol(k)*: Given k, find a subset of size k with maximum volume.
2. *Max-Subset(V)*: Given V and that A has unit norm vectors, find the largest subset $A' \subseteq A$ with volume at least V.

Our reduction in the NP-hardness proof of Max-Vol produces a gap, which provides a hardness of approximation result for Max-Vol(k).

Theorem 2. *Max-Vol(k) is NP-Hard to approximate within* $\frac{2\sqrt{2}}{3} + \epsilon$.

Proof. We already proved that an instance of X3C is true if and only if the maximum volume in the Max-Vol instance is 1. Assume X3C instance is not true. Then, we have at least one overlapping element between two sets. Any collection of size $\frac{m}{3}$ will have two sets v_1, v_2 which have non-zero intersection. The corresponding columns in A' have $d(v_1, v_2) = \|v_1 - (v_1 \cdot v_2) v_2\| = \|v_1 - \frac{1}{3} v_2\| \leq \frac{2\sqrt{2}}{3}$, where $d(v_1, v_2)$ is the orthogonal part of v_1 with respect to v_2. Since $Vol(A') \leq d(v_1, v_2)$, we have $Vol(A') \leq \frac{2\sqrt{2}}{3}$. A polynominal time algorithm with a $\frac{2\sqrt{2}}{3} + \epsilon$ approximation factor for Max-Vol would decide the X3C instance in this case, which would imply $P = NP$. ■

3 The Greedy Approximation Algorithm

Since there is no PTAS for Max-Vol(k), the next natural question is whether there exists a simple heuristic with some approximation guarantee. One obvious strategy is the following greedy algorithm:

Algorithm 1 Greedy(A, k)

```
S ← ∅
while |S| < k do
  Select largest norm vector v ∈ A
  Remove the projection of v from every element of A
  S ← S ∪ v
end while
```

This algorithm has already been proposed to compute QR decomposition of matrices by [1]. Although it is known to work well in practice, there was no provable result for the algorithm, specifically with respect to the non-singularity of the selected columns. Our analysis of Greedy provides an approximation ratio for the maximum volume as well as bounds on the singular values in a rank revealing QR factorization when combined with the results in [10]. In this paper, we do not relate our analysis to that context as the results are rather weak and better algorithms have been proposed as cited in the introduction.

The outline of the remainder of this section is as follows: In Section 3.1, we analyze performance ratio of Greedy. Section 3.2 provides a lower bound Greedy. We also analyze Greedy for Max-Subset(V) in Section 3.3 where we require the columns of the matrix be unit norm since in that case one should guarantee that the volume is either monotonically non-increasing or non-decreasing in the number of vectors chosen by the algorithm. When all the vectors have unit norm, the volume is monotonically non-increasing in the number of vectors chosen, and we analyze the algorithm in this case.

3.1 Approximation Ratio of Greedy

We consider Greedy after k steps. First, we assume that the dimension of the space spanned by the column vectors in A is at least k, since otherwise there is nothing to prove. Let $span(S)$ denote the space spanned by the vectors in the set S and let $\pi_S(v)$ be the projection of v onto $span(S)$. In this section, for brevity, we denote $\|v - \pi_S(v)\|$, the norm of the orthogonal part of v from $span(S)$, by $d(v, S)$. Let $V_k = \{v_1, \dots, v_k\}$ be the set of vectors in order that have been chosen by the greedy algorithm at the end of the k^{th} step. Let $W_k = \{w_1, \dots, w_k\}$ be a set of k vectors of maximum volume. Our main result in this subsection is the following theorem:

Theorem 3. *$Vol(V_k) \geq 1/k! \cdot Vol(W_k)$.*

We prove the theorem through a sequence of lemmas. The basic idea is to show that at the j^{th} step, Greedy loses a factor of at most j to the optimal. Theorem 3 then follows by an elementary induction. First, define $\alpha_i = \pi_{(V_{k-1})}(w_i)$ for $i = 1, \dots, k$. α_i is the projection of w_i onto $span(V_{k-1})$ where $V_{k-1} = \{v_1, \dots, v_{k-1}\}$. Let $\beta_i = w_i - \pi_{(V_{k-1})}(w_i)$. Hence, we have

$$w_i = \alpha_i + \beta_i \qquad for\ i = 1, \dots, k. \tag{1}$$

Note that the dimension of $span(V_{k-1})$ is $k-1$, which means that the α_i's are linearly dependent. We will need some stronger properties of the α_i's.

Definition 4. *A set of m vectors is said to be in general position, if they are linearly dependent and any $m-1$ element subset of them are linearly independent.*

It's immediate from Definition 4 that

Remark 5. *Let $U = \{\gamma_1, \dots, \gamma_m\}$ be a set of m vectors in general position. Then, γ_i can be written as a linear combination of the other vectors in U, i.e.*

$$\gamma_i = \sum_{l \neq i} \lambda_l^i \gamma_l \tag{2}$$

for $i = 1, \dots, m$. λ_l^i's are the coefficients of γ_l in the expansion of γ_i.

Lemma 6. *Let $U = \{\gamma_1, \dots, \gamma_m\}$ be a set of m vectors in general position. Then, there exists a γ_i such that $|\lambda_j^i| \leq 1$ for all $j \neq i$.*

Proof. Assume, without loss of generality that $A = \{\gamma_2, \gamma_3, \ldots, \gamma_m\}$ has the greatest volume among all possible $m-1$ element subsets of U. We claim that γ_1 has the desired property. Consider the set $B_j = \{\gamma_1, \ldots \gamma_{j-1}, \gamma_{j+1}, \ldots, \gamma_m\}$ for $2 \leq j \leq m$. Let $C_j = A - \{\gamma_j\} = B_j - \{\gamma_1\}$. Then, since A has the greatest volume, $Vol(A) = Vol(C_j) \cdot d(\gamma_j, C_j) \geq Vol(B_j) = Vol(C_j) \cdot d(\gamma_1, C_j)$. Hence, we have $d(\gamma_j, C_j) \geq d(\gamma_1, C_j)$. Then, using (2), we can write

$$\gamma_1 = \lambda_j^1 \gamma_j + \sum_{l \neq j, l \neq 1} \lambda_l^1 \gamma_l \tag{3}$$

Denoting $\delta_j = \pi_{C_j(\gamma_j)}$ and $\theta_j = \gamma_j - \delta_j$, (3) becomes

$$\gamma_1 = \left(\lambda_j^1 \delta_j + \sum_{l \neq j, l \neq 1} \lambda_l^j \gamma_l \right) + \lambda_j^1 \theta_j$$

where the first two terms are in $span(C_j)$. Hence, $\theta_1 = \gamma_1 - \pi_{C_j}(\gamma_1) = \lambda_j^1 \theta_j$ and so $\|\theta_1\| = |\lambda_j^1| \|\theta_j\|$. Note that $\|\theta_1\| = d(\gamma_1, C_j)$ and $\|\theta_j\| = d(\gamma_j, C_j)$, so $d(\gamma_1, C_j) = |\lambda_j^1| d(\gamma_j, C_j)$. Since $d(\gamma_1, C_j) \geq d(\gamma_j, C_j)$, we have $|\lambda_j^1| \leq 1$. ∎

Lemma 7. *If $\|\alpha_i\| > 0$ for $i = 1, \ldots, k$ and $k \geq 2$, then there exists a set of m vectors $U = \{\alpha_{i_1}, \ldots, \alpha_{i_m}\} \subseteq \{\alpha_1, \ldots, \alpha_k\}$ with $m \geq 2$ that are in general position.*

Proof. Note that the cardinality of a set U with the desired properties should be at least 2, since otherwise there is nothing to prove. We argue by induction on k. For the base case $k = 2$, we have two vectors α_1 and α_2 spanning a 1-dimensional space and clearly any one of them is linearly independent since neither is 0. Assume that, as the induction hypothesis, any set of $k \geq 2$ non-zero vectors $\{\alpha_1, \ldots, \alpha_k\}$ spanning at most a $k-1$ dimensional space has a non-trivial subset in general position. Consider a $k+1$ element set $A = \{\alpha_1, \ldots, \alpha_{k+1}\}$ with $dim(span(A)) \leq k$. If the vectors in A are not in general position, then there is a k element subset A' of A which is linearly dependent. Hence, $dim(span(A')) \leq k-1$ for which, by the induction hypothesis, we know that there exists a non-trivial subset in general position. ∎

The existence of a subset in general position guaranteed by Lemma 7 will be needed when we apply the next lemma.

Lemma 8. *Assume $\|\alpha_i\| > 0$ for $i = 1, \ldots, k$. Then, there exists an α_{i_j} such that $d(\alpha_{i_j}, W'_{k-1}) \leq (m-1) \cdot d(v_k, V_{k-1})$, where $W'_{k-1} = W_k - \{w_{i_j}\}$.*

Proof. Let $U = \{\alpha_{i_1}, \ldots, \alpha_{i_m}\} \subseteq \{\alpha_1, \ldots, \alpha_k\}$ be in general position where $m \geq 2$ (the existence of U is given by Lemma 7). Assume α_{i_1} has the property given by Lemma 6. Let $U' = \{w_{i_2}, \ldots, w_{i_m}\}$. We claim that α_{i_1} has the desired property. First, note that $d(\alpha_{i_1}, W'_{k-1}) \leq d(\alpha_{i_1}, U')$, since $span(U')$ is a subspace of $span(W'_{k-1})$. We seek a bound on $d(\alpha_{i_1}, W'_{k-1})$. Using (2) and (1), we have

$$\alpha_{i_1} = \sum_{l \neq 1} \lambda_{i_l}^1 \alpha_{i_l} = \sum_{l \neq 1} \lambda_{i_l}^1 (w_{i_l} - \beta_{i_l}).$$

where α_{i_l}'s are the vectors in U and β_{i_l}'s are their orthogonal parts. Rearranging,

$$\sum_{l\neq 1}\lambda^1_{i_l}\beta_{i_l} = \sum_{l\neq 1}\lambda^1_{i_l}w_{i_l} - \alpha_{i_1}.$$

Note that the right hand side is an expression for the difference between a vector in $span(U')$ and α_{i_1}. Hence,

$$\begin{aligned}
d(\alpha_{i_1}, W'_{k-1}) &\leq d(\alpha_{i_1}, U') \\
&= \min_{v\in span(U')} \|v - \alpha_{i_1}\| \\
&\leq \|\sum_{l\neq 1}\lambda^1_{i_l}w_{i_l} - \alpha_{i_1}\| \\
&= \|\sum_{l\neq 1}\lambda^1_{i_l}\beta_{i_l}\| \\
&\leq \sum_{l\neq 1}\lambda^1_{i_l}\|\beta_{i_l}\| \\
&\leq (m-1)\cdot max_{1\leq l\leq m}\|\beta_{i_l}\| \\
&\leq (m-1)\cdot d(v_k, V_{k-1}).
\end{aligned}$$

where the last two inequalities follow from Lemma 6 and the greedy property of the algorithm, respectively. ■

Before stating the final lemma, which gives the approximation factor of the algorithm at each round, we need the following simple observation.

Lemma 9. *Let u be a vector, let V and W be subspaces and let $\alpha = \pi_V(u)$. Then $d(u, W) \leq d(u, V) + d(\alpha, W)$.*

Proof. Let $\gamma = \pi_W(\alpha)$. By triangle inequality for vector addition, we have
$\|u - \gamma)\| \leq \|u - \alpha\| + \|\alpha - \gamma\| = d(u, V) + d(\alpha, W)$. The result follows since $d(u, W) \leq \|u - \gamma\|$. ■

Lemma 10. *At the k^{th} step of the algorithm, there exists a w_i such that*
$d(w_i, W'_{k-1}) \leq k \cdot d(v_k, V_{k-1})$ where $W'_{k-1} = W_k - \{w_i\}$.

Proof. For $k = 1$, there's nothing to prove. For $k \geq 2$, there are two cases.

1. One of the w_i's is orthogonal to V_{k-1} ($\|\alpha_i\| = 0$). In this case, by the greedy property, $d(v_k, V_{k-1}) \geq \|w_i\| \geq d(w_i, W'_{k-1})$, which gives the result.

2. For all w_i, $\|\alpha_i\| > 0$, i.e., all w_i have non-zero projection on V_{k-1}. Assuming that $\alpha_1 = \pi_{V_{k-1}}(w_1)$ has the desired property proved in Lemma 8, we have for the corresponding w_1

$$\begin{aligned}
d(w_1, W'_{k-1}) &\leq d(w_1, V_{k-1}) + d(\alpha_1, W'_{k-1}) \\
&\leq \|\beta_1\| + d(\alpha_1, W'_{k-1}) \\
&\leq \|\beta_1\| + (m-1)\cdot d(v_k, V_{k-1}) \\
&\leq m \cdot d(v_k, V_{k-1}).
\end{aligned}$$

The first inequality is due to Lemma 9. The last inequality follows from the greedy property of the algorithm, i.e. the fact that $d(v_k, V_{k-1}) \geq \|\beta_1\|$. The lemma follows since $m \leq k$. ∎

The last lemma immediately leads to the result of Theorem 3, with a simple inductive argument as follows:

Proof. The base case is easily established since $Vol(V_1) = Vol(W_1)$. Assume that $Vol(V_{k-1}) \geq \frac{1}{(k-1)!} \cdot Vol(W_{k-1})$ for some $k > 2$. By Lemma 10, we have a w_i such that $d(w_i, W'_{k-1}) \leq k \cdot d(v_k, V_{k-1})$ where $W'_{k-1} = W_k - \{w_i\}$. It follows that

$$\begin{aligned} Vol(V_k) &= d(v_k, V_{k-1}) \cdot Vol(V_{k-1}) \\ &\geq \frac{d(w_i, W'_{k-1})}{k} \cdot \frac{Vol(W_{k-1})}{(k-1)!} \\ &\geq \frac{d(w_i, W'_{k-1})}{k!} \cdot Vol(W'_{k-1}) \\ &= \frac{Vol(W_k)}{k!}. \end{aligned}$$

∎

3.2 Lower Bound for the Greedy Algorithm

We give a lower bound of $1/2^{k-1}$ for the approximation factor of Greedy by explicitly constructing a bad example. We will inductively construct a set of unit vectors satisfying this lower bound. It will be the case that the space spanned by the vectors in the optimal solution is the same as the space spanned by the vectors chosen by Greedy. We will first consider the base case $k = 2$: let the matrix $A = [v_1 w_1 w_2]$ where $dim(A) = 2$ and $d(v_1, w_1) = d(v_1, w_2) = \delta$ for some $1 > \delta > 0$ such that θ, the angle between w_1 and w_2 is twice the angle between v_1 and w_1, i.e. v_1 is 'between' w_1 and w_2. If the greedy algorithm first chooses v_1, then $\lim_{\delta \to 0} Vol(V_2)/Vol(W_2) = \frac{1}{2} \cos \frac{\theta}{2} = \frac{1}{2}$. Hence, for $k = 2$, there's a set of vectors for which $Vol(W_2) = (2 - \epsilon) \cdot Vol(V_2)$ for arbitrarily small $\epsilon > 0$.

For arbitrarily small $\epsilon > 0$, assume that there is an optimal set of k vectors $W_k = \{w_1, \ldots, w_k\}$ such that $Vol(W_k) = (1 - \epsilon)2^{k-1} \cdot Vol(V_k)$ where $V_k = \{v_1, \ldots, v_k\}$ is the set of k vectors chosen by the algorithm. The vectors in W_k and V_k span a subspace of dimension k, and assume $w_i \in \mathbb{R}^d$ where $d > k$. Let $d(v_2, V_1) = \epsilon_1 = \delta$ for some $1 > \delta > 0$, and $d(v_{i+1}, V_i) = \epsilon_i = \delta\epsilon_{i-1}$ for $i = 2, \ldots, k-1$. Thus, $Vol(V_k) = \delta^{\frac{k(k-1)}{2}}$ and $Vol(W_k) = (1 - \epsilon)2^{k-1}\delta^{\frac{k(k-1)}{2}}$. Assume further that for all w_i in W_k, $d(w_i, V_j) \leq \epsilon_j$ for $j = 1, \ldots, k-2$ and $d(w_i, V_{k-1}) = \epsilon_{k-1}$ so that there exists an execution of Greedy where no $\{v_1, \ldots, v_k\}$ is chosen.

We will now construct a new set of vectors $W_{k+1} = W'_k \cup \{w_{k+1}\} = \{w'_1, \ldots, w'_k, w_{k+1}\}$ which will be the optimal solution. Let $w_i^j = \pi_{V_j}(w_i)$, and let $e_i^j = \pi_{V_j}(w_i) - \pi_{V_{j-1}}(w_i)$ for $j = 2, \ldots, k$ and $e_i^1 = w_i^1$. Namely, e_j^i is the component of w_i which is in V_j, but perpendicular to V_{j-1} and e_i^1 is the component of w_i which in in the span of v_1. (Note that $\|e_i^k\| = \epsilon_{k-1}$.) Let u be a unit vector perpendicular to $span(W_k)$. For each w_i we define a new vector $w'_i = (\sum_{j=1}^{k-1} e_i^j) + \sqrt{1-\delta^2} e_i^k + \delta\epsilon_{k-1}u$. Intuitively, we are defining a set of new vectors which are first rotated towards V_{k-1} and then towards u such that they are $\delta\epsilon_{k-1}$ away from V_k. Introduce another vector $w_{k+1} = \sqrt{1-\delta^2} v_1 - \delta\epsilon_{k-1}u$. Intuitively, this new vector is v_1 rotated towards the negative direction of u. Note that, in this setting $\epsilon_k = \delta\epsilon_{k-1}$. We finally choose $v_{k+1} = w_{k+1}$.

Lemma 11. *For any $w \in W_{k+1}$, $d(w, V_j) \le \epsilon_j$ for $j = 1, \ldots, k-1$ and $d(w, V_k) = \epsilon_k$.*

Proof. For $w = w_{k+1}$, $d(w_{k+1}, V_j) = \epsilon_k \le \epsilon_j$ for $j = 1, \ldots, k$. Let $w = w'_i$ for some $1 \le i \le k$. Then, for any $1 \le j \le k-1$, we have $d(w'_i, V_j)^2 = \sum_{l=j+1}^{k-1} \|e_i^l\|^2 + (1-\delta^2)\|e_i^k\|^2 + \delta^2\|e_i^k\|^2 = \sum_{l=j+1}^{k} \|e_i^l\|^2 = d(w_i, V_j)^2 \le {\epsilon_j}^2$ by the induction hypothesis. ∎

Lemma 11 ensures that $\{v_1, \ldots, v_{k+1}\}$ is a valid output of Greedy. What remains is to show that for any $\epsilon > 0$, we can choose δ sufficiently small so that $Vol(W_{k+1}) \ge (1-\epsilon)2^k \cdot Vol(V_{k+1})$. In order to show this, we will need the following lemmas.

Lemma 12. $\lim_{\delta \to 0} Vol(W_{k+1}) = 2\epsilon_k \cdot Vol(W_k)$.

Proof. With a little abuse of notation, let W_{k+1} denote the matrix of coordinates for the vectors in the set W_{k+1}.

$$W_{k+1} = \begin{pmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,k} & \sqrt{1-\delta^{2k}} \\ w_{2,1} & w_{2,2} & \cdots & w_{2,k} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \sqrt{1-\delta^2}w_{k,1} & \sqrt{1-\delta^2}w_{k,2} & \cdots & \sqrt{1-\delta^2}w_{k,k} & 0 \\ \delta^k & \delta^k & \cdots & \delta^k & -\delta^k \end{pmatrix}$$

where $w_{i,j}$ is the i^{th} coordinate of w_j, which is in W_k. (Note that this is exactly how U is constructed in the inductive step). Expanding on the right-most column of the matrix, we have

$$Vol(W_{k+1}) = |det(W_{k+1})| = |\sqrt{1-\delta^{2k}} \cdot det(A) + (-1)^{k+1}\delta^k \cdot det(B)| \tag{4}$$

where A and B are the corresponding minors of the coefficients, i.e. the left-most lower and upper $k \times k$ sub-matrices of W_{k+1}, respectively. Clearly, we have $det(B) = \sqrt{1-\delta^2} \cdot det(W_k)$ where W_k is the matrix of coordinates for the vectors in the set W_k. Let C be the matrix obtained by replacing each $w_{1,i}$ by 1 in W_k. Then, using row interchange operations on A, we can move the last row of A to the top. This gives a sign change of $(-1)^{k-1}$. Then, factoring out $\sqrt{1-\delta^2}$ and δ^k in the first and last rows respectively, we have $det(A) = (-1)^{k-1}\delta^k\sqrt{1-\delta^2} \cdot det(C)$. Hence, (4) becomes

$$|det(W_{k+1})| = (\delta^k\sqrt{1-\delta^2})|\sqrt{1-\delta^{2k}} \cdot det(C) + det(W_k)| \tag{5}$$

We will need the following lemma to compare $det(W_k)$ and $det(C)$.

Lemma 13. $\lim_{\delta \to 0} \frac{\det(C)}{\det(W_k)} = 1$.

Proof. For $i > 1$, the elements of the i^{th} rows of both W_k and C has δ^{i-1} as a common coefficient by construction. Factoring out these common coefficients, we have $det(W_k) = \delta^{\frac{k(k-1)}{2}} \cdot det(U)$ and $det(C) = \delta^{\frac{k(k-1)}{2}} \cdot det(U')$ where U and U' are matrices with non-zero determinants as δ approaches 0. Furthermore, $lim_{\delta \to 0} det(U) = det(U')$ as the elements in the first row of U approaches 1. The result then follows. ∎

Using Lemma 13 and (5), we have

$$lim_{\delta \to 0} Vol(W_{k+1}) = lim_{\delta \to 0}|det(W_{k+1})| = 2\delta^k|det(W_k)| = 2\epsilon_k \cdot Vol(W_k)$$

∎

Theorem 14. *$Vol(W_{k+1}) \geq (1-\epsilon)2^k \cdot Vol(V_{k+1})$ for arbitrarily small $\epsilon > 0$.*

Proof. Given any $\epsilon' > 0$ we can choose δ small enough so that $Vol(W_{k+1}) \geq 2\epsilon_k(1-\epsilon') \cdot Vol(W_k)$, which is always possible by Lemma 12. Given any ϵ'', we can apply induction hypothesis to obtain V_k and W_k such that $Vol(W_k) \geq (1-\epsilon'')2^{k-1} \cdot Vol(V_k)$. Thus,

$$\begin{aligned} Vol(W_{k+1}) &\geq 2\epsilon_k(1-\epsilon') \cdot Vol(W_k) \\ &\geq 2\epsilon_k(1-\epsilon')(1-\epsilon'')2^{k-1} \cdot Vol(V_k) \\ &= (1-\epsilon')(1-\epsilon'')2^k \cdot Vol(V_{k+1}), \end{aligned}$$

where we have used $Vol(V_{k+1}) = \epsilon_k \cdot Vol(V_k)$. Choosing ϵ' and ϵ'' small enough such that $(1-\epsilon')(1-\epsilon'') > 1-\epsilon$ gives the result. ■

3.3 Maximizing the Number of Unit norm Vectors Attaining A Given Volume

In this section, we give a result on approximating the maximum number of unit norm vectors which can be chosen to have at least a certain volume. This result is essentially a consequence of the previous approximation result. We assume that all the vectors in A have unit norm, hence the volume is non-increasing in the number of vectors chosen by the algorithm. Let OPT_k denote the optimal volume for k vectors. Note that $OPT_k \geq OPT_{k+1}$ and the number of vectors m, chosen by Greedy attaining volume at least OPT_k is not greater than k. Our main result states that, if the optimal volume of k vectors is $2^{-\Omega(k)}$, then Greedy chooses $\Omega(\frac{k}{\log k})$ vectors having at least that volume. Thus, Greedy gives a $\log k$-approximation to the optimal number of vectors. We prove the result through a sequence of lemmas. The following lemma is an immediate consequence of applying Greedy on W_k.

Lemma 15. *Let $W_k = \{w_1, \ldots, w_k\}$ be a set of k vectors of optimal volume OPT_k. Then there exists a permutation π of the vectors in W_k such that $d_{\pi(k)} \leq d_{\pi(k-1)} \leq \ldots \leq d_{\pi(2)}$ where $d_{\pi_i} = d(w_{\pi_i}, \{w_{\pi_1}, \ldots, w_{\pi_{i-1}}\})$ for $k \geq i \geq 2$.*

We use this existence result to prove the following lemma.

Lemma 16. *$OPT_m \geq (OPT_k)^{\frac{m-1}{k-1}}$ where $m \leq k$.*

Proof. Let $W_k = \{w_1, \ldots, w_k\}$ be a set of vectors of optimal volume OPT_k. By Lemma 15, we know that there exists an ordering of vectors in W_k such that $d_{\pi(k)} \leq d_{\pi(k-1)} \leq \ldots \leq d_{\pi(2)}$ where $d_{\pi_i} = d(w_{\pi_i}, \{w_{\pi_1}, \ldots, w_{\pi_{i-1}}\})$ for $k \geq i \geq 2$. Let $W_m{}' = \{w_{\pi(1)}, \ldots, w_{\pi(m)}\}$. Then, we have $OPT_m \geq Vol(W_m{}') = \prod_{i=2}^{m} d_{\pi_i} \geq (\prod_{i=2}^{k} d_{\pi_i})^{\frac{m-1}{k-1}} = (OPT_k)^{\frac{m-1}{k-1}}$. ■

Lemma 17. *Suppose $OPT_k \leq 2^{\frac{(k-1)m\log m}{m-k}}$. Then, the greedy algorithm chooses at least m vectors whose volume is at least OPT_k.*

Proof. We are seeking a condition for OPT_k which will provide a lower bound for m such that $\frac{OPT_m}{m!} \geq OPT_k$. If this holds, then $Vol(Greedy_m) \geq \frac{Opt_m}{m!} \geq OPT_k$ and so Greedy can choose at least m vectors which have volume at least OPT_k. It suffices to find such an m satisfying $\frac{(OPT_k)^{\frac{m-1}{k-1}}}{m!} \geq OPT_k$ by Lemma 16. This amounts to $\frac{1}{m!} \geq (OPT_k)^{1-\frac{m-1}{k-1}}$. Since $\frac{1}{m!} \geq \frac{1}{m^m}$ for $m \geq 1$, we require

$$\frac{1}{m^m} \geq (OPT_k)^{1-\frac{m-1}{k-1}}.$$

Taking logarithms of both sides and rearranging, we have

$$-\frac{(k-1)m}{k-m}\log m \geq \log OPT_k.$$

Taking exponents of both sides yields

$$2^{\frac{(k-1)m\log m}{m-k}} \geq OPT_k.$$

■

In order to interpret this result, we will need to restrict OPT_k. Otherwise, for example if $OPT_k = 1$, the greedy algorithm may never get more than 1 vector to guarantee a volume of at least OPT_k since it might be possible to miss guess the first vector. In essence, the number of vectors chosen by the algorithm depends on OPT_k. First, we discuss what is a reasonable condition on OPT_k. Consider n vectors in m dimensions which defines a point in $\mathbb{R}^{m\times n}$. The set of points in which any two vectors are orthogonal has measure 0. Thus, define $2^{-\alpha} = \max_{ij} d(v_i, v_j)$. Then, it is reasonable to assume $\alpha > 0$, in which case $OPT_k \leq 2^{-\alpha k} = 2^{-\Omega(k)}$. Hence, we provide the following theorem which follows from the last lemma under the reasonable assumption that the optimal volume decreases by a constant factor with the addition of one more vector.

Theorem 18. *If $OPT_k \leq 2^{-\Omega(k)}$, then the greedy algorithm chooses $\Omega\left(\frac{k}{\log k}\right)$ vectors having volume at least OPT_k.*

Proof. For some α, $OPT_k \leq 2^{-\alpha k}$. Thus, we solve for m such that

$$2^{-\alpha k} \leq 2^{\frac{(k-1)m\log m}{m-k}}.$$

Suitable rearrangements yield

$$m \leq \frac{\alpha k(k-m)}{(k-1)\log m} \leq \frac{2\alpha k}{\log m}.$$

For m, the largest integer such that $m \leq \frac{2\alpha k}{\log m}$, we have

$$m \approx \frac{2\alpha k}{\log\left(\frac{2\alpha k}{\log m}\right)} = \frac{2\alpha k}{\log(2\alpha k) - \log\log m} = \Omega\left(\frac{k}{\log k}\right).$$

■

In reality, for a random selection of n vectors in m dimensions, α will depend on n and so the result is not as strong as it appears.

4 Discussion

Our analysis of the approximation ratio relies on finding the approximation factor at each round of Greedy. Indeed, we have found examples for which the volume of the vectors chosen by the greedy algorithm falls behind the optimal volume by as large a factor as $1/k$, making Lemma 10 tight. But it might be possible to improve the analysis by correlating the 'gains' of the greedy algorithm between different steps. Hence, one of the immediate questions is that whether one can close the gap between the approximation ratio and the lower bound for the greedy algorithm.

We list other open problems as follows:

- Do there exist efficient non-greedy algorithms with better guarantees for this problem? Max-Vol(k) does not appear to admit any canonical techniques like LP or SDP.
- There is a huge gap between the approximation ratio of the greedy algorithm we have analyzed and the inapproximability result. Can this gap be closed on the inapproximability side by using more advanced techniques?
- Volume seems to play an important role in constructing a low-rank approximation to a matrix. Solutions proposed thus far consider only randomized algorithms. Can this work be extended to find a deterministic algorithm for matrix reconstruction?

We would like to note that the approximation ratio of the greedy algorithm is considerably small because of the 'multiplicative' nature of the problem. Another important problem which resembles Max-Vol in terms of behavior (but not necessarily in nature) is the Shortest Vector Problem (SVP), which is not known to have a polynomial factor approximation algorithm. Indeed, the most common algorithm which works well in practice has $2^{O(n)}$ approximation ratio [13] and non-trivial hardness results for this problem are difficult to find.

References

[1] P. A. Businger and G. H. Golub. Linear least squares solutions by Householder transformations. *Numerische Mathematik*, (7):269–276, 1965.

[2] F. R. de Hooga and R. M. M. Mattheijb. Subset selection for matrices. *Linear Algebra and its Applications*, (422):349–359, 2007.

[3] A. Deshpande, L. Rademacher, S. Vempala, and G. Wang. Matrix approximation and projective clustering via volume sampling. In *SODA '06*, pages 1117–1126. ACM Press, 2006.

[4] A. Deshpande and S. Vempala. Adaptive sampling and fast low-rank matrix approximation. In *RANDOM'06*, pages 292–303. Springer, 2006.

[5] P. Drineas, R. Kannan, and M. W. Mahoney. Fast monte carlo algorithms for matrices II: Computing a low-rank approximation to a matrix. *SIAM Journal on Computing*, 36(1):158–183, 2006.

[6] P. Drineas, R. Kannan, and M. W. Mahoney. Fast monte carlo algorithms for matrices III: Computing a compressed approximate matrix decomposition. *SIAM Journal on Computing*, 36(1):184–206, 2006.

[7] A. Frieze, R. Kannan, and S. Vempala. Fast monte-carlo algorithms for finding low-rank approximations. *Journal of the Association for Computing Machinery*, 51(6):1025–1041, 2004.

[8] M. R. Garey and D. S. Johnson. *Computers and Intractability.* W. H. Freeman, 1979.

[9] G. H. Golub and C. V. Loan. *Matrix Computations.* Johns Hopkins U. Press, 1996.

[10] M. Gu and S. C. Eisenstat. Efficient algorithms for computing a strong rank-revealing QR factorization. *SIAM Journal on Scientific Computing*, 17(4):848–869, 1996.

[11] Y. P. Hong and C.-T. Pan. Rank-revealing QR factorizations and the singular value decomposition. *Mathematics of Computation*, 58:213–232, 1992.

[12] R. M. Karp. Reducibility among combinatorial problems. In R. E. Miller and J. W. Thatcher, editors, *Complexity of Computer Computations*, pages 85–103. Plenum Press, 1972.

[13] A. K. Lenstra, H. W. Lenstra, and L. Lovasz. Factoring polynomials with rational coefficients. *Mathematische Annalen*, (261):515–534, 1982.

[14] C.-T. Pan. On the existence and computation of rank-revealing *LU* factorizations. *Linear Algebra and its Applications*, 316(1-3):199–222, 2000.