RC24696 (W0811-137) November 24, 2008
Mathematics

IBM Research Report

Minkowski and KZ Reduction of Nearly Orthogonal Lattice Bases

Sanjeeb Dash, Ramesh Neelamani*, Gregory B. Sorkin
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598

*ExxonMobil Upstream Research Company
3319 Mercer
Houston, TX 77027

Research Division
Almaden - Austin - Beijing - Cambridge - Haifa - India - T. J. Watson - Tokyo - Zurich

Minkowski and KZ reduction of nearly orthogonal lattice bases

Sanjeeb Dash, Ramesh Neelamani, Gregory B. Sorkin *

November 24, 2008

Abstract

We prove that if a lattice basis is *nearly orthogonal* (the angle between any basis vector and the linear subspace spanned by the other basis vectors is at least $\frac{\pi}{3}$ radians), then the basis is nearly KZ-reduced for some ordering of the basis vectors. It follows that a KZ-reduced basis can be obtained from a nearly orthogonal basis in polynomial time. We also show that if a nearly orthogonal lattice basis has *nearly equal* vector lengths (they are within a certain constant factor of one another), then the basis is Minkowski reduced. We use this result to show that m i.i.d. random vectors drawn from a uniform distribution over the unit ball in $\mathbb{R}^n$ form a Minkowski-reduced basis of the lattice generated by the vectors almost surely as n tends to infinity, if $m \leq cn$ for a constant $c < 1/4$. This result extends a result of Donaldson (1979) who proved the above property for fixed m as n tends to infinity.

Keywords

lattices, shortest lattice vector, random lattice, Minkowski reduced.

1 Introduction

A lattice $\mathcal{L}$ in $\mathbb{R}^n$ is the set of all integer linear combinations of a finite set of linearly independent vectors; that is, $\mathcal{L} = \{u_1 b_1 + u_2 b_2 + \cdots + u_m b_m \,|\, u_i \in \mathbb{Z}\}$ where $b_1, \ldots, b_m$ are in $\mathbb{R}^n$ and are linearly independent. The set of vectors $\mathcal{B} = \{b_1, \ldots, b_m\}$ is called a *basis* of the lattice $\mathcal{L}$. A lattice is said to be m-dimensional if a basis contains m vectors.

Neelamani, Dash and Baraniuk [8] define a lattice basis to be θ-orthogonal if the angle between any basis vector and the linear subspace spanned by the remaining basis vectors is at least θ. They call a θ-orthogonal basis *nearly orthogonal* if θ is at least $\frac{\pi}{3}$ radians. They prove that a $\frac{\pi}{3}$-orthogonal basis always contains a shortest non-zero lattice vector. Thus, the shortest lattice vector problem (SVP) becomes trivial for a $\frac{\pi}{3}$-orthogonal basis, just as it is trivial for an orthogonal basis.

We prove additional properties of $\frac{\pi}{3}$-orthogonal bases in this paper. We show that such a basis is "nearly" *KZ-reduced*, in the sense that a $\frac{\pi}{3}$-orthogonal basis can be transformed into a KZ-reduced basis in polynomial time. Secondly, we show that if all vectors of a θ-orthogonal ($\theta \geq \frac{\pi}{3}$) basis have lengths no more than $2/\cos\theta$ times the length of the shortest basis vector, then the basis is *Minkowski reduced* for some ordering of the vectors.

Various authors have studied lattice problems in the context of random lattices, i.e., lattices generated by a set of random vectors. Donaldson [4] proved that for fixed m, a random $n \times m$ matrix chosen uniformly

*S. Dash and G.B. Sorkin are with the IBM T. J. Watson Research Center, Yorktown Heights, NY 10598; Email: {sanjeebd,sorkin}@us.ibm.com. R. Neelamani is with the ExxonMobil Upstream Research Company, 3319 Mercer, Houston, TX 77027; Email: neelsh@rice.edu; Fax: 713 431 6161.

from the unit sphere in $\mathbb{R}^{nm}$ defines a Minkowski reduced basis of the lattice generated by its columns almost surely as $n \to \infty$. It follows that as $n \to \infty$, one of the columns of the random matrix is a shortest lattice vector almost surely. In earlier work, Daudé and Vallée [3] showed that in a random lattice generated by n vectors chosen independently and uniformly from the unit ball in $\mathbb{R}^n$, the expected number of iterations of the LLL algorithm [6] is $O(n^2 \log n)$. Now consider an m-dimensional lattice in $\mathbb{R}^n$ generated by m vectors chosen independently and uniformly from the unit ball in $\mathbb{R}^n$; these vectors form a basis because they are linearly independent almost surely. Recently, Akhavi, Marckert and Rouault [2] proved that as $n - m \to \infty$, the probability that this basis is "nearly LLL-reduced" [6] tends to 1. More precisely, they show that such a basis satisfies the *size condition* in the LLL algorithm (see condition (ii) in the definition of LLL reduction in Section 3). Neelamani, Dash and Baraniuk [8] show that for $c = 0.071$, if $m \leq cn$, then as $n \to \infty$, this basis is almost surely $\frac{\pi}{3}$-orthogonal. In Section 4, we improve the value of c to 1/4. We then show that if c is any constant less than 1/4, then as $n \to \infty$, $m \leq cn$ i.i.d. vectors chosen from a uniform distribution over the unit ball in $\mathbb{R}^n$ form a Minkowski reduced basis almost surely (for some re-ordering of the vectors) and also a KZ-reduced basis almost surely. These results imply Donaldson's result cited above.

2 Definitions

Consider an m-dimensional lattice $\mathcal{L}$ in $\mathbb{R}^n$, $m \leq n$. By an *ordered basis* of $\mathcal{L}$, we mean a basis with a given ordering of the basis vectors. We represent such a basis by an ordered set, for example $(b_1, \ldots, b_m)$, or as the columns of a matrix $\mathcal{B}$. We represent an unordered basis as $\{b_1, \ldots, b_m\}$. We refer to the Euclidean norm of a vector v as its length, and denote it by $\|v\|$. We denote the length of a shortest non-zero vector in a lattice $\mathcal{L}$ by $\lambda(\mathcal{L})$,

Definition 1 (Minkowski reduced). *An ordered basis $(b_1, \ldots, b_m)$ is* Minkowski reduced *if b_1 is a shortest lattice vector, and for $i = 2, 3, \ldots, m$, b_i is a shortest vector among lattice vectors not in the span of $\{b_1, \ldots, b_{i-1}\}$.*

In definitions like the above, in our randomized setting we could replace "a shortest vector" with "the shortest vector", as it is almost certainly unique up to sign, but we will stay with the more general phrasing.

The QR decomposition of a matrix $\mathcal{B}$ is its representation as $\mathcal{B} = QR$ where Q is an orthonormal matrix and R is an upper-triangular matrix. Let an ordered basis be given by the columns of a matrix $\mathcal{B} = QR$.

Definition 2 (Proper). *A basis $\mathcal{B} = QR$, $R = (r_{ij})$, is* proper *[5] if $r_{ij} \in (-|r_{ii}|/2, |r_{ii}|/2]$ for all $i, j \in 1, \ldots, m$.*

Lovász [7] refers to a proper basis as a *weakly reduced* basis.

Definition 3 (LLL-reduced). *Given a number δ in the interval $(0, \sqrt{3}/2)$, a basis $\mathcal{B} = QR$, $R = (r_{ij})$, is* δ-LLL-reduced *if*

(i) for $i = 1, \ldots, m-1$, $|r_{i+1,i+1}| \geq \delta |r_{ii}|$ and
(ii) it is proper.

Condition (i) is what we called the *size condition* in Section 1. For any fixed δ in the interval $(0, \sqrt{3}/2)$, the LLL algorithm [6] obtains a δ-LLL-reduced basis in polynomial time with the property that the first basis vector has length at most $\lambda(\mathcal{L})/\delta^{n-1}$.

We define $\text{proj}_{\perp b_1, \ldots, b_i}(v)$ to be the projection of the vector v on the subspace orthogonal to $b_1, \ldots, b_i$. The next definition is from Kannan [5].

Definition 4 (KZ reduced). *An ordered basis* $(b_1, \ldots, b_n)$ *is* KZ-reduced *if:*
(i) b_1 *is a shortest lattice vector;*
(ii) in the lattice generated by $\mathcal{B}' = \{\text{proj}_{\perp b_1}(b_2), \ldots, \text{proj}_{\perp b_1}(b_n)\}$, $\mathcal{B}'$ *is a KZ-reduced basis; and*
(iii) the basis is proper.

The letters K and Z in KZ-reduced stand for Korkin and Zolotarev. Kannan [5] observes that a KZ-reduced basis is also LLL-reduced (for some $\delta \in (0, \sqrt{3}/2)$).

Definition 5 (nearly KZ reduced). *A basis is* nearly KZ-reduced *if it satisfies (i) in Definition 4 and satisfies (ii) with the words KZ-reduced replaced by nearly KZ-reduced.*

It is well-known [5] that a nearly KZ-reduced basis can be transformed to a KZ-reduced basis (i.e., properness attained) in polynomial time, via the procedure used in LLL-reduction to make a basis proper.

Definition 6 (θ-orthogonal). *A set of vectors* $\{b_1, \ldots, b_m\}$ *is* θ*-orthogonal if for* $i = 1, \ldots, m$*, the angle between* b_i *and the subspace spanned by* $\{b_1, \ldots, b_{i-1}, b_{i+1}, \ldots, b_m\}$ *lies in the range* $[\theta, \frac{\pi}{2}]$.

Neelamani, Dash and Baraniuk prove the following result.

Theorem 1. *[8] A* $\frac{\pi}{3}$*-orthogonal basis contains a shortest non-zero lattice vector.*

3 Main results

Theorem 2. *If* $\mathcal{B} = \{b_1, \ldots, b_m\}$ *is a* $\frac{\pi}{3}$*-orthogonal basis for a lattice* $\mathcal{L}$*, then some ordering of* $\mathcal{B}$ *is nearly KZ-reduced.*

Proof: Define $b_i' = \text{proj}_{\perp b_1}(b_i)$ for $i = 2, \ldots, m$. We will first prove that if $\mathcal{B}$ is a θ-orthogonal collection of vectorsfor some θ, then so is $\{b_2', \ldots, b_m'\}$. We will then invoke Theorem 1 to prove Theorem 2.

For any vector b_i with $i \geq 2$, we next show that the angle between b_i' and the subspace spanned by $\{b_2', \ldots, b_{i-1}', b_{i+1}', \ldots, b_m'\}$ lies in the range $[\theta, \frac{\pi}{2}]$. Let $\mathcal{C} = (c_1, c_2, \ldots, c_m)$ be an ordered basis obtained by permuting the columns of $\mathcal{B}$ such that $c_1 = b_1$ and $c_m = b_i$. Then $\mathcal{C}$ is θ-orthogonal. Define $c_i' = \text{proj}_{\perp c_1} c_i$ for $i = 2, \ldots, m$. Clearly, $(c_2', \ldots c_m')$ is a permutation of $(b_2', \ldots, b_m')$ with $c_m' = b_i'$. We can assume that $\mathcal{C} = (r_{ij})$ is an upper triangular matrix with non-zeros on the diagonal (by taking its QR decomposition and using the columns of Q to define the coordinate system). Let $\theta' \geq \theta$ be the angle between c_m and the subspace spanned by $c_1, \ldots, c_{m-1}$. Now

$$\sin^2(\theta') = \frac{r_{mm}^2}{\sum_{k=1}^{m} r_{km}^2}.$$

Let θ'' denote the angle between c_m' and the subspace spanned by $c_2', \ldots, c_{m-1}'$. The vectors $c_2', \ldots, c_m'$ are the same as the vectors $c_2, \ldots, c_m$ except for the first coordinate (i.e., r_{1i} is deleted from c_i). Therefore

$$\sin^2(\theta'') = \frac{r_{mm}^2}{\sum_{k=2}^{m} r_{km}^2} \geq \sin^2(\theta').$$

We conclude that $\theta'' \geq \theta' \geq \theta$. Therefore $\{b_2', \ldots, b_m'\}$ is θ-orthogonal.

Define a *length-ordered* basis to be one where (i) b_1 is a shortest basis vector, and (ii) $\mathcal{B}' = \{b_2', \ldots, b_m'\}$ is a length-ordered basis of the lattice spanned by its columns. Any basis can be length-ordered in polynomial time: just choose a shortest basis vector as the first, then recursively length-order the projections of the

others. Observe that $\frac{\pi}{3}$-orthogonality is independent of ordering; therefore the columns of a $\frac{\pi}{3}$-orthogonal basis can be permuted to obtain a length-ordered $\frac{\pi}{3}$-orthogonal basis.

We will prove by induction on the dimension of the lattice that a length-ordered $\frac{\pi}{3}$-orthogonal basis is nearly KZ-reduced. This statement is clearly true for one-dimensional lattices. Assume we have proved the statement for all $(m-1)$-dimensional lattices. Let $\mathcal{B} = (b_1, \ldots, b_m)$ be a length-ordered $\frac{\pi}{3}$-orthogonal basis of an m-dimensional lattice. Invoking Theorem 1, b_1 is a shortest lattice vector. By definition, $\mathcal{B}' = (b'_2, \ldots, b'_m)$ is length-ordered. Further, since $\mathcal{B}$ is $\frac{\pi}{3}$-orthogonal, so is $\mathcal{B}'$ (as proved earlier). By induction it forms a nearly KZ-reduced basis of the lattice generated by its columns. Therefore $\mathcal{B}$ also forms a nearly KZ-reduced basis. □

Therefore, a KZ-reduced basis can be obtained from a $\frac{\pi}{3}$-orthogonal basis in polynomial time.

In the following theorem, note that as θ increases the first hypothesis becomes stricter and the second one more relaxed.

Theorem 3. *Let $\mathcal{B} = \{b_1, \ldots, b_m\}$ be a θ-orthogonal basis for a lattice $\mathcal{L}$ with $\theta \geq \frac{\pi}{3}$. Further, suppose that for all $i \in \{1, 2, \ldots, m\}$,*

$$2\cos\theta \, \|b_i\| \leq \min_{j \in \{1,2,\ldots,m\}} \|b_j\| . \tag{1}$$

Then some ordering of the basis is Minkowski reduced.

Proof: Let $\mathcal{B}$ denote a basis that satisfies the conditions of Theorem 3. Assume $\mathcal{B}$ is ordered such that

$$\|b_1\| \leq \|b_2\| \leq \cdots \leq \|b_m\| ,$$

and $2\cos\theta \, \|b_m\| \leq \|b_1\|$. We can assume that $\|b_1\| = 1$. We will show that for any $S \subseteq \{1, 2, \ldots, m\}$,

$$\left\| \sum_{i \in S} u_i b_i \right\| \geq \max_{i \in S} \|b_i\| , \tag{2}$$

if $u_i \in \mathbb{Z}$ and $|u_i| \geq 1$ for all $i \in S$. Consider a subset of indices S with k being the largest index in S. Let $v = \sum_{i \in S, i \neq k} u_i b_i$, where $u_i \in \mathbb{Z}$ and $|u_i| \geq 1$ for all $i \in S \setminus \{k\}$, and let θ' be the angle between b_k and v. Define $l_1 = \|v\|$ and $l_2 = \|b_k\|$. Obviously $l_1, l_2 \geq \|b_1\| = 1$ and $\theta' \geq \theta$. We will show that for any non-zero integer u_k

$$\left\| \sum_{i \in S} u_i b_i \right\|^2 - \|b_k\|^2 \geq 0. \tag{3}$$

Observe that

$$\|v + u_k b_k\|^2 = (l_1 \pm u_k l_2 \cos\theta')^2 + (u_k l_2 \sin\theta')^2 =$$
$$l_1^2 + u_k^2 l_2^2 \pm 2 u_k l_1 l_2 \cos\theta' \;\geq\; l_1^2 + u_k^2 l_2^2 - |u_k| l_1,$$

as $2 l_2 \cos\theta' \leq 1$. Therefore

$$\|v + u_k b_k\|^2 - \|b_k\|^2 \geq l_1^2 + u_k^2 l_2^2 - |u_k| l_1 - l_2^2. \tag{4}$$

As $l_1 \geq 1$, the right hand side in (4) is non-negative for $|u_k| = 1$. Further

$$l_1^2 + u_k^2 l_2^2 - |u_k| l_1 - l_2^2 \;\geq\; l_1^2 + u_k^2 l_2^2 - |u_k| l_1 l_2 - l_2^2 =$$
$$(l_1 - \frac{1}{2}|u_k| l_2)^2 + \frac{3}{4} u_k^2 l_2^2 - l_2^2.$$

If $|u_k| \geq 2$ then $u_k^2 \geq 4$ and $\frac{3}{4}u_k^2 l_2^2 - l_2^2 > 0$. This proves (3) and therefore (2).

We now show that for $i = 2, 3, \ldots, m$, b_i is the shortest vector in $\mathcal{L}$ not contained in the span of $\{b_1, \ldots, b_{i-1}\}$. Assume that this is not true. Then there exists some set of indices S not entirely contained in $\{1, 2, \ldots, i-1\}$, and non-zero integers $u_i, i \in S$, such that $v = \sum_{i \in S} u_i b_i$ has length less than $\|b_i\|$. But if k is the maximum index in S, then $k \geq i$ and therefore, $\|b_k\| \geq \|b_i\|$. But (3) implies that $\|v\| \geq \|b_k\|$, which is a contradiction. This proves that $\mathcal{B}$ is Minkowski-reduced. □

The result is easy to show for $\theta = \frac{\pi}{2}$ and $\theta = \frac{\pi}{3}$. In the first case, if the basis vectors are ordered by increasing lengths, than the ordered basis is Minkowski-reduced. In the second case, $2\cos\theta = 1$, and all the basis vectors have lengths equal to $\lambda(\mathcal{L})$ and the basis is therefore Minkowski-reduced.

4 Random lattices

For a given probability distribution of vectors in $\mathbb{R}^n$, we define an m-dimensional random lattice – with $m \leq n$ – as one which is generated by m i.i.d. vectors drawn from the distribution. The distributions we will consider in this paper are: uniform over the unit ball in $\mathbb{R}^n$; the vector entries are standard normal random variables; the vector entries are Bernoulli random variables with success probability 1/2. When $m \leq n$, m vectors drawn from the above distributions are almost surely linearly independent as $n \to \infty$ and define a basis of the generated lattice. We will also consider lattices generated by columns of an $n \times m$ matrix drawn uniformly from the unit ball in $\mathbb{R}^{nm}$; the columns of this matrix form a basis almost surely as $n \to \infty$. The ideas and techniques we use here are similar to, and draw upon those in Daudé and Vallée [3] and Akhavi, Marckert and Rouault [2]. However, all our results are for the case $m < n$, and thus apply only to non full-dimensional lattices.

Neelamani, Dash and Baraniuk [8] showed that lattice bases defined by m vectors drawn from the first three distributions above are almost surely $\frac{\pi}{3}$-orthogonal as n tends to infinity with $m \leq 0.071n$. We improve the ratio of m to n in the case of the first two distributions based on the following lemma.

Lemma 4. *Let* $\mathbf{X}$ *be a vector* $(X_1, \ldots, X_n)$ *whose entries are independent standard normal random variables,* $X_i \sim N(0,1)$. *Let* $c \in (0,1)$. *As* $n \to \infty$, *the (random) angle between* $\mathbf{X}$ *and the subspace* $\mathbf{L}$ *spanned by* m *similar, independent random vectors* $\mathbf{X}_1, \ldots, \mathbf{X}_m$ *with* $m = \lfloor cn \rfloor$ *is almost surely*

$$\tan^{-1}\left(\sqrt{\frac{1}{c} - 1}\right).$$

Proof: $\mathbf{L}$ is almost surely m-dimensional. It is well known that the squared length of $\mathbf{X}$ has chi-squared distribution χ_n^2, and its orientation is uniform over a unit sphere in $\mathbb{R}^n$. By the latter property, the (random) angle between $\mathbf{X}$ and $\mathbf{L}$ is identically distributed to the angle between $\mathbf{X}$ and any fixed m-dimensional subspace $\mathbf{L}'$. For convenience, let $\mathbf{L}'$ be the subspace spanned by the m unit vectors $(1, 0, 0, \ldots)$, $(0, 1, 0, \ldots)$ *etc.*

The projection of $\mathbf{X}$ onto $\mathbf{L}'$ is

$$\mathbf{X}' = (X_1, \ldots, X_m, 0, \ldots, 0),$$

and the orthogonal component of $\mathbf{X}$ with respect to $\mathbf{L}'$ is

$$\mathbf{X}'' = (0, \ldots, 0, X_{m+1}, \ldots, X_n).$$

The angle ϕ between $\mathbf{X}$ and $\mathbf{L}'$ is precisely the angle between $\mathbf{X}$ and $\mathbf{X}'$, which is given by $\tan(\phi) = \|\mathbf{X}''\| \,/\, \|\mathbf{X}'\|$. Immediately,

$$\phi = \tan^{-1}\left(\sqrt{L_{n-m}^2/L_m^2}\right), \tag{5}$$

where L_{n-m}^2 and L_m^2 are a pair of *independent* chi-squared random variables representing the squared lengths of X'' and X' respectively.

If n and m are both large (for example if $m = \lfloor cn \rfloor$ and $n \to \infty$) then both L_{n-m}^2 and L_m^2 are concentrated random variables, and their ratio will also be concentrated. (A chi-squared random variable with k degrees of freedom has expectation k and variance $2k$.) Stronger statements are possible, but a simple concentration inequality (see Appendix A) is that for any fixed $\epsilon > 0$, as $k \to \infty$,

$$\mathbb{P}\left((1-\epsilon)k < L_k^2 < (1+\epsilon)k\right) \geq 1 - \exp(-\Omega(k)).$$

Here $\Omega(f(k))$ stands for a function greater than $cf(k)$ for some constant $c > 0$. Thus, with $m = \lfloor cn \rfloor$ and $n \to \infty$,

$$\mathbb{P}\left((1-\epsilon)^2 < \left[L_{n-m}^2/L_m^2\right]/\left[(n-m)/m\right] < (1+\epsilon)^2\right) \geq 1 - \exp(-\Omega(n)).$$

From (5) and because $\tan^{-1}$ has bounded derivative, setting

$$\bar{\phi} := \tan^{-1}\left(\sqrt{n/m - 1}\right) = \tan^{-1}\left(\sqrt{\frac{1}{c} - 1}\right), \tag{6}$$

we have that for any $\epsilon > 0$,

$$\mathbb{P}\left((1-\epsilon)\bar{\phi} < \phi < (1+\epsilon)\bar{\phi}\right) \geq 1 - \exp(-\Omega(n)). \tag{7}$$

□

As a corollary, let $\mathbf{X}_1, \ldots, \mathbf{X}_m$ be i.i.d. vectors as above, and let ϕ_i be the angle formed between $\mathbf{X}_i$ and the subspace spanned by the other vectors. Then the inequality (7) applies simultaneously to the entire collection of angles ϕ_i in place of the single angle ϕ. This is easy to see. For each i, the probability that ϕ_i violates the inequality is $\exp(-\Omega(n))$, and by the union bound, the probability that *any* angle violates it is at most $m\exp(-\Omega(n))$. From $m < n$, this is again $\exp(-\Omega(n))$. Further, solving for $\tan^{-1}\left(\sqrt{1/c - 1}\right) = \sqrt{3}$, one gets $c = 1/4$. Therefore, for any constant $c' < 1/4$, matrices with $\lfloor c'n \rfloor$ i.i.d. columns are $\frac{\pi}{3}$-orthogonal almost surely as n tends to infinity.

In the results above, if $m \leq cn$ and does not tend to ∞, then the random variable L_m^2 is not a concentrated random variable. However, the proof in Appendix A can be easily modified to show that $L_m^2 \leq cm(1+\epsilon)$ with probability greater than $1 - \exp(-\Omega(m))$. Then, if $\mathbf{X}_1, \ldots, \mathbf{X}_m$ are i.i.d. vectors as in the previous Lemma, they form a $\bar{\phi}$-orthogonal collection almost surely, except that the angle between any vector and the subspace spanned by the remaining vectors will not be concentrated around $\bar{\phi}$, but will be larger than $\bar{\phi}$ almost surely. We have thus proved the following result.

Theorem 5. *Let $c \in (0,1)$. An $n \times m$ matrix whose entries are independent standard normal random variables is almost surely $\bar{\phi}$-orthogonal, with $\bar{\phi}$ given by (6), as $n \to \infty$ and $m \leq cn$. If $c < 1/4$, then the $n \times m$ matrix is almost surely $\frac{\pi}{3}$-orthogonal.*

Definition 7. *We call the following distributions for $n \times m$ matrices* simple Gaussian distributions.
(i) All entries of the matrix are i.i.d. standard normal random variables.
(ii) Each column of the matrix is chosen independently and uniformly from the unit ball in $\mathbb{R}^n$.
(iii) The matrix is chosen uniformly from the unit ball in $\mathbb{R}^{nm}$.

A matrix B drawn from any of the three distributions above has the following property: given a constant $c \in (0,1)$, as $n \to \infty$ with $m \leq cn$, for any $\epsilon \in (0,1)$ almost surely every pair of columns b_i and b_j of B satisfy $\|b_i\| \leq (1+\epsilon)\|b_j\|$. We explicitly prove this property for distribution (i) in Definition 7 in the appendix. For the second distribution, all columns of B have length at most 1. Further, with probability at most $\exp(-\Omega(n))$, a vector drawn uniformly from the unit ball in $\mathbb{R}^n$ has length $1/(1+\epsilon)$ or less (most of the volume of a high-dimensional unit ball is concentrated near its boundary). Therefore, by the union bound, the probability that any column of B has length less than $1/(1+\epsilon)$ is at most $\exp(-\Omega(n))$. Finally, for the third distribution, it is easy to see that the columns of B behave (as $n \to \infty$) like i.i.d. vectors drawn from the ball with radius $1/\sqrt{m}$.

Theorem 6. *Let $\mathcal{B}$ denote an $n \times m$ matrix drawn from a simple Gaussian distribution. If $m \leq cn$, where $c < 1/4$, then as $n \to \infty$, some ordering of $\mathcal{B}$ forms a Minkowski reduced basis of the lattice generated by its columns. If $c < 1/5$, then some ordering of $\mathcal{B}$ forms a KZ-reduced basis almost surely.*

Proof: Assume $c < 1/4$. Then for some fixed $\epsilon > 0$, $\mathcal{B}$ is almost surely $(\frac{\pi}{3} + \epsilon)$-orthogonal. Let

$$1 + \epsilon' = \frac{1}{2\cos(\frac{\pi}{3} + \epsilon)}.$$

We know, based on the earlier discussion, that almost surely the column-lengths of $\mathcal{B}$ are within a factor of $(1+\epsilon')$ of one another. Applying Theorem 3, we conclude that some ordering of $\mathcal{B}$ is almost surely Minkowski reduced.

Assume $c < 1/5$. Then for some fixed $\epsilon > 0$, $\mathcal{B}$ is almost surely $(\phi + \epsilon)$-orthogonal, where $\tan(\phi) = \sqrt{1/c - 1} = 2$. Also, almost surely the column-lengths of $\mathcal{B}$ are within a factor of $(1+\epsilon')$ of one another, for any $\epsilon' > 0$. Let $\mathcal{B}'$ be obtained by arbitrarily permuting columns of $\mathcal{B}$; it is also $(\phi+\epsilon)$-orthogonal. Consider its decomposition $\mathcal{B}' = QR$, where $R = (r_{ij})$ is an $m \times m$ upper triangular matrix. Let the columns of $\mathcal{B}'$ be $a_1, \ldots, a_m$. Let θ_i be the angle between a_i and the subspace spanned by the preceding columns of $\mathcal{B}'$.

For any $i, j \in 1, \ldots, m$, $|r_{ii}| \geq \|a_i\| \sin\theta_i$, while $|r_{ij}| \leq \|a_j\| \cos\theta_j$. As $\theta_i, \theta_j \geq \phi + \epsilon$,

$$\frac{|r_{ij}|}{|r_{ii}|} \leq \frac{\|a_j\| \cos\theta_j}{\|a_i\| \sin\theta_i} \leq \frac{\|a_j\| \cos(\phi+\epsilon)}{\|a_i\| \sin(\phi+\epsilon)} < \frac{1}{2},$$

for appropriately chosen ϵ. Therefore $\mathcal{B}'$ is a proper matrix. We can assume (by re-ordering columns of $\mathcal{B}$) that $\mathcal{B}'$ is length-ordered; then it is also KZ-reduced almost surely. □

Note that the matrices satisfying the conditions of the previous theorem are almost surely LLL-reduced; this follows from the fact that they are KZ-reduced. Donaldson[4] proved that matrices drawn from the third simple Gaussian distribution discussed above are almost surely Minkowski reduced as $n \to \infty$ but m is fixed. Thus Theorem 6 is a significant extension of his result. As we mentioned earlier, Akhavi, Marckert and Rouault proved that matrices drawn from the second simple Gaussian distribution satisfy the size-reduction property of the LLL-algorithm even with many more columns (they only require that $n - m \to \infty$). Our results thus imply that if the number of columns is reduced then much stronger reduction properties can be proved.

Finally, observe that matrices with ± 1 entries drawn from a Bernoulli distribution with success probability 1/2 have equal-length columns. Therefore the following is an immediate consequence of the $\frac{\pi}{3}$-orthogonality of such matrices for $m \leq 0.071n$ as $n \to \infty$.

Theorem 7. *An $n \times m$ matrix with ± 1 entries drawn from a Bernoulli distribution with success probability 1/2 is almost surely Minkowski reduced as $n \to \infty$ with $m \leq 0.071n$.*

A Column Lengths of a Gaussian Random Basis

We will prove that given an $\epsilon > 0$, all columns of an $n \times m$ random matrix $\mathcal{B} = (b_1, \ldots, b_m)$ whose entries are i.i.d. Gaussian random variables with mean 0 and variance 1 satisfy

$$n(1-\epsilon) \leq \|b_i\|_2^2 \leq n(1+\epsilon) \tag{8}$$

with probability at least $1 - c_1 m\sqrt{n}e^{-c_2 n}$ for large n. During the proof, w.l.o.g. we will assume that $n = 2k$, for some integer k.

Let b be a vector in $\mathbb{R}^n$, whose components i.i.d. Gaussian random variables with mean 0 and variance 1. Then $\|b\|^2$ has a *chi-squared* distribution with $2k$ degrees of freedom. Its probability density function (PDF) is given by

$$f(x; 2k) = \frac{(1/2)^k e^{-x/2} x^{k-1}}{(k-1)!}$$

with $(k-1)!$ denoting $(k-1)$ factorial. First, note that

$$f(2k; 2k) = \frac{1}{2}\frac{e^{-k}k^{k-1}}{(k-1)!} = \frac{1}{2}\frac{e^{-k}k^k}{k!}$$

By Stirling's formula, for large k

$$k! \approx e^{-k}k^k\sqrt{2\pi k} \Rightarrow f(2k; 2k) \approx \frac{1}{2\sqrt{2\pi k}} < 1.$$

The approximate equality in Stirling's formula can be replaced by a more precise inequality [10]

$$k! > e^{-k}k^k\sqrt{2\pi k}\, e^{1/(12k+1)} \forall k \geq 1,$$

to infer that $f(2k; 2k) < \frac{1}{\sqrt{k}}$ for large k. Define

$$g(x) = \frac{f(x; 2k)}{\sqrt{k} f(2k; 2k)} = \frac{1}{\sqrt{k}} e^{\frac{-x+2k}{2}} \left(\frac{x}{2k}\right)^{k-1}.$$

Then $f(x; 2k) < g(x)$. Therefore,

$$\begin{aligned} P\left(\|b_i\|^2 > 2k(1+\epsilon)\right) &= \int_{2k(1+\epsilon)}^{\infty} f(x; 2k)dx \\ &\leq \int_{2k(1+\epsilon)}^{\infty} g(x)dx \\ &= \int_{2k(1+\epsilon)}^{\infty} \frac{1}{\sqrt{k}} e^{\frac{-x+2k}{2}} \left(\frac{x}{2k}\right)^{k-1} dx. \end{aligned}$$

Substituting x by $2k(1+\delta)$, we get

$$\int_{2k(1+\epsilon)}^{\infty} g(x)dx = \int_{\epsilon}^{\infty} 2\sqrt{k}e^{-k\delta}(1+\delta)^{k-1}d\delta. \tag{9}$$

To obtain bounds on the above integral, we derive a bound on $1+x$. We know that $1+x < e^x$ for $x > 0$. Thus $1+\epsilon < e^{\epsilon}$. Choose some $r > 0$ such that $1+\epsilon = e^{\epsilon}/e^r = e^{\epsilon - r}$. Then

$$1 + x \leq e^{x-r}, \qquad \forall\, x \geq \epsilon.$$

This is because $1 + x = e^{x-r}$ at $x = \epsilon$ and $\frac{d}{dx}(1+x) < \frac{d}{dx}e^{x-r}$ for all $x \geq \epsilon$. Similarly, one can show that $1 - x < e^{-x}$ for $x \geq 0$, and there is some positive $s > 0$ such that

$$1 - x \leq e^{-x-s}, \qquad \forall\, x \geq \epsilon.$$

Therefore

$$(1+\delta)^{k-1} \leq e^{(\delta - r)(k-1)}, \qquad \forall\, \delta \geq \epsilon.$$

Substituting this upper bound in (9), we get

$$\begin{aligned} \int_{\epsilon}^{\infty} 2\sqrt{k}e^{-k\delta}(1+\delta)^{k-1}d\delta &\leq \int_{\epsilon}^{\infty} 2\sqrt{k}e^{-r(k-1)}e^{-\delta}d\delta \\ &= 2ke^{-\epsilon}e^{-r(k-1)}. \end{aligned}$$

Similarly the probability that $\|b\|^2 < 2k(1-\epsilon)$ is bounded above by

$$\int_{0}^{2k(1-\epsilon)} g(x)dx = \int_{\epsilon}^{1} 2\sqrt{k}g(2k(1-\delta))d\delta = \int_{\epsilon}^{1} 2\sqrt{k}e^{k\delta}(1-\delta)^{k-1}d\delta, \tag{10}$$

where the last two terms are obtained by the substitution of variables $x = 2k(1-\delta)$ for δ between ϵ and 1. We can use the fact that $1 - \delta \leq e^{-\delta - s}$ to bound the integral in (10) by

$$\int_{\epsilon}^{1} 2\sqrt{k}e^{k\delta}e^{-(\delta+s)(k-1)}d\delta = \int_{\epsilon}^{1} 2\sqrt{k}e^{-s(k-1)}e^{\delta}d\delta = 2ke^{-s(k-1)}(e - e^{\epsilon}).$$

Therefore, there are positive constants c_1, c_2 such that

$$P\left(\|b\|^2 \notin [n(1-\epsilon), n(1+\epsilon)]\right) \leq c_1\sqrt{n}e^{-c_2 n}. \tag{11}$$

Let $\mathcal{I} = [n(1-\epsilon), n(1+\epsilon)]$. It follows from (11) that

$$P\left(\|b_i\|^2 \notin \mathcal{I}\right) \leq c_1\sqrt{n}e^{-c_2 n}, \qquad \text{for } i = 1, 2, \ldots, m.$$

Consequently,

$$P\left(\bigcup_{i=1}^{m} \left\{\|b_i\|^2 \in \mathcal{I}\right\}\right) > 1 - c_1 m\sqrt{n}e^{-c_2 n}.$$

The expression on the right hand side of the above inequality is $1 - \exp(-\Omega(n))$, as long as $m = O(n^k)$ for any fixed positive integer k.

References

[1] A. Akhavi. Random lattices, threshold phenomena and efficient reduction algorithms, *Theoretical Computer Science* **287** (2002), 359–385.

[2] A. Akhavi, J. F. Marckert, and A. Rouault, On the reduction of a random basis, Proceedings of SIAM-ALENEX/ANALCO07, New Orleans, January 07.

[3] H. Daudé and B. Vallée, An upper bound on the average number of iterations of the LLL algorithm, *Theoretical Computer Science* **123** (1994), 95–115.

[4] J. L. Donaldson, Minkowski reduction of integral matrices, *Mathematics of Computation* **33**(no. 145) (1979), 201–216.

[5] R. Kannan, Algorithmic geometry of numbers, *Annual Review of Comp. Sci.* **2** (1987), 231–267.

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

[7] L. Lovász, An algorithmic theory of numbers, graphs and convexity, CBMS-NSF REgional conference series in applied mathematics, SIAM (1986).

[8] R. Neelamani, S. Dash, and R. G. Baraniuk, On nearly orthogonal lattice bases and random lattices, *SIAM Journal on Discrete Mathematics*, vol. 21 , no. 1, pp. 199–219, Feb. 2007.

[9] P. Nguyen, D. Stehle. LLL on the average, *Proceedings of the 7th Algorithmic Number Theory Symposium (ANTS VII)*, LNCS **4076**, pp. 238–256, Springer, (2006).

[10] H. Robbins, A remark on Stirling's Formula *Amer. Math. Monthly* **62** (1955), 26–29.