An Elementary Introduction to Smoothing Parameter Bounds

Pritam Chandra

Ashoka University, Sonipat, India, 131029

Abstract

We exposit the essential background in the theory of lattices that builds up to a relation between the *smoothing parameter* and the *Gram Schmidt basis length* of a lattice, a result crucial for defining Gaussian distributions on lattices [3]. Apart from elementary linear algebra and some standard results in Fourier analysis, this note is self–contained.

Table of contents

Conventions and useful results. Throughout this note, lattices under discussion will be sitting inside $\mathbb{R}^n$. Vectors will be columns. A matrix with linearly independent columns will be called full rank.

When an $m \times n$ matrix is denoted by any upper case, say A, its columns will automatically be assumed to be the lower cases $a_1, \ldots, a_n$, unless mentioned otherwise. In such cases A could denote either the matrix or the set of column vectors $\{a_1, \ldots, a_n\}$ as established from the context. Thus in this case the subspace $\operatorname{ran} A$ is same as $\operatorname{span}\{a_1, \ldots, a_n\}$.

An $n \times n$ matrix U will be called unitary if $U^T U = I$.

Listed below are some basic results from linear algebra which will be used in several arguments.

1. An $n \times n$ matrix A is 0 if and only if $\langle x, Ay \rangle = 0$ for all $x, y \in \mathbb{R}^n$.

2. For any $m \times n$ matrix A, we have $(\operatorname{ran} A)^{\perp} = \ker A^T$.

3. Let A, B be two rectangular matrices such that AB is defined. Then $AB = 0$ if and only if $\operatorname{ran} B \subset \ker A$.

4. Let P be the projection onto the subspace $S \subset \mathbb{R}^n$. Then $\operatorname{ran} P = S$ and $\ker P = S^{\perp}$. Further, $P^T = P$, and $P^2 = P$.

5. Let P and Q be two orthoprojections onto the spaces S and T respectively. Then $PQ = 0$ if and only if $S \cap T = \{0\}$.

The letters x, y, z will usually denote lattice vectors and integer vectors, and u, v, w points in $\mathbb{R}^n$.

1 The dual lattice

The dual space of an n dimensional vector space $\mathcal{H}$ is the set $\mathcal{H}^* = \{f |\ f: \mathcal{H} \to \mathbb{R} \quad \text{linear}\}$, that is, the space of all linear functionals on $\mathcal{H}$. If $\mathcal{H}$ is a Hilbert space (a vector space equipped with an inner product), then every linear functional f is of the form

$$f(w) = \langle \varphi, w \rangle \quad \text{for all } w, \tag{1}$$

where $\varphi = [\ f(e_1) \quad \cdots \quad f(e_n)\ ]^T$. This says that a linear functional f can be represented by a vector φ in V. Conversely, every vector $\varphi \in \mathbb{R}^n$ gives rise to a linear functional on $\mathcal{H}$ via equation (1). Thus, in this case, the dual of $\mathcal{H}$ is bijective with $\mathcal{H}$ itself. We carry forward this intuition to lattices.

A lattice is a discrete additive subgroup of $\mathbb{R}^n$. The lattices under discussion are "like" vector spaces, but allowing linear combinations with coefficients only from $\mathbb{Z}$. The name for such a space built over a ring (generally $\mathbb{Z}$) is a module. Given a set of linearly independent vectors $A = \{a_1, \ldots, a_n\}$, the lattice generated by A is denoted by $\mathcal{L}(A)$, where[1]

$$\mathcal{L}(A) = \{\alpha_1 a_1 + \cdots + \alpha_n a_n |\ \alpha_j \in \mathbb{Z}\}.$$

Definition 1. Let $L = \mathcal{L}(B)$ be a lattice. The **dual lattice** L^* is defined as

$$L^* = \{w \in \operatorname{span}(B) \,|\, \langle w, x \rangle \in \mathbb{Z} \quad \text{for all } x \in L\}.$$

Thus the dual lattice L^* is an analogue of the dual space in the sense that L^* is the space of all integer valued linear functinoals on L. Pick any $y \in L^*$. Let $L_{y,i} = \{x \in L |\ \langle x, y \rangle = i\}$. Then y partitions L into "layers" as follows

$$L = \bigcup_{i \in \mathbb{Z}} L_{y,i}.$$

1. It is useful to consider lattices generated by $m < n$ linearly independent vectors stacked into the $n \times m$ column matrix A. The lattice $\mathcal{L}(A)$ would now be a discrete additive subgroup of the m-dimensional space $\operatorname{ran} A$. Restricting to this space one can find an $m \times m$ matrix B such that $\mathcal{L}(A)$ is isomorphic to $\mathcal{L}(\mathrm{B})$. Therefore, for our discussion it is sufficient to study lattices which are generated by a full-rank square matrix. We call such lattices full-rank.

In particular $L_{y,0} = L \cap y^{\perp}$. The layer $L_{y,0}$ always contains the origin. However, for $i \neq 0$, the layer $L_{y,i}$ maybe empty. Note that $L_{y,i}$ sits inside the affine subspace $\{w \in \mathbb{R}^n | \langle w, y \rangle = i\}$. Then the distance between two layers $L_{y,i}$ and $L_{y,j}$ is defined as the distance between the corresponding affine subspaces.[2]

Pick any w such that $\langle w, y \rangle = 1$. Then for any $u \in y^{\perp}$, by Cauchy-Schwarz inequality, we have

$$1 = \langle w, y \rangle = \langle w, y - u \rangle \leq \|w\| \, \|y - u\| \Longrightarrow \frac{1}{\|w\|} \leq \|y - u\| \, .$$

Further, $w = {}^{y}/_{\|y\|^2}$ satisfies $\langle w, y \rangle = 1$, and w is orthogonal to $y^{\perp}$. So, the distance of w from $y^{\perp}$ is exactly $\|w\| = {}^{1}/_{\|y\|}$. Thus, this vector w achieves equality in the above inequality. This shows that

$$\mathrm{dist}(L_{y,0}, L_{y,1}) = \frac{1}{\|y\|} \, .$$

A similar calculation shows that for $i \neq j$,

$$\mathrm{dist}(L_{y,i}, L_{y,j}) = \frac{|i - j|}{\|y\|}.$$

Heuristically, this says that if we can produce a vector $y \in L^*$ such that $\|y\|$ is small, then the lattice L is spread out, owing to a large distance between the layers. Conversely, if $\|y\|$ is large, then the lattice L is closely packed.[3] The idea gives a peek into an expansive theory of relating the density of a lattice to the size of vectors in its dual.

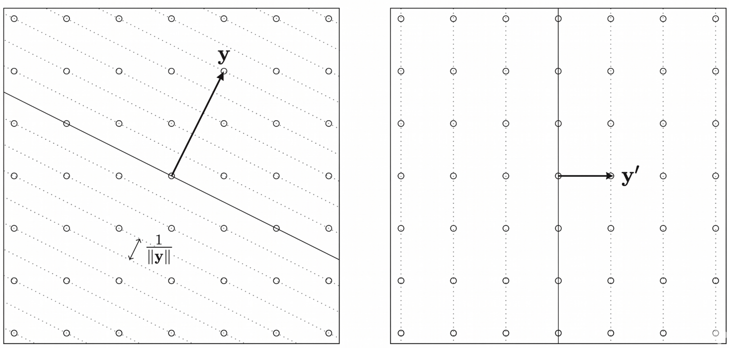

Figure 1. In this figure $L = L^* = \mathbb{Z}^2$. We look at the layers generated by the dual vector $y = (1, 2)$ and the dual vector $y' = (1, 0)$.

2. For two subsets S, T of some metric space, the distance between S and T is defined as the shortest distance between a point taken from S and a point taken from T. That is,

$$\mathrm{dist}(S, T) = \inf_{s \in S, t \in T} \|s - t\|.$$

3. As such one can always find $y \in L^*$ such that $\|y\|$ is arbitrarily large. But this will force many of the layers to be empty, thereby not ensuring short vectors. For example, consider $L = \mathbb{Z}^2$ which is its own dual. If $y = [\, n \;\; 0 \,]^T$, then $L_{y,i}$ is non-empty only when i is a multiple of n. However, if one picks w from some basis set of L^*, then every layer is non-empty.

Definition 2. Suppose $L=\mathcal{L}(B)$ be a full rank lattice in $\mathbb{R}^n$. Then we denote the set $\{Bx: \|x\|_\infty = 1\} = B\,[0,1]^n$ by $\mathcal{P}(B)$, and call it **the funadamental parallelepiped** of L with respect to the basis B.

Definition 3. The determinant of the lattice L, denoted by $\det L$, is the volume of $\mathcal{P}(B)$. This is also $[\det(B^T B)]^{\frac{1}{2}}$. Even though $\mathcal{P}(B)$ is dependent on the basis B, the quanity $\det L$ is a property of the lattice, and therefore, invariant under change of the lattice basis.

Theorem 1. *Suppose B is a basis of the lattice $L\subset\mathbb{R}^n$. Let D be a full rank matrix such that*

$$\operatorname{span}(B)=\operatorname{span}(D) \quad \textit{and} \quad B^T D = I.$$

In this case we say D is the dual of B. Then D is a basis of L^*.

Proof. Pick any $Dz\in\mathcal{L}(D)$ and $Bx\in L$. Then $\langle Bx, Dz\rangle = \langle x, z\rangle \in \mathbb{Z}$ as both x and z have integral coordinates. This shows that $Dz\subset L^*$ and thus $\mathcal{L}(D)\subset L^*$.

Conversely, suppose $y\in L^*$. Recall that L^* and L lie in the same vector space which is spanned by B, and also by D. So, there exists $w\in\mathbb{R}^n$ such that $y=Dw$. And

$$y = Dw = DB^T Dw = D(B^T y)\in\mathcal{L}(D),$$

as $B^T y\in\mathbb{Z}^n$. Thus $L^*\subset\mathcal{L}(D)$. □

From this theorem we immediately derive the following corollaries.

Corollary 1. *If B is a basis of the lattice L, then the Moore-Penrose inverse $B^+ = B\,(B^T B)^{-1}$ of B is a basis of L^*. If L is full rank, then B^{-T} is a basis of L^*.*

Corollary 2. *For any lattice L, the dual of L^* is L. Succintly, $(L^*)^* = L$.*

Corollary 3. *For $\alpha\neq 0$, the dual of αL is $\frac{1}{\alpha} L^*$.*

Corollary 4. *Suppose B and D are duals of each other. Then for any unitary U, the matrices UB and UD are duals of each other, and therefore, the dual of the lattice UL is the lattice UL^*.*

Corollary 5. *For any lattice L we have $\det L = \frac{1}{\det L^*}$.*

Proof. Recall that $\det L = \sqrt{\det(B^T B)}$ where B is some basis of L. We know that B^+ is a basis of L^*. And

$$(B^+)^T B^+ = (B^T B)^{-T} B^T B\,(B^T B)^{-1} = (B^T B)^{-1}.$$ □

Definition 4. Let B be any full (column) rank $m \times n$ matrix. Let P_i be the projection onto $\operatorname{span}\{b_1, \ldots, b_{i-1}\}^{\perp}$. Let $\tilde{b_1} = b_1$ and $\tilde{b_i} = P_i b_i$. Let $\tilde{B}$ be the matrix whose columns are $\tilde{b_i}$. Then $\tilde{B}$ is called the **Gram-Schmidt orthogonalization (GSO)** of B.

GSO can equivalently be formulated as a matrix decomposition. That is, given any full rank $m \times n$ matrix B, there exists a $m \times n$ matrix $\tilde{B}$ with orthogonal columns, and a $n \times n$ upper triangular matrix R with diagonal entries 1s, such that $B = \tilde{B} R$. The conventional GSO can be extracted by equating the individual columns of both sides.

Proposition 1. *Suppose B and D are two linearly independent collections in $\mathbb{R}^n$ which are duals of each other. For $i \leq j$, let P_i and Q_j be the projections onto $\operatorname{span}\{b_1, \ldots, b_{i-1}\}^{\perp}$ and $\operatorname{span}\{d_{j+1}, \ldots, d_n\}^{\perp}$ respectively. Then $P_i \begin{bmatrix} b_i & \cdots & b_j \end{bmatrix}$ and $Q_j \begin{bmatrix} d_i & \cdots & d_j \end{bmatrix}$ are duals of each other.*

Proof. Let B_{ij} be the $n \times (j - i + 1)$ matrix whose column representation is $\begin{bmatrix} b_i & \cdots & b_j \end{bmatrix}$. Similarly, define D_{ij}. We are required to show that $P_i B_{ij}$ is the dual of $Q_j B_{ij}$.

Define $S_i = \operatorname{span}\{b_1, \ldots, b_{i-1}\}$ and $T_j = \operatorname{span}\{d_{j+1}, \ldots, d_n\}$. Let $\bar{P}_i = I - P_i$ and $\bar{Q}_j = I - Q_j$. Then $P_i, \bar{P}_i, Q_j, \bar{Q}_j$ are projections onto $S_i^{\perp}, S_i, T_j^{\perp}$ and T_j respectively. Note that

$$\begin{aligned}
(P_i B_{ij})^T (Q_j D_{ij}) &= B_{ij}^T P_i Q_j D_{ij} \\
&= B_{ij}^T D_{ij} - B_{ij}^T \bar{P}_i D_{ij} - B_{ij}^T \bar{Q}_j D_{ij} + B_{ij}^T \bar{P}_i \bar{Q}_j D_{ij} \\
&= B_{ij}^T D_{ij} \\
&= I.
\end{aligned}$$

Here the last equality follows by noting that $\langle b_k, d_l \rangle = \delta_{k,l}$, and the second last inequality holds because $\bar{P}_i D_{ij}, B_{ij}^T \bar{Q}_j, \bar{P}_i \bar{Q}_j$ are all 0. To see this note that for any $x, y \in \mathbb{R}^n$ we have

$$\langle x, \bar{P}_i D_{ij} y \rangle = \langle \bar{P}_i x, D_{ij} y \rangle = \sum_{k=1}^{i-1} \sum_{l=i}^{j} \alpha_k y_l \langle b_k, d_l \rangle = 0,$$

where α_ks are appropriate real numbers obtained by expanding $\bar{P}_i x$ with respect to $b_1, \ldots, b_{i-1}$. A similar calculation shows that $B_{ij}^T \bar{Q}_j, \bar{P}_i \bar{Q}_j$ are also 0.

Now, it remains to show that $\operatorname{ran} P_i B_{ij} = \operatorname{ran} Q_j D_{ij}$. We immediately see that

$$\operatorname{ran} P_i B_{ij} \subset \operatorname{ran} P_i = S_i^{\perp}.$$

We also note using our previous calculations that $B_{ij}^T P_i \bar{Q}_j = B_{ij}^T \bar{Q}_j - B_{ij}^T \bar{P}_i \bar{Q}_j = 0$. This shows

$$\operatorname{ran} \bar{Q}_j = T_j \subset \ker B_{ij}^T P_i = (\operatorname{ran} P_i B_{ij})^{\perp} \Longrightarrow \operatorname{ran} P_i B_{ij} \subset T_j^{\perp}.$$

This shows $\operatorname{ran} P_i B_{ij} \subset S_i^{\perp} \cap T_j^{\perp} = (S_i + T_j)^{\perp}$. Now $\bar{P}_i \bar{Q}_j = 0$ means $S_i \cap T_j = \{0\}$. So,

$$\dim (S_i + T_j)^{\perp} = n - \dim (S_i + T_j) = n - (i-1) - (n-j) = j - i + 1.$$

We also claim that the columns of $P_i B_{ij}$ are linearly independent. Indeed

$$\begin{array}{rrcl} \beta_i P_i b_i + \cdots + \beta_j P_i b_j = 0 \implies & P_i(\beta_i b_i + \cdots + \beta_j b_j) & = & 0 \\ \implies & \beta_i b_i + \cdots + \beta_j b_j & \in & S_i \\ \implies & \beta_i = \cdots = \beta_j & = & 0, \end{array}$$

where the final implication follows from the linear independence of b_js. So dim $\operatorname{ran} P_i B_{ij} = j - i + 1$. Therefore, we conclude $\operatorname{ran} P_i B_{ij} = (S_i + T_j)^{\perp}$. The symmetric argument shows that $\operatorname{ran} Q_j D_{ij} = (S_i + T_j)^{\perp}$. This concludes the proof. $\square$

Choosing $i = j$ we get $P_i b_i$ is the dual of $Q_i d_i$. The two conditions of theorem 1 in this case translates to the following theorem.

Theorem 2. *Suppose B and D are duals of each other. Define the collection F such that $f_i = d_{n-i+1}$. Let $\tilde{B}, \tilde{F}$ be the GSOs of B and F respectively. Then $\tilde{b}_i$ and $\tilde{f}_{n-i+1}$ are duals of each other. That is*

$$\frac{\tilde{b}_i}{\|\tilde{b}_i\|} = \frac{\tilde{f}_{n-i+1}}{\|\tilde{f}_{n-i+1}\|} \quad \textit{and} \quad \|\tilde{b}_i\| \, \|\tilde{f}_{n-i+1}\| = 1.$$

Further, $\frac{\tilde{b}_n}{\|\tilde{b}_n\|^2} = f_1 = d_n$ is a vector in the dual lattice.

The last statement is derived from $d_n = f_1 = \|f_1\| \frac{\tilde{b}_n}{\|\tilde{b}_n\|} = \frac{\tilde{b}_n}{\|\tilde{b}_n\|^2}$. This states that there is a dual vector of size $\frac{1}{\|\tilde{b}_n\|}$. Equivalently, one can partition L into layers such that the distance between the layers is $\|\tilde{b}_n\|$.

2 More length relations between the primal and the dual

We have seen that short lattice vectors are of significant interest, as they reveal information about the density of the dual lattice. In general, short lattice vectors are difficult to find, and hence, various inequalities become useful. To build this theory we will introduce two important quantities: successive minima and Gram-Schmidt basis length, and establish relations between them.

Definition 5. Given a lattice L, we define λ_k as the smallest number r such that L has at least k linearly independent vectors of Euclidean norm less than r. In particular, λ_1 is the length of the shortest nonzero vector in L. To specify the lattice we may denote λ_k as $\lambda_k(L)$.

In general, we could replace the Euclidean norm in the definition by any p-norm. Then we specify the choice of norm in the superscript as $\lambda_k^{(p)}$. These numbers λ_k (or $\lambda_k^{(p)}$) are known as the **successive minima** of L.

Theorem 3. *(Minkowski's theorem) Suppose L is a full rank lattice in $\mathbb{R}^n$. Let S be any bounded, symmetric, convex set in $\mathbb{R}^n$ such that*

$$\operatorname{vol} S \geq 2^n \det L.$$

Then, $S \cap L$ has a nonzero vector.

Proof. Let $T = \frac{1}{2} S$. Note that $\operatorname{vol} S = \frac{1}{2^n} \operatorname{vol}(S) = \det L$.

Now, consider the map $T \longrightarrow \mathcal{P}(B)$ such that $t \longmapsto w$ such that $t - w \in L$. Since S is bounded, this map is a union of finitely many translations (isometries) and is therefore volume preserving. Since $\operatorname{vol} T \geq \operatorname{vol} \mathcal{P}(B) = \det L$, there exist points t_1, $t_2 \in T$ such

$$t_1 = x_1 + w \quad \text{and} \quad t_2 = x_2 + w \quad \text{and} \quad x_1 \neq x_2.$$

Then the point $x_1 - x_2 = t_1 - t_2 = \frac{s_1}{2} + \frac{(-s_2)}{2} \in S$. This is a nonzero lattice vector. □

Theorem 4. *(Hermite's bound, theorem 8.1 in [4]) Suppose L is a lattice of rank n. Then*

$$\lambda_1^{(\infty)}(L) \leq (\det L)^{\frac{1}{n}}, \quad \textit{and therefore,} \quad \lambda_1(L) \leq \sqrt{n}\,(\det L)^{\frac{1}{n}}.$$

Proposition 2. *If L is a lattice of rank n, then*

$$\frac{1}{\lambda_n(L^*)} \leq \lambda_1(L) \leq \frac{n}{\lambda_1(L^*)}\,.$$

Proof. Club theorem 4 with corollary 5 to observe

$$\lambda_1(L)\,\lambda_1(L^*) \leq \sqrt{n}\,(\det L)^{\frac{1}{n}}\sqrt{n}\,(\det L^*)^{\frac{1}{n}} \leq n.$$

This shows the upper bound. Now, suppose u is the shortest nonzero vector in L and say $v_1, \ldots, v_n$ are a set of linearly independent vectors in L^* such that $\|v_i\| \leq \lambda_n(L^*)$ for all i. Put the v_is in the rows of the matrix V. Since V is nonsingular, $Vu \neq 0$. Hence there exists an i such that $\langle u, v_i \rangle \neq 0$. But $\langle u, v_i \rangle \in \mathbb{Z}$. Therefore,

$$1 \leq |\langle u, v_i \rangle| \leq \|u\|\,\|v_i\| \leq \lambda_1(L)\,\lambda_n(L^*)\,.$$

□

This adds to the intuition we built in section 1 in our discussion about the distance between lattice layers and their effect on the geometry of the lattice. The proposition above says that the lengths of the shortest vectors of L and L^* are inversely related to each other.

Remark 1. The upper bound in proposition 2 can be significantly improved. For example, the inequality $\lambda_1(L)\,\lambda_n(L^*) \leq n$ holds. More generally, the following was proved by Banaszczyk in [1].

$$\lambda_i(L)\,\lambda_{n-i+1}(L^*) \leq n$$

Definition 6. Let $L \subset \mathcal{H}$ be any lattice. Then we call the quantity $\tilde{\mathrm{bl}}(L)$ as the **Gram-Schmidt basis length** of L and define it as

$$\tilde{\mathrm{bl}}(L) = \min_{\substack{B \\ \mathcal{L}(B)=L}} \max_i \|\tilde{b}_i\|\,.$$

In words, suppose the "size" of a basis is the length of the longest vector in its GSO. Then $\tilde{\mathrm{bl}}(L)$ is the size of the smallest basis of L.

Remark 2. The quantity $\tilde{\mathrm{bl}}(L)$ is invariant under application of unitary transformations to L. That is $\tilde{\mathrm{bl}}(L) = \tilde{\mathrm{bl}}(UL)$ for any unitary U. This follows from the fact that if $\tilde{B}$ is the GSO of B, then $U\tilde{B}$ is the GSO of UB. Thus

$$\tilde{\mathrm{bl}}(UL) = \min_{\substack{B \\ \mathcal{L}(B)=L}} \max_i \|U\tilde{b}_i\| = \min_{\substack{B \\ \mathcal{L}(B)=L}} \max_i \|\tilde{b}_i\| = \tilde{\mathrm{bl}}\,(L)\,.$$

Proposition 3. *For every lattice L there exists a unitary U such that*

$$\tilde{\mathrm{bl}}(L)\,\lambda_1^{(\infty)}(UL^*) \geq 1.$$

Proof. Let B be the basis of L such that $\tilde{\mathrm{bl}}(L) = \max_i \|\tilde{b}_i\|$. Suppose U is the unitary such that the column vectors of $U\tilde{B}$ are parallel to the canonical basis vectors. Then $\mathcal{L}(UB) = UL$ and its dual is UL^*. Let F be the dual basis of UB such that $U\tilde{b}_i$ and $\tilde{f}_{n-i+1}$ are duals of each other, for all i. That is $\tilde{f}_{n-i+1} = \frac{U\tilde{b}_i}{\|U\tilde{b}_i\|^2}$.

Let v be any vector in UL^*. Then v can be written as

$$v = \alpha_1 f_1 + \cdots + \alpha_j f_j \quad \text{where}\, \alpha_j \neq 0.$$

Note that

$$\begin{aligned} v_{n-j+1} &= \left\langle v, \frac{U\tilde{b}_{n-j+1}}{\|U\tilde{b}_{n-j+1}\|} \right\rangle = \left\langle \sum_{k=1}^{j} \alpha_k f_k, \frac{U\tilde{b}_{n-j+1}}{\|\tilde{b}_{n-j+1}\|} \right\rangle \\ &= \frac{1}{\|\tilde{b}_{n-j+1}\|} \sum_{k=j}^{n} \alpha_k \frac{\langle U\tilde{b}_{n-k+1}, U\tilde{b}_{n-j+1}\rangle}{\|\tilde{b}_{n-k+1}\|^2} \\ &= \frac{\alpha_j}{\|\tilde{b}_{n-j+1}\|}\,. \end{aligned}$$

Now, choose v such that $\lambda_1^{(\infty)}(UL^*) = \|v\|_\infty$. Then

$$\lambda_1^{(\infty)}(UL^*) \geq |v_{n-j+1}| = \frac{|\alpha_j|}{\|\tilde{b}_{n-j+1}\|} \geq \frac{1}{\|\tilde{b}_{n-j+1}\|} \geq \frac{1}{\tilde{\mathrm{bl}}(L)}. \qquad \square$$

3 Anti-concentration bounds for multivariate Gaussians

The Gaussian distribution on $\mathbb{R}^n$ has parameters μ, Σ and the density function

$$N(x) = N_{\mu,\Sigma}(x) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} \exp\Big[-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\Big].$$

Here $\mu \in \mathbb{R}^n$ is called the mean of the distribution, and the positive semidefinite matrix Σ is called the covariance matrix. A random variable X with this density function is denoted by $X \sim N(\mu, \Sigma)$. If $X = [\, X_1 \ \cdots \ X_n \,]^T$, then we have

$$\mathbb{E} X_i = \mu_i \ \text{ and } \operatorname{cov}(X_i, X_j) = \Sigma_{ij}.$$

Now, for $s > 0$, define Γ_s as follows

$$\Gamma_s = N\Big(0, \frac{s^2}{2\pi} I\Big). \tag{2}$$

If $Y \sim \Gamma_s$, then $\mathbb{E}\, \|Y\|^2 = \frac{ns^2}{2\pi}$. Our attention will be on a slightly modified map ρ_s defined for $s > 0$ as

$$\rho_s(x) = s^n \Gamma_s(x) = \exp\big(-\pi \, \|\tfrac{x}{s}\|^2\big). \tag{3}$$

When s is omitted we take $s = 1$. We naturally extend the definition of ρ_s to subsets of $\mathbb{R}^n$ as

$$\rho_s(A) := \sum_{x \in A} \exp\big(-\pi \, \|\tfrac{x}{s}\|^2\big) \quad \text{for any } A \subset \mathbb{R}^n. \tag{4}$$

Going forward we look at a selection of Banaszczyk's anticoncentration bounds for ρ evaluated on the points of a lattice. An anticoncentration bound, in this context, gives the probability that an arbitrary lattice point is "far" from a given point. In other words, this is the probability that the lattice points are not concentrated around a given point.

Definition 7. Let L be a lattice in $\mathbb{R}^n$ and r some positive number. We define $L_r^{(p)}$ and $\bar{L}_r^{(p)}$ as the sets

$$L_r^{(p)} = \{x \in L \colon \|x\|_p > r\} \ \text{ and } \ \bar{L}_r^{(p)} = \{x \in L \colon \|x\|_p \leq r\}.$$

We have $L = L_r^{(p)} \cup \bar{L}_r^{(p)}$. And $L_0^{(p)} = L - \{0\}$. When $p = 2$, we write $L_r^{(p)}$ as just L_r.

Lemma 1. *(Lemma 1.5, [1]) For any lattice $L\subset\mathbb{R}^n$ we have*

$$\frac{\rho(L_{\sqrt{n}})}{\rho(L)}<2^{-n}.$$

The lemma states that the contribution to $\rho(L)$ from lattice vectors of size larger than $\sqrt{n}$ is negligible. In a similar vein we have the following.

Lemma 2. *(Lemma 2.10, [2]) Let $L\subset\mathbb{R}^n$ be a lattice and $r>0$. Then*

$$\frac{\rho(L_r^{(\infty)})}{\rho(L)}<2n\,e^{-\pi r^2}.$$

4 Fourier transform and the dual lattice

Definition 8. Suppose $V\subset\mathbb{R}^n$, and $f\colon V\to\mathbb{R}$ is differentiable.[4] Then the function $\hat{f}$, called the **Fourier transform** of f on V, is defined as

$$\hat{f}(y)=\int_V f(x)\,e^{2\pi i\langle x,y\rangle}\,\mathrm{d}x.$$

The Fourier transform behaves neatly with lattice structures. As such, suppose $V=\mathcal{P}(B)$, and $L=\mathcal{L}(B)$. Then, for any $x\in\mathcal{P}(B)$, we have

$$f(x)=\sum_{y\in L^*}\hat{f}(y)\,e^{-2\pi i\langle x,y\rangle}.$$

Another useful property of Fourier transforms on lattices is that a function summed over a lattice connects to its Fourier transform summed over the dual lattice.

Lemma 3. *Suppose $f\colon\mathbb{R}^n\to\mathbb{R}$ is differentiable. Then for any lattice $L\subset\mathbb{R}^n$ and any point $x\in\mathbb{R}^n$ we have*

$$\sum_{z\in L}f(x+z)=\det L^*\sum_{y\in L^*}\hat{f}(y)e^{2\pi i\langle x,y\rangle}.$$

In particular, when $x=0$, we get $f(L)=\det L^*\,\hat{f}(L^*)$.

Proposition 4. *With ρ as defined in equation (4), we have $\hat{\rho}_s=s^n\rho_{\frac{1}{s}}$ on $\mathbb{R}^n$.*

4. Fourier transforms exist for more general class of functions. But for our purposes the seemingly strong assumption of differentiability is sufficient.

5 Bounds on the smoothing parameter

Recall the distribution Γ_s on $\mathbb{R}^n$ from equation (2), which is a Gaussian random variable with mean 0 and covariance matrix $\frac{s^2}{2\pi} I$.

Let $L = \mathcal{L}(B) \subset \mathbb{R}^n$ be a lattice. We define the distribution γ_s on $\mathcal{P}(B)$ as $\Gamma_s \bmod \mathcal{P}(B)$, which in turn is defined as follows

$$\gamma_s(x) = \sum_{y \in L} \Gamma_s(x+y)\,. \tag{5}$$

Intuitively, sampling from this distribution involves the following steps. Sample $z \in \mathbb{R}^n$ with the distribution Γ_s. Then find the unique vector $x \in \mathcal{P}(B)$ such that $y = z - x \in L$. Then x is a sample from $\mathcal{P}(B)$ with the distribution γ_s.

With this we are equipped to define the smoothing parameter of L.

Definition 9. The **smoothing parameter** $\eta_\varepsilon = \eta_\varepsilon(L)$ of the n-dimensional lattice L is the smallest positive number s such that

$$\|\gamma_s - \sqcup\|_1 < \varepsilon\,.$$

where $\sqcup$ is the uniform distribution over $\mathcal{P}(B)$. This means, for $s > \eta_\varepsilon$, if one samples points from Γ_s, then the corresponding distribution γ_s is "almost" indistinguishable from the uniform distribution on $\mathcal{P}(B)$.

Lemma 4. *Let $L = \mathcal{L}(B)$ be any lattice, and let γ_s be the distributioin $\Gamma_s \bmod \mathcal{P}(B)$. Let $\sqcup$ be the uniform distribution on $\mathcal{P}(B)$. Then*

$$\|\gamma_s - \sqcup\|_1 \leq \rho_{\frac{1}{s}}(L_0^*).$$

Proof. First note that for any $x \in \mathcal{P}(B)$ we have

$$\begin{aligned}
\gamma_s(x) - \sqcup(x) &= \sum_{y \in L} \frac{1}{s^n} \rho_s(x+y) - \frac{1}{\det L} && \text{[equation (3), (5)]} \\
&= \frac{1}{\det L} \sum_{z \in L^*} \frac{1}{s^n} \hat{\rho}_s(z)\, e^{2\pi i \langle x, z\rangle} - \frac{1}{\det L} && \text{[lemma 3]} \\
&= \frac{1}{\det L} \sum_{z \in L^*} \rho_{\frac{1}{s}}(z)\, e^{2\pi i \langle x, z\rangle} - \frac{1}{\det L} && \text{[proposition 4]} \\
&= \frac{1}{\det L} \sum_{z \in L_0^*} \rho_{\frac{1}{s}}(z)\, e^{2\pi i \langle x, z\rangle}\,.
\end{aligned}$$

Therefore, we calculate (assuming the sum and integral can be interchanged)

$$\begin{aligned}
\|\Upsilon_s - \sqcup\|_1 &= \int_{\mathcal{P}(B)} |\Upsilon_s(x) - \sqcup(x)|\, \mathrm{d}x \\
&\leq \int_{\mathcal{P}(B)} \frac{1}{\det L} \sum_{z \in L_0^*} \left| \rho_{\frac{1}{s}}(z)\, e^{2\pi i \langle x, y\rangle} \right| \mathrm{d}x \\
&= \sum_{z \in L_0^*} \left| \rho_{\frac{1}{s}}(z) \right| \left[\frac{1}{\det L} \int_{\mathcal{P}(B)} \mathrm{d}x \right] \\
&= \sum_{z \in L_0^*} \rho_{\frac{1}{s}}(z) \\
&= \rho_{\frac{1}{s}}(L_0^*)\,.
\end{aligned}$$

□

Definition 10. Owing to lemma 4, the smoothing parameter $\eta_\varepsilon(L)$ of L is alternatively defined as the smallest s for which $\rho_{\frac{1}{s}}(L_0^*) < \varepsilon$. Further note that this definition removes the apparent dependence of the parameter on B, the lattice basis.

Remark 3. Note that for any $x \in \mathbb{R}^n$ and any unitary U we have

$$\rho_{\frac{1}{s}}(Ux) = \exp(-\pi s\, \|Ux\|^2) = \exp(-\pi s\, \|x\|^2) = \rho_{\frac{1}{s}}(x).$$

Therefore, the smoothing parameter of a lattice is invariant under unitary transformations of the lattice. That is, $\eta_\varepsilon(L) = \eta_\varepsilon(UL)$ for every unitary U.

Below we give two kinds of upper bounds for the smoothing parameter, relating it to short vectors of the dual lattice.

Proposition 5. *Let $\varepsilon = 2^{-n}$. For any lattice $L \subset \mathbb{R}^n$, we have $\eta_\varepsilon(L) \leq \frac{\sqrt{n}}{\lambda_1(L^*)}$.*

Proof. For any lattice Λ we note that $\rho(\Lambda) = \rho\big(\Lambda_{\sqrt{n}}\big) + \rho\big(\bar{\Lambda}_{\sqrt{n}}\big)$. Lemma 1 says

$$\rho\big(\Lambda_{\sqrt{n}}\big) < 2^{-n} \rho(\Lambda).$$

Eliminating $\rho(\Lambda)$ using the two expressions we get

$$\rho\big(\Lambda_{\sqrt{n}}\big) < \frac{2^{-n}}{1 - 2^{-n}} \rho\big(\bar{\Lambda}_{\sqrt{n}}\big) = \frac{1}{2^n - 1} \rho\big(\bar{\Lambda}_{\sqrt{n}}\big) < 2^{-n} \rho\big(\bar{\Lambda}_{\sqrt{n}}\big)\,.$$

Suppose $s > 0$ be such that $s > \frac{\sqrt{n}}{\lambda_1(L^*)}$. Then no nonzero vector in the lattice sL^* has size smaller than $\sqrt{n}$. Therefore, $(sL^*)_0 = (sL^*)_{\sqrt{n}}$. And

$$\rho_{\frac{1}{s}}(L_0^*) = \rho((sL^*)_0) = \rho\big((sL^*)_{\sqrt{n}}\big) < 2^{-n}\rho\big(\overline{(sL^*)}_{\sqrt{n}}\big) = 2^{-n}\rho(0) = 2^{-n},$$

where the third inequality is obtained by putting $\Lambda = sL^*$. $\square$

The second bound we derive is with respect to the infinity norm and for arbitrary ε. For presentation we define

$$\beta_{n,\varepsilon} = \sqrt{\frac{\ln(2n[1 + 1/\varepsilon])}{\pi}}\,.$$

Clearly $\beta_{n,\varepsilon}$ is increasing in n and decresing in ε.

Theorem 5. *Let $L \subset \mathbb{R}^n$ be a lattice and $\varepsilon > 0$. Then*

$$\eta_\varepsilon(L) \leq \frac{\beta_{n,\varepsilon}}{\lambda_1^{(\infty)}(L_*)}.$$

Proof. Let $\lambda < \lambda_1^{(\infty)}(L^*)$. No nonzero vector in sL^* has infinity norm smaller than $s\lambda$. Thus

$$\rho_{\frac{1}{s}}(L_0^*) = \rho(sL^* - \{0\}) = \rho((sL^*)_0^\infty) = \rho\big((sL^*)_{s\lambda}^{(\infty)}\big).$$

Now pick $s = \frac{\beta_{n,\varepsilon}}{\lambda} > \frac{\beta_{n,\varepsilon}}{\lambda_1^{(\infty)}(L_*)}$. Then

$$\begin{aligned} \rho_{\frac{1}{s}}(L_0^*) &= \rho\big((sL^*)_{s\lambda}^{(\infty)}\big) \leq 2ne^{-\pi s^2\lambda^2} && \text{[lemma 2]} \\ &= 2ne^{-\ln\left[2n\left(1+\frac{1}{\varepsilon}\right)\right]}\rho(sL^*) \\ &= \frac{\varepsilon}{1+\varepsilon}[1 + \rho(sL^* - \{0\})] \\ &= \frac{\varepsilon}{1+\varepsilon}\Big(1 + \rho_{\frac{1}{s}}(L_0^*)\Big). \end{aligned}$$

Rearranging gives $\rho_{\frac{1}{s}}(L_0^*) < \varepsilon$. Therefore, $\eta_\varepsilon(L) \leq s$. $\square$

This finally leads to the smoothing parameter bound appearing in seminal work [3] of Gentry, Peikert and Vaikunthanathan.

Theorem 6. *Let $L \subset \mathbb{R}^n$ be any lattice. Then $\eta_\varepsilon(L) \leq \tilde{\mathrm{bl}}(L)\,\beta_{n,\varepsilon}$.*

Proof. By proposition 3 we know that there exists an unitary U such that

$$\frac{1}{\lambda_1^{(\infty)}(UL^*)} \leq \tilde{\mathrm{bl}}(L).$$

Apply theorem 5 to the lattice UL to get

$$\eta_\varepsilon(UL) \leq \frac{\beta_{n,\varepsilon}}{\lambda_1^{(\infty)}(UL^*)}.$$

But, by remark 3, we have $\eta_\varepsilon(UL) = \eta_\varepsilon(L)$. Combining these equations we obtain the required bound. □

Bibliography

[1] Wojciech Banaszczyk. New bounds in some transference theorems in the geometry of numbers. *Mathematische Annalen*, 296(4):625-635, 1993.

[2] Wojciech Banaszczyk. Inequalities for convex bodies and polar reciprocal lattices in $\mathbb{R}^n$. *Discrete & Computational Geometry*, 13(2):217-232, 1995.

[3] Craig Gentry, Chris Peikert, and Vinod Vaikuntanathan. Trapdoors for hard lattices and new cryptographic constructions. In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing (STOC)*, pages 197-206. ACM, 2008.

[4] Jeffrey Hoffstein, Jill Pipher, and Joseph H. Silverman. *An Introduction to Mathematical Cryptography*. Springer, New York, 2 edition, 2014.