CSE 206A: Lattice Algorithms and Applications Fall 2017

Minkowski's theorem

Instructor: *Daniele Micciancio* UCSD CSE

There are many important quantities associated to a lattice. Some of them, like the determinant, can be efficiently computed given any basis for the lattice. In this lecture we will study other quantities for which no efficient algorithm is known. While no efficient algorithm is known to efficiently compute these quantities, we will establish easily computable upper and lower bounds, and describe a simple mathematical application based on these efficiently computable bounds.

1 Minimum Distance

Definition 1 *For any lattice Λ, the minimum distance of Λ is the smallest distance between any two lattice points:*

$$\lambda(\Lambda) = \inf\{\|\mathbf{x}-\mathbf{y}\| \colon \mathbf{x},\mathbf{y} \in \Lambda, \mathbf{x} \neq \mathbf{y}\}.$$

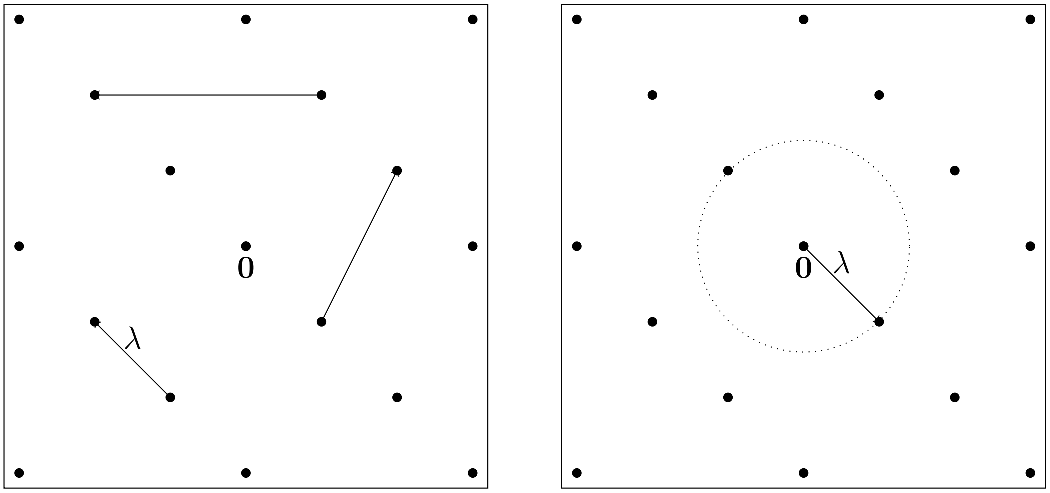

Figure 1: The minimum distance of a lattice λ is the smallest distance between any two lattice points, and equals the length of the shortest nonzero vector

We observe that the minimum distance can be equivalently defined as the length of the shortest nonzero lattice vector:

$$\lambda(\Lambda) = \inf\{\|\mathbf{v}\| \colon \mathbf{v} \in \Lambda \setminus \{\mathbf{0}\}\}.$$

This follows from the fact that lattices are additive subgroups of $\mathbb{R}^n$, i.e., they are closed under addition and subtraction. So, if $\mathbf{x}$ and $\mathbf{y}$ are distinct lattice points, then $\mathbf{x}-\mathbf{y}$ is a

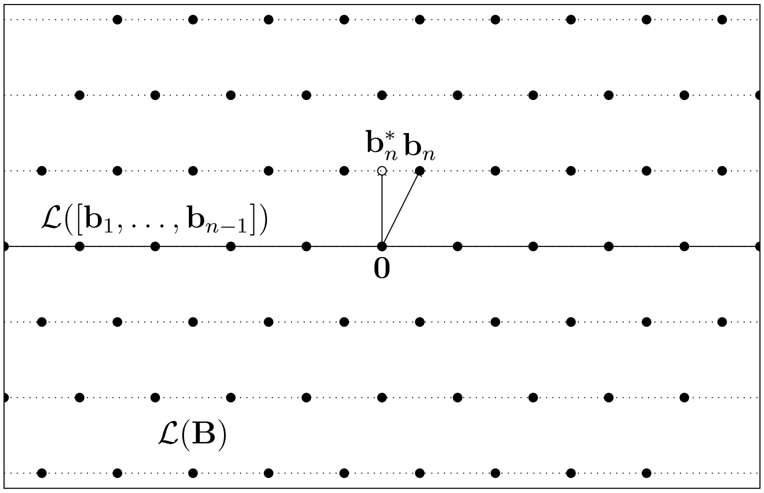
$\mathbf{b}_n^*$
$\mathbf{b}_n$
$\mathcal{L}([\mathbf{b}_1,\ldots,\mathbf{b}_{n-1}])$
$\mathbf{0}$
$\mathcal{L}(\mathbf{B})$

Figure 2: The lattice $\mathcal{L}(\mathbf{B})$ can be decomposed into layers orthogonal to $\mathbf{b}_n^*$, at distance $\|\mathbf{b}_n^*\|$ from each other. In particular, all points in layers other than the one through the origin, have length at least $\|\mathbf{b}_n^*\|$.

nonzero lattice point. The first thing we want to prove about the minimum distance is that it is always achieved by some lattice vector, i.e., there is a lattice vector $\mathbf{x} \in \Lambda$ of length exactly $\|\mathbf{x}\| = \lambda(\Lambda)$. To prove this, we need first to establish a lower bound on $\lambda(\Lambda)$.

Consider a lattice basis $\mathbf{B}$, and its Gram-Schmidt orthogonalization $\mathbf{B}^*$. The lattice $\Lambda = \mathcal{L}(\mathbf{B})$ can be partitioned into layers $\mathcal{L}([\mathbf{b}_1,\ldots,\mathbf{b}_{n-1}]) + c \cdot \mathbf{b}_n$ where $c \in \mathbb{Z}$. (See Figure 2.) Clearly, all points on layers other $c = 0$ are at distance at least $\|\mathbf{b}_n^*\|$ from the origin. The remaining lattice points belong to a lower dimensional lattice $\mathcal{L}([\mathbf{b}_1,\ldots,\mathbf{b}_{n-1}])$. It follows by induction that all nonzero lattice points in $\mathcal{L}(\mathbf{B})$ have length at least $\min_i \|\mathbf{b}_i^*\|$.

An alternative way to see that $\min_i \|\mathbf{b}_i^*\|$ is a lower bound on the minimum distance of a lattice is to consider the fundamental region $C = \mathcal{C}(\mathbf{B}^*)$ defined by the (centered) orthogonalized parallelepiped. The open ball $\mathcal{B}(r) = \{\mathbf{x}\colon \|\mathbf{x}\| < r\}$ of radius $r = \frac{1}{2}\min_i \|\mathbf{b}_i^*\|$ is completely contained in C. Since C is a fundamental region of $\mathcal{L}(\mathbf{B})$, the shifted regions $\mathbf{v} + C$ (with $\mathbf{v} \in \mathcal{L}(\mathbf{B})$) are disjoint. So, also the balls $\mathbf{v} + \mathcal{B}(r) \subset \mathbf{v} + C$ (for $\mathbf{v} \in \mathcal{L}(\mathbf{B})$) are disjoint, and any two lattice points $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{L}(\mathbf{B})$ must be at distance at least $2r = \min_i \|\mathbf{b}_i^*\|$. The following theorem gives still another proof that $\min_i \|\mathbf{b}_i^*\|$ is a lower bound on the minimum distance of a lattice, using simple linear algebra.

Theorem 2 *For every lattice basis* $\mathbf{B}$ *and its Gram-Schmidt orthogonalization* $\mathbf{B}^*$, $\lambda(\mathcal{L}(\mathbf{B})) \geq \min_i \|\mathbf{b}_i^*\|$.

Proof. Let us consider a generic lattice vector $\mathbf{Bx}$, where $\mathbf{x} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}$ and let k be the biggest index such that $x_k \neq 0$. We prove that

$$\|\mathbf{Bx}\| \geq \|\mathbf{b}_k^*\| \geq \min_i \|\mathbf{b}_i^*\|. \tag{1}$$

In order to prove (1), we take the scalar product of our lattice vector and $\mathbf{b}_k^*$. Using the orthogonality of $\mathbf{b}_k^*$ and $\mathbf{b}_i$ (for $i < k$) we get

$$\langle \mathbf{B}\mathbf{x}, \mathbf{b}_k^* \rangle = \sum_{i \leq k} \langle \mathbf{b}_i x_i, \mathbf{b}_k^* \rangle = x_k \langle \mathbf{b}_k, \mathbf{b}_k^* \rangle = x_k \|\mathbf{b}_k^*\|^2.$$

By Cauchy-Shwartz, $\|\mathbf{B}\mathbf{x}\| \cdot \|\mathbf{b}_k^*\| \geq |\langle \mathbf{B}\mathbf{x}, \mathbf{b}_k^* \rangle| \geq |x_k| \cdot \|\mathbf{b}_k^*\|^2$. Using $|x_k| \geq 1$ and dividing by $\|\mathbf{b}_k\|^*$, we get $\|\mathbf{B}\mathbf{x}\| \geq \|\mathbf{b}_k^*\|$. □

An immediate consequence of Theorem 2 is that the minimum distance of a lattice $\lambda(\Lambda) > 0$ is strictly positive, and the lattice Λ is a *discrete* subgroup of $\mathbb{R}^n$, i.e., it satisfies the following properties:

(subgroup) Λ is closed under addition and subtraction,[1]

(discrete) there is an $\epsilon > 0$ such that any two distinct lattice points $\mathbf{x} \neq \mathbf{y} \in \Lambda$ are at distance at least $\|\mathbf{x} - \mathbf{y}\| \geq \epsilon$.

Notice that not every subgroup of $\mathbb{R}^n$ is a lattice. For example, $\mathbb{Q}^n$ is a subgroup of $\mathbb{R}^n$, but it is not a lattice because it is not discrete. However, it can be shown that every discrete subgroup of $\mathbb{R}^n$ is a lattice, i.e., it can be expressed as $\mathcal{L}(\mathbf{B})$ for some basis $\mathbf{B}$.

The definition $\mathcal{L}(\mathbf{B}) = \{\mathbf{B}\mathbf{x} : \mathbf{x} \in \mathbb{Z}^n\} \subset \mathbb{R}^d$ can be extended to matrices $\mathbf{B}$ whose columns are not linearly independent, and the resulting set is always a subgroup of $\mathbb{R}^d$. However, it is not always a lattice because it may not be discrete. Still, we will see that if $\mathbf{B}$ is a matrix with integer or rational entries, then $\mathcal{L}(\mathbf{B})$ is always a lattice.

Exercise 1 *Find a matrix $\mathbf{B} \in \mathbb{R}^{d \times n}$ such that $\mathcal{L}(\mathbf{B})$ is not a lattice.* [Hint: $\mathbf{B}$ can be as small as a 1-by-2 matrix.]

Notice that the lower bound $\min_i \|\mathbf{b}_i^*\|$ depends on the choice of the basis. We will see later in the course that some bases give better lower bounds than others, but at this point any nonzero lower bound will suffice. We want to show that there is a lattice vector of length λ. Consider a sphere of radius $\frac{3}{2}\lambda > \lambda$. Clearly, in the definition of $\lambda = \inf\{\|\mathbf{x}\| : \mathbf{x} \in \Lambda \setminus \{0\}\}$, we can restrict $\mathbf{x}$ to range over all lattice vectors inside the sphere of radius $\frac{3}{2}\lambda$. We observe that (by a volume argument, see Figure 3) the sphere contains only finitely many lattice points. It follows that we can replace the infimum over the whole lattice with a minimum over a finite subset, and there is a point in the set achieving the smallest possible norm.

This shows that any lattice Λ has a vector of length $\lambda(\Lambda)$. Finding such a vector is a central problem in the algorithmic study of lattices.

Definition 3 *The Shortest Vector Problem (*SVP*), given a lattice basis $\mathbf{B}$, asks to find a shortest nonzero lattice vector, i.e., a vector $\mathbf{v} \in \mathcal{L}(\mathbf{B})$ with $\|\mathbf{v}\| = \lambda(\mathcal{L}(\mathbf{B}))$.*

[1] Technically, closure under subtraction is enough because addition can be expressed as $a + b = a - (-b)$.

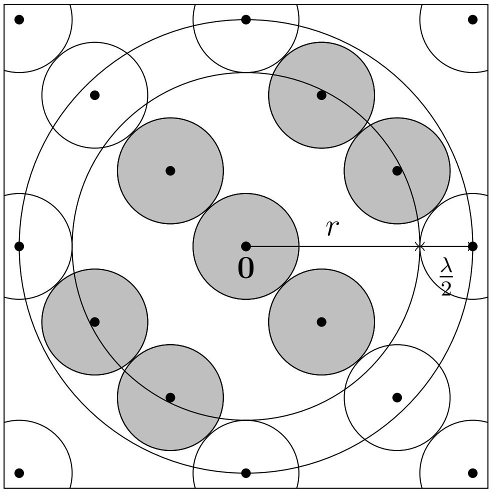

Figure 3: Consider a lattice with minimum distance λ, and a ball S of radius r. Open balls of radius $\lambda/2$ centered around lattice points are disjoint, and they are all contained in ball S' of slightly larger radius $r + \lambda/2$. So, since the volume of the small balls cannot exceed the volume of S', the number of lattice points in S is at most $\frac{(r+\lambda/2)^n}{(\lambda/2)^n} = (1 + 2r/\lambda)^n$. The same argument applies even when S is not centered around the origin.

The SVP can be defined with respect to any norm, but the euclidean norm $\|\mathbf{x}\| = \sqrt{\sum_i x_i^2}$ is the most common, and the one we assume unless otherwise stated. Other important norms in the study of lattice problems are the ℓ_1 norm $\|\mathbf{x}\|_1 = \sum_i |x_i|$, and the ℓ_∞ norm $\|\mathbf{x}\|_\infty = \max_i |x_i|$. As no efficient algorithm is known to solve SVP exactly, it is interesting to study the problem of finding approximately shortest vectors in a lattice.

Definition 4 *For any $\gamma \geq 1$, the γ-approximate* SVP*, given a basis $\mathbf{B}$, asks to find a nonzero lattice vector $\mathbf{v} \in \mathcal{L}(\mathbf{B}) \setminus \{\mathbf{0}\}$ of norm at most $\|\mathbf{v}\| \leq \gamma \cdot \lambda(\mathcal{L}(\mathbf{B}))$.*

Notice that exact SVP is equivalent to γ-approximate SVP for $\gamma = 1$. As SVP gets harder as the lattice dimension increases, the approximation factor γ is usually a (monotonically increasing) function of the dimension of the lattice.

One can also define a decision (promise) problem naturally associated to SVP, which informally corresponds to the task of determining (or approximating) the value of λ, without necessarily finding a lattice vector achieving that length.

Definition 5 GapSVP$_\gamma$*, given a basis $\mathbf{B}$ and a positive real*[2] *d, asks to determine if $\lambda(\mathcal{L}(\mathbf{B})) \leq d$ or $\lambda(\mathcal{L}(\mathbf{B})) > \gamma d$.*

In the above problem definition, when $d < \lambda(\mathcal{L}(\mathbf{B})) \leq \gamma d$, any answer is acceptable. In other words, an algorithm attempting to solve GapSVP is only required to work when the

[2] When $\mathbf{B}$ is an integer basis, one can assume without loss of generality that d^2 is an integer, and therefore it admits a short simple description.

minimum distance of the lattice is either small, or much larger. This gap between positive and negative instances of the problem capture the task of computing λ only approximately, within a factor γ.

Exercise 2 *Prove that for any $\gamma \geq 1$,* $\textsc{GapSVP}_\gamma$ *can be efficiently reduced to the problem of approximating $\lambda(\mathcal{L}(\mathbf{B}))$ within a factor γ. (We say that an algorithm approximates λ within a factor γ, if, on any input lattice, it outputs a real in the range $[\lambda, \gamma\lambda)$.)*

Exercise 3 *Prove that the problem of approximating $\lambda(\mathcal{L}(\mathbf{B}))$ within a factor γ can be efficiently reduced to* $\textsc{GapSVP}_\gamma$*.* [Hint: use binary search.]

Exercise 4 *Prove that* $\textsc{GapSVP}_\gamma$ *efficiently reduces to* SVP_γ*.*

Exercise 5 *Prove that* SVP_1 *efficiently reduces to* $\textsc{GapSVP}_1$*.*

Notice that the last exercise asks to reduce the search version to the decision version of SVP only for their exact version $\gamma = 1$. Finding a reduction from SVP_γ to $\textsc{GapSVP}_\gamma$ for arbitrary γ is an important open problem in the complexity of lattice problems, with interesting cryptographic implications. (More about this later on in the course.)

2 Minkowski's theorem

We now turn to estimating the value of λ from above. Clearly, for any basis $\mathbf{B}$, we have $\lambda(\mathbf{B}) \leq \min_i \|\mathbf{b}_i\|$, because each column of $\mathbf{B}$ is a nonzero lattice vector. We would like to get a better bound, and, specifically, a bound that does not depend on the choice of the basis. Clearly, lattices with arbitrarily large minimum distance can be easily obtained simply by scaling any given lattice by a constant $c > 0$ to obtain $\lambda(c \cdot \Lambda) = c \cdot \lambda(\Lambda)$. What if we normalize the lattice so that $\det(\Lambda) = 1$? By definition of determinant, these are lattices with density 1, i.e., with about one lattice point per each unit volume of space. Can the lattice still have arbitrarily large minimum distance? Equivalently, we are asking if it is possible to bound the ratio $\lambda(\Lambda)/\det(\Lambda)^{1/n}$ for any n-dimensional lattice Λ. (Notice that the quantity $\lambda(\Lambda)/\det(\Lambda)^{1/n}$ is invariant under linear scaling because $\det(c \cdot \Lambda) = c^n \cdot \det(\Lambda)$.) For historical reasons[3], mathematicians have defined and studied the square of this quantity, which is called the *Hermite factor* of a lattice.

Definition 6 *The* Hermite factor *of an n-dimensional lattice Λ is the quantity $\gamma(\Lambda) = (\lambda(\Lambda)/\det(\Lambda)^{1/n})^2$. The* Hermite constant *in dimension n is the supremum $\gamma_n = \sup_\Lambda \gamma(\Lambda)$, where Λ ranges over all n-dimensional lattices.*

The upper bound on γ_n we are going to prove is due to Minkowski, but in proving it we follow a different path, by first proving a theorem of Blichfeldt, and then deriving Minkowski's theorem as a corollary.

[3] These problems were originally formulated and studied in the equivalent language of positive definite quadratic forms.

Theorem 7 *Given a lattice* $\Lambda = \mathcal{L}(\mathbf{B})$ *and a set* $S \subseteq span(\mathbf{B})$, *if* $vol(S) > \det(\Lambda)$ *then* S *contains two points* $\mathbf{z}_1, \mathbf{z}_2 \in S$ *such that* $\mathbf{z}_1 - \mathbf{z}_2 \in \Lambda$.

Proof. Consider the sets $S_\mathbf{x} = S \cap (\mathbf{x} + \mathcal{P}(\mathbf{B}))$, where $\mathbf{x} \in \Lambda$. Notice that these sets form a partition of S (see Figure 4,) i.e., they are pairwise disjoint and

$$S = \bigcup_{\mathbf{x} \in \Lambda} S_\mathbf{x}.$$

In particular we have

$$\text{vol}(S) = \sum_{\mathbf{x} \in \Lambda} \text{vol}(S_\mathbf{x}).$$

Notice that the shifted sets $S_\mathbf{x} - \mathbf{x} = (S - \mathbf{x}) \cap \mathcal{P}(\mathbf{B})$ are all contained in $\mathcal{P}(\mathbf{B})$. We want to prove that the $S_\mathbf{x}$ cannot be all mutually disjoint. Since $\text{vol}(S_\mathbf{x}) = \text{vol}(S_\mathbf{x} - \mathbf{x})$, we have

$$\text{vol}(\mathcal{P}(\mathbf{B})) = \det(\Lambda) < \text{vol}(S) = \sum_{\mathbf{x} \in \Lambda} \text{vol}(S_\mathbf{x}) = \sum_{\mathbf{x} \in \Lambda} \text{vol}(S_\mathbf{x} - \mathbf{x}).$$

The facts that $S_\mathbf{x} - \mathbf{x} \subseteq \mathcal{P}(\mathbf{B})$ and $\sum_{\mathbf{x} \in \Lambda} \text{vol}(S_\mathbf{x} - \mathbf{x}) > \text{vol}(\mathcal{P}(\mathbf{B}))$ imply that these sets cannot be disjoint, i.e. there exist two distinct vectors $\mathbf{x} \neq \mathbf{y} \in \Lambda$ such that $(S_\mathbf{x} - \mathbf{x}) \cap (S_\mathbf{y} - \mathbf{y}) \neq 0$.

Let $\mathbf{z}$ be any vector in the (non-empty) intersection $(S_\mathbf{x} - \mathbf{x}) \cap (S_\mathbf{y} - \mathbf{y})$ and define

$$\mathbf{z}_1 = \mathbf{z} + \mathbf{x} \in S_\mathbf{x} \subseteq S$$

$$\mathbf{z}_2 = \mathbf{z} + \mathbf{y} \in S_\mathbf{y} \subseteq S.$$

These two vectors satisfy

$$\mathbf{z}_1 - \mathbf{z}_2 = \mathbf{x} - \mathbf{y} \in \Lambda.$$

□

As a corollary to Blichfeldt theorem we easily get the following theorem of Minkowski that will be used to bound the length of the shortest vector in a lattice.

Corollary 8 *[Minkowski's convex body theorem] Let* Λ *be a full dimensional lattice. If* $S \subset \mathbb{R}^n$ *is a symmetric convex body of volume* $vol(S) > 2^n \det(\Lambda)$, *then* S *contains a nonzero lattice point.*

Proof. Consider the set $S/2 = \{\mathbf{x} : 2\mathbf{x} \in S\}$. The volume of $S/2$ satisfies

$$\text{vol}(S/2) = 2^{-n}\text{vol}(S) > \det(\Lambda)$$

By Blichfeldt theorem there exist $\mathbf{z}_1, \mathbf{z}_2 \in S/2$ such that $\mathbf{z}_1 - \mathbf{z}_2 \in \Lambda \setminus \{\mathbf{0}\}$. By definition of $S/2$, $2\mathbf{z}_1, 2\mathbf{z}_2 \in S$. Since S is symmetric, we also have $-2\mathbf{z}_2 \in S$ and by convexity,

$$\mathbf{z}_1 - \mathbf{z}_2 = \frac{2\mathbf{z}_1 - 2\mathbf{z}_2}{2} \in S$$

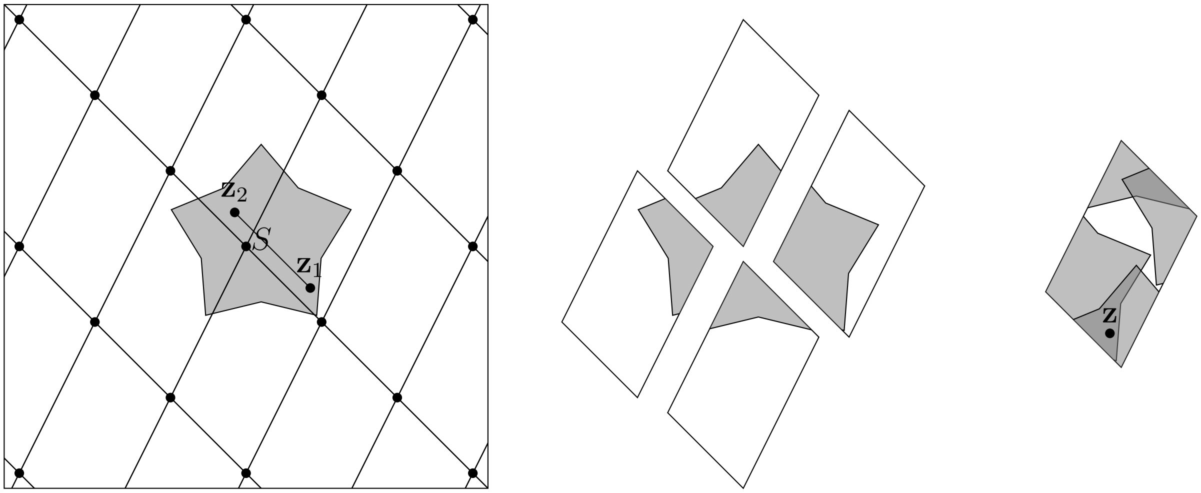

Figure 4: Illustration of Blichfeldt's theorem.

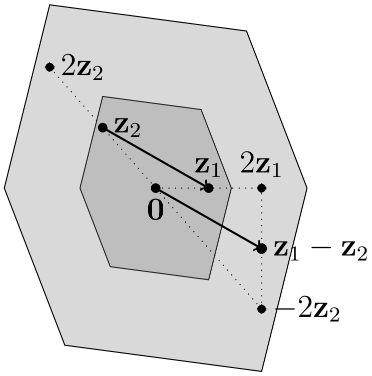

Figure 5: Minkowski's theorem

is a non-zero lattice vector contained in the set S. (See Figure 5.) □

The relation between Minkowski theorem and bounding the length of the shortest vector in a lattice is easily explained. Consider first the ℓ_∞ norm: $\|\mathbf{x}\|_\infty = \max_i |x_i|$. We show that every (full rank, n-dimensional) lattice Λ always contains a nonzero vector $\mathbf{x}$ such that $\|\mathbf{x}\|_\infty \leq \det(\Lambda)^{1/n}$. Let $l = \min\{\|\mathbf{x}\|_\infty : \mathbf{x} \in \Lambda \setminus \{\mathbf{0}\}\}$ and assume for contradiction $l > \det(\Lambda)^{1/n}$. Take the hypercube $C = \{\mathbf{x} : \|\mathbf{x}\| < l\}$. Notice that C is convex, symmetric, and has volume $\mathrm{vol}(C) = (2l)^n > 2^n \det(\Lambda)$. So, by Minkowski's theorem, C contains a nonzero lattice vector $\mathbf{x}$. By definition of C, we have $\|\mathbf{x}\|_\infty < l$, a contradiction to the minimality of l. This gives the following corollary.

Corollary 9 *For any full dimensional lattice $\Lambda \subset \mathbb{R}^n$ there exists a lattice point $\mathbf{x} \in \Lambda/\{0\}$ such that*

$$\|\mathbf{x}\|_\infty \leq \det(\Lambda)^{1/n}.$$

Using the inequality $\|\mathbf{x}\| \leq \sqrt{n}\|\mathbf{x}\|_\infty$ (valid for any n-dimensional vector $\mathbf{x}$), we get a corresponding bound in the ℓ_2 norm. It is easy to see that for Euclidean norm the full dimensionality condition is not necessary because one can embed any lattice $\Lambda \subset \mathbb{R}^d$ of rank n into $\mathbb{R}^n$ by a simple orthogonal projection operation.

Corollary 10 *Hermite constant is at most $\gamma_n \leq n$, i.e., for any n-dimensional lattice Λ there exists a lattice point $\mathbf{x} \in \Lambda \setminus \{\mathbf{0}\}$ such that*

$$\|\mathbf{x}\|_2 \leq \sqrt{n} \det(\Lambda)^{1/n}.$$

We could have proved the bound for the Euclidean norm directly, using a sphere instead of a cube, and then plugging in the formula for the volume of an n-dimensional sphere. This can be useful to get slightly better bounds, but only by a constant (independent of n) factor. For example, in two dimensions, for any lattice Λ, the disk $S = \{\mathbf{x} \colon \|\mathbf{x}\| < \lambda(\Lambda)\}$ contains no nonzero lattice point. So, by Minkowski's theorem, the area of S can be at most $2^n \det(\Lambda) = 4\det(\Lambda)$. But we know that the area of S is $\pi\lambda^2$. So, $\lambda(\Lambda) \leq 2\sqrt{\det(\Lambda)/\pi}$, which is strictly smaller than $\sqrt{2}\det(\Lambda)^{1/n}$. This yields the bound $\gamma_2 \leq 4/\pi \approx 1.27 < 2$. In fact, $\gamma_2 = 2/\sqrt{3} \approx 1.15$ is even smaller, but we will not prove this.

Exercise 6 *Find a lattice $\Lambda \subset \mathbb{R}^2$ such that $\gamma(\Lambda) = 2/\sqrt{3}$.* [Hint: The solution is commonly called the "hexagonal" lattice.]

Exercise 7 *Use Minkowski's convex body theorem to prove an upper bound on the shortest nonzero vector in any full rank n-dimensional lattice $\Lambda \subseteq \mathbb{R}^n$ with respect to the ℓ_1 norm $\|\mathbf{x}\| = \sum_i |x_i|$. How does it compare to the bound $\|\mathbf{x}\| \leq n \det(\Lambda)^{1/n}$ obtainable from Corollary 9 using norm relations $\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_1 \leq n\|\mathbf{x}\|_\infty$?*

We remark that a lattice Λ can contain vectors arbitrarily shorter than Minkowski's bound $\sqrt{n}\det(\Lambda)^{1/n}$. Consider for example the two dimensional lattice generated by the vectors $(1, 0)$ and $(0, D)$, where D is a large integer. The lattice contains a short vector of length $\lambda = 1$. However, the determinant of the lattice is D, and Minkowski's bound $\sqrt{2D}$ is much larger than 1.

It can also be shown that Minkowski's bound cannot be asymptotically improved, in the sense that there is a constant c such that for any dimension n there is a n-dimensional lattice Λ_n such that $\gamma_n > c \cdot n$. So, up to constant factors, $O(\sqrt{n})\det(\Lambda)^{1/n}$ is the best upper bound one can possibly prove on the length of the shortest vector of any n-dimensional lattice as a function of the determinant.

3 Successive Minima

The length of the shortest nonzero vector in a lattice can be equivalently defined as the radius of the smallest ball containing a nonzero lattice point. This definition is easily generalized to define a sequence of parameters $\lambda = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$, called the successive minima of the lattice. (See Figure 6.)

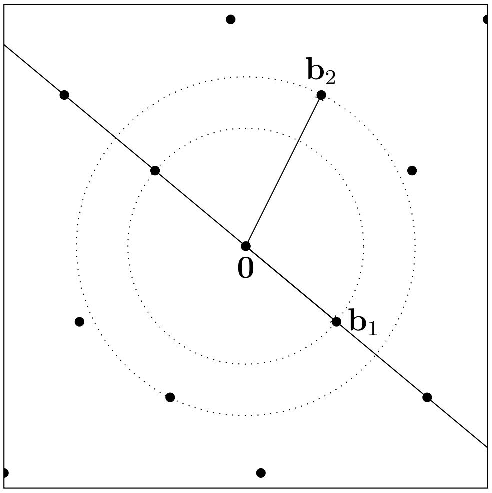

Figure 6: The successive minima of a lattice. λ_2 is the radius of the smallest ball containing two linearly independent lattice vectors.

Definition 11 *For any n-dimensional lattice Λ and integer $k \leq n$, let $\lambda_k(\Lambda)$ be the smallest $r > 0$ such that Λ contains at least k linearly independent vectors of length at most r.*

The successive minima of a lattice generalize the minimum distance $\lambda = \lambda_1$. By a volume argument similar to the one used to show that there exist vectors of length λ, one can show that there exist (linearly independent) lattice vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ of lengths $\lambda_1, \ldots, \lambda_k$. Minkowski's theorem can also be generalized to provide a bound not just on λ_1, but on the geometric mean of all successive minima.

Theorem 12 *For any lattice Λ, $\left(\prod_{i=1}^n \lambda_i\right)^{1/n} \leq \sqrt{\gamma_n} \cdot \det(\Lambda)^{1/n}$, where γ_n is Hermite's constant.*

Proof. Let Λ be an n-dimensional lattice with successive minima $\lambda_1, \ldots, \lambda_n$, and let $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \Lambda$ be linearly independent lattice vectors such that $\|\mathbf{x}_i\| = \lambda_i$. Consider the orthogonalized vectors $\mathbf{x}_i^*$ and define the transformation

$$T\left(\sum c_i \mathbf{x}_i^*\right) = \sum (c_i/\lambda_i)\mathbf{x}_i^*$$

that scales each direction $\mathbf{x}_i^*$ by a factor λ_i. We will show that the scaled lattice $T(\Lambda)$ has minimum distance at least 1. It follows, by the definition of Hermite constant, that

$$1 \leq \sqrt{\gamma_n} \cdot \det(T(\Lambda))^{1/n} \leq \sqrt{\gamma_n} \cdot \left(\frac{\det(\Lambda)}{\prod_i \lambda_i}\right)^{1/n}$$

which proves the Theorem.

We need to prove that the minimum distance of $T(\Lambda)$ is at least 1. (Equivalently, this could be formulated as showing that the open ellipsoid $E = \{\mathbf{x} \in \|T(\mathbf{x})\| < 1\}$ does not contain any nonzero lattice point from Λ.) To this end, let $\mathbf{v} \in T(\Lambda)$ be an arbitrary nonzero vector in the scaled lattice. Write $\mathbf{v}$ as $\mathbf{v} = T(\mathbf{w})$ for $\mathbf{w} \in \Lambda$, and express $\mathbf{v}, \mathbf{w}$ in terms of the orthogonalized vectors $\mathbf{v} = \sum_i (c_i/\lambda_i)\mathbf{x}_i^*$ and $\mathbf{w} = \sum_i c_i \mathbf{x}_i^*$. Take the largest $k \in 1, \ldots, n$ such that $\lambda_k \leq \|\mathbf{w}\|$. Notice that if $k < n$, then it must be $\lambda_k \leq \|\mathbf{w}\| < \lambda_{k+1}$, and by definition of λ_{k+1}, the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_k, \mathbf{w}$ must be linearly dependent. This is trivially true also when $k = n$. Since $\mathbf{x}_1, \ldots, \mathbf{x}_k$ are linearly independent, it must be that $\mathbf{w} \in \operatorname{span}(\mathbf{x}_1 \ldots, \mathbf{x}_k)$, and $c_i = 0$ for all $i > k$. It follows that

$$\|\mathbf{v}\|^2 = \sum_{i \leq k} (c_i^2/\lambda_i^2)\|\mathbf{x}_i^*\|^2 \geq \sum_{i \leq k} c_i^2 \|\mathbf{x}_i^*\|^2/\lambda_k^2 = \|\mathbf{w}\|^2/\lambda_k^2 \geq 1.$$

□

Exercise 8 *Prove that any lattice achieving Hermite's constant $\gamma_n = (\lambda(\Lambda)/\det(\Lambda)^{1/n})^2$ must necessarily have n linearly independent vectors of length $\lambda(\Lambda)$. (Equivalently, all its successive minima are the same $\lambda_1 = \lambda_2 = \ldots = \lambda_n$.)* [Hint: Use Minkowski's second theorem and Hadamard's inequality]

As for λ_1, one can define (both exact for $\gamma = 1$, and approximate for $\gamma > 1$) computational problems naturally associated to λ_n, or the problem of simultaneously achieving all successive minima.

Definition 13 *For $\gamma \geq 1$, the γ-approximate Shortest Independent Vectors Problem (*SIVP_γ*), given a basis $\mathbf{B}$ of an n-dimensional lattice, asks to find linearly independent vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathcal{L}(\mathbf{B})$ such that $\max_i \|\mathbf{v}_i\| \leq \gamma \cdot \lambda_n(\mathcal{L}(\mathbf{B}))$.*

As usual, there is also a decision (promise) problem, corresponding to the task to approximating the value of λ_n without necessarily finding short lattice vectors.

Definition 14 $\mathrm{GAPSIVP}_\gamma$*, given a basis $\mathbf{B}$ and a positive real d, asks to determine if $\lambda_n(\mathcal{L}(\mathbf{B})) \leq d$ or $\lambda_n(\mathcal{L}(\mathbf{B})) > \gamma d$.*

Sometime, the following variant of SIVP is considered, which asks to find linearly independent vectors simultaneously achieving all successive minima.

Definition 15 *For $\gamma \geq 1$, the γ-approximate Successive Minima Problem (*SMP*), given a basis $\mathbf{B}$ of an n-dimensional lattice, asks to find linearly independent vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathcal{L}(\mathbf{B})$ such that $\|\mathbf{v}_i\| \leq \gamma \cdot \lambda_i(\mathcal{L}(\mathbf{B}))$ for all i.*

Just like for SVP, it is easy to give a reduction from the decision problem $\mathrm{GAPSIVP}_\gamma$ to the search problem SIVP_γ, but giving a reduction in the other direction is an open problem.

Exercise 9 *Give a reduction from* $\mathrm{GAPSIVP}_\gamma$ *to* SIVP_γ.

Clearly, both SVP_γ and SIVP_γ reduce to SMP_γ. Reducing SMP_γ to SVP_γ or SIVP_γ (preserving the approximation factor γ) is currently an open problem.

4 A simple application in Number Theory

As an application of Minkowski's theorem we show that any prime number p congruent to 1 mod 4 can be written as the sum of two squares.

Theorem 16 *For every prime* $p \equiv 1 \bmod 4$ *there exist integers* $a, b \in \mathbb{Z}$ *such that* $p = a^2 + b^2$.

Proof. The idea is to set up a two dimensional integer lattice $\Lambda \subseteq \mathbb{Z}^2$ such that $\lambda(\Lambda)^2 = p$. Then, if (a, b) is the shortest vector in the lattice, we have $p = a^2 + b^2$.

Let $p \in \mathbb{Z}$ be a prime such that $p \equiv 1 \pmod 4$. Then $\mathbb{Z}_p^*$ is a group such that $4 \mid o(\mathbb{Z}_p^*) = p - 1$. Therefore, there exists an element of multiplicative order 4, and -1 is a quadratic residue modulo p, i.e. there exists an integer i such that $i^2 \equiv -1 \pmod p$. It immediately follows that

$$p \mid i^2 + 1. \tag{2}$$

Consider the lattice basis

$$\mathbf{B} = \begin{bmatrix} 1 & 0 \\ i & p \end{bmatrix}.$$

The lattice $\mathcal{L}(\mathbf{B})$ has determinant p. Therefore, by Minkowski's theorem there exists a nonzero lattice vector $\mathbf{v} \in \mathcal{L}(\mathbf{B})$ with squared norm at most $\|\mathbf{v}\|^2 \leq r^2$, for $\pi r^2 = 2^2 \det(\Lambda) = 4p$. Notice that $0 < \|\mathbf{v}\|^2 \leq (4/\pi)p < 2p$. In order to conclude that $\|\mathbf{v}\|^2 = p$ we observe that the squared norm of any lattice vector must be a multiple of p. In fact, for any integer vector $\mathbf{x} = (x_1, x_2)$,

$$\|\mathbf{B}\mathbf{x}\|^2 = x_1^2 + (ix_1 + px_2)^2 = (1 + i^2)x_1^2 + p(2ix_1x_2 + px_2^2)$$

which is a multiple of p by our choice of i. □

This application shows how lattices can be used to prove non-trivial facts in number theory. A similar theorem that can be proved with the same lattice techniques is the following.

Theorem 17 *Every positive integer* n *can be written as the sum of four squares* $n = a^2 + b^2 + c^2 + d^2$ *(with* $a, b, c, d \in \mathbb{Z}$*).*

The proof is left to the reader as an exercises. As you can easily guess, the proofs involves a 4-dimensional lattice.

5 Packing Radius and Covering Radius

The successive minima $\lambda_1, \ldots, \lambda_n$ provide much information about the n-dimensional geometric structure of a lattice. Two other useful parameters are the *packing radius* and the *covering radius* of the lattice. The packing radius $r(\Lambda)$ is defined as the largest radius $r > 0$ such that the open balls $\mathcal{B}(\mathbf{v}, r) = \{\mathbf{x} \colon \|\mathbf{x} - \mathbf{v}\| < r\}$ centered around all lattice points do not intersect. It is immediate to see that the packing radius of a lattice equals precisely half the minimum distance.

Exercise 10 *Show that for any lattice Λ, the packing radius of a lattice equals $\frac{1}{2}\lambda_1(\Lambda)$.*

The covering radius $\mu(\Lambda)$ is defined as the smallest radius $\mu > 0$ such that the closed balls $\bar{\mathcal{B}}(\mathbf{v}, \mu) = \{\mathbf{x} \colon \|\mathbf{x} - \mathbf{v}\| \leq \mu\}$ cover the entire space:

$$\mathrm{span}(\Lambda) \subseteq \bigcup_{\mathbf{v} \in \Lambda} \bar{\mathcal{B}}(\mathbf{v}, \mu).$$

Clearly, the covering radius is at least as large as the packing radius, i.e., $\mu(\Lambda) \geq \frac{1}{2}\lambda_1(\Lambda)$ (and the inequality is always strict, except in dimension $n = 1$), but it can be arbitrarily larger than the first minimum λ_1. In fact, the covering radius is closely related to λ_n, but only approximately, within a factor $\sqrt{n}$. This time, $\frac{1}{2}\lambda_n$ is only a lower bound, and the covering radius can be as large as $\frac{\sqrt{n}}{2}\lambda_n$.

Lemma 18 *For any lattice Λ, the covering radius is at least $\mu(\Lambda) \geq \frac{1}{2}\lambda_n(\Lambda)$.*

Proof. Fix a lattice Λ, and assume for contradiction the covering radius is $\mu < \frac{1}{2}\lambda_n$, i.e., $\epsilon = \frac{1}{2}\lambda_n(\Lambda) - \mu > 0$. We construct a set of linearly independent vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \Lambda$ of length at most $\max_i \|\mathbf{v}_i\| \leq 2\mu + \epsilon$. This is a contradiction because $2\mu + \epsilon = \lambda_n - \epsilon < \lambda_n$. The vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are build inductively as follows. For $i = 1, \ldots, n$, let $\mathbf{t}_i \in \mathrm{span}(\Lambda)$ be a vector of length $\mu + \epsilon$ orthogonal to $\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}$. By definition of covering radius, there is a lattice point within distance μ from $\mathbf{t}$. Let $\mathbf{v}_i$ be any such lattice point. By triangle inequality $\|\mathbf{v}_i\| \leq \|\mathbf{t}_i\| + \mu = 2\mu + \epsilon$. Also by triangle inequality, the distance of $\mathbf{v}_i$ from $\mathrm{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{i-1})$ is at least $\|\mathbf{t}_i\| - \mu = \epsilon > 0$. In particular, $\mathbf{v}_i \notin \mathrm{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{i-1})$ and $\mathbf{v}_i$ is linearly independent from $\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}$. This proves that the lattice vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent, and they have all length at most $2\mu + \epsilon$. □

The upper bound $\mu(\Lambda) \leq \frac{\sqrt{n}}{2}\lambda_n(\Lambda)$ easily follows from the fact that the orthogonalized parallelepiped $\mathcal{C}(\mathbf{B}^*)$ is a fundamental region.

Exercise 11 *Show that for any n-dimensional lattice Λ, the covering radius is at most $\mu(\Lambda) \leq \frac{\sqrt{n}}{2}\lambda_n(\Lambda)$.* [Hint: let $\mathbf{V} = \mathbf{v}_1, \ldots, \mathbf{v}_n$ linearly independent lattice vectors of length at most λ_n, and bound the covering radius of the sublattice $\mathcal{L}(\mathbf{V})$ using the fundamental region $\mathcal{C}(\mathbf{V}^*)$.]

Both the upper and lower bounds on the covering radius are essentially tight.

Exercise 12 *Show that for any n, there exists an n-dimensional lattice Λ such that $\mu(\Lambda) = \frac{\sqrt{n}}{2}\lambda_n(\Lambda)$.*

Exercise 13 *Show that for any real $c > \frac{1}{2}$ and any integer $n \geq 1$, there exists an n-dimensional lattice Λ such that $\mu(\Lambda) = c\lambda_n(\Lambda)$.*

In summary, the successive minima of a lattice and the covering radius satisfy

$$\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n \leq 2\mu \leq \sqrt{n}\lambda_n.$$

We have seen that there are lattices where $\lambda_n = \lambda_1$, and lattices where λ_n is almost as large as 2μ. But can λ_n be simultaneously (i.e., for the same lattice) close to both λ_1 and the covering radius? In other words, are there lattices such that the covering radius μ is not much larger (say, by a constant factor) than λ_1? We will see that this is indeed possible, although this time finding explicit constructions of lattices with $\mu = O(\lambda_1)$ is not an easy task.

To begin with, we observe that just like Minkowski's theorem gives an upper bound on the minimum distance (and packing radius) of a lattice $\frac{1}{2}\lambda_1(\Lambda) \leq (\det(\Lambda)/V_n)^{1/n}$ (where V_n is the volume of the unit ball in $\mathbb{R}^n$,) a similar volume argument can be used to prove a lower bound on the covering radius.

Exercise 14 *Prove that for any n-dimensional lattice Λ, the covering radius is at least $\mu(\Lambda) \geq (\det(\Lambda)/V_n)^{1/n}$, where V_n is the volume of the unit ball in $\mathbb{R}^n$.* [Hint: use an argument similar to the proof of Blichfeldt theorem.]

We will build a lattice Λ such that $\mu(\Lambda) = O(\lambda_1(\Lambda))$, starting from an arbitrary lattice (say $\Lambda_0 = \mathbb{Z}^n$), and then adding more points to it, so to reduce its determinant (up to the point that the lattice must necessarily have a small covering radius) without decreasing its minimum distance.

Exercise 15 *Let $\Lambda \subseteq \mathbb{R}^n$ be a full rank lattice and $\mathbf{v} \in \Lambda$ be a lattice vector. Prove that $\Lambda' = \Lambda \cup (\Lambda + \frac{\mathbf{v}}{2})$ is also a lattice. In particular, if $\frac{\mathbf{v}}{2}$ is at distance at least $\lambda_1(\Lambda)$ from Λ, then Λ' is a lattice with minimum distance $\lambda_1(\Lambda') = \lambda_1(\Lambda)$ and determinant $\det(\Lambda') = \det(\Lambda)/2$.*

Now, let Λ be a full rank lattice with covering radius $\mu(\Lambda) > 2\lambda_1(\Lambda)$. By definition of covering radius, there exists[4] a point $\mathbf{h}$ at distance μ from the lattice. (Such a point is called a *deep hole.*) Also, by definition of covering radius, there is a lattice point $\mathbf{v} \in \Lambda$ within distance μ from $2\mathbf{h}$. It follows by triangle inequality that the distance of $\frac{1}{2}\mathbf{v}$ from the lattice is at least

$$\text{dist}\left(\frac{\mathbf{v}}{2}, \Lambda\right) \geq \text{dist}(\mathbf{h}, \Lambda) - \left\|\frac{\mathbf{v}}{2} - \mathbf{h}\right\| = \mu - \frac{\|\mathbf{v} - 2\mathbf{h}\|}{2} \geq \mu - \frac{\mu}{2} = \frac{\mu}{2} \geq \lambda_1(\Lambda).$$

Exercise 16 *Prove that for any full rank lattice Λ there is a lattice $\Lambda' \supseteq \Lambda$ such that $\mu(\Lambda') \leq 2\lambda_1(\Lambda) = 2\lambda_1(\Lambda')$.* [Hint: start from Λ, and build a sequence of denser and denser lattices by adding vectors of the form $\mathbf{v}/2$.

This shows that there are lattices Λ' such that the covering radius $\mu \leq 2\lambda_1$ is within a small constant from the packing radius. You can even improve the constant.

[4] It is easy to show that the maximal distance μ is always achieved by some point. Still, if you are concerned about the existence of a point at distance exactly μ, you can relax this requirement, and simply ask $\mathbf{h}$ to be at distance at least $\frac{1}{2}\mu + \lambda_1 < \mu$ from the lattice.

Exercise 17 *Prove that for any n, there exists an n-dimensional lattice Λ such that $\mu(\Lambda) \leq 1.5 \cdot \lambda_1(\Lambda')$.* [Hint: follow the same approach as in the previous exercise, but using vectors of the form $\frac{\mathbf{v}}{3}$.]

A simple consequence of the existence of lattices with $\mu = O(\lambda_1)$ is that the bound on the minimum distance provided by Minkowski's convex body theory is tight up to a small constant factor.

Exercise 18 *Prove that $\gamma_n = \Theta(n)$. What are the best constants $c_1 \geq c_2 > 0$ for which you can show $c_2 n \leq \gamma_n \leq c_1 n$?*