**Lecture 2**
**SVP, Gram-Schmidt, LLL**

Instructor: Chris Peikert
Scribe: Hank Carter

# 1 Shortest Vector Problem

Last time we defined the minimum distance $\lambda_1(\mathcal{L})$ of a lattice $\mathcal{L} \subset \mathbb{R}^n$, and showed that it is upper bounded by $\sqrt{n} \cdot \det(\mathcal{L})^{1/n}$ (Minkowski's theorem), but this bound is often very loose. Some natural computational questions are: given a lattice (specified by some basis), can we compute its minimum distance? Can we find a vector that achieves this distance? Can we find good approximations to these? And how efficiently can we perform these tasks? These are all versions of the *Shortest Vector Problem*, which we now define formally.

**Definition 1.1 (Shortest Vector Problem, exact form).** The *exact* form of SVP has three common variants, which we restrict to *integer* lattices (and so integral bases) without any major loss of generality:[1]

1. Decision: given a lattice basis $\mathbf{B}$ and a real $d > 0$, distinguish between the cases $\lambda_1(\mathcal{L}(\mathbf{B})) \leq d$ and $\lambda_1(\mathcal{L}(\mathbf{B})) > d$.[2]
2. Calculation: given a lattice basis $\mathbf{B}$, find $\lambda_1(\mathcal{L}(\mathbf{B}))$.
3. Search: given a lattice basis $\mathbf{B}$, find a nonzero $\mathbf{v} \in \mathcal{L}(\mathbf{B})$ such that $\|\mathbf{v}\| = \lambda_1(\mathcal{L}(\mathbf{B}))$.

It is obvious that solving the Calculation version immediately lets us solve the Decision version. More formally, we say that "Decision reduces to Calculation" and write Decision $\leq$ Calculation; note the directionality of these statements. The reverse direction (Calculation $\leq$ Decision) also holds, since if we have an oracle for Decision, we can solve Calculation via binary search, by varying the choice of $d$. The only subtlety here is that in order for the search to succeed in polynomial time, the number of possible values for $\lambda_1$ must be bounded by $2^{\mathrm{poly}(|\mathbf{B}|)}$, where $|\mathbf{B}|$ is the bit length of the given basis. This is indeed the case, because the minimum distance is the square root of an integer, and is between $1$ and $\sqrt{n}\det(\mathbf{B})^{1/n}$ by Minkowski's theorem. The latter is bounded by $2^{\mathrm{poly}(|\mathbf{B}|)}$ because the determinant can be computed in polynomial time. Finally, it also turns out that the Search version is equivalent to the other two versions; we will see the proof of this later.

Also of great interest and wide applicability are *approximate* versions of SVP.

**Definition 1.2 (Approximate SVP).** The $\gamma$-approximate Shortest Vector Problem, where $\gamma = \gamma(n) \geq 1$ is a function of the dimension $n$, has the following variants (again restricted to integer lattices):

1. Decision ($\mathsf{GapSVP}_\gamma$): given a lattice basis $\mathbf{B}$ and a real $d > 0$, distinguish between the cases $\lambda_1(\mathcal{L}(\mathbf{B})) \leq d$ and $\lambda_1(\mathcal{L}(\mathbf{B})) > \gamma \cdot d$.[3]
2. Estimation ($\mathsf{EstSVP}_\gamma$): given a lattice basis $\mathbf{B}$, compute $\lambda_1(\mathcal{L}(\mathbf{B}))$ up to a $\gamma$ factor, i.e., output some $d \in [\lambda_1(\mathcal{L}(\mathbf{B})), \gamma \cdot \lambda_1(\mathcal{L}(\mathbf{B}))]$.
3. Search ($\mathsf{SVP}_\gamma$): given a lattice basis $\mathbf{B}$, find a nonzero $\mathbf{v} \in \mathcal{L}(\mathbf{B})$ such that $\|\mathbf{v}\| \leq \gamma \cdot \lambda_1(\mathcal{L}(\mathbf{B}))$.

Observe that taking $\gamma = 1$ corresponds to the exact versions of the problems (as defined above), and also that the problems cannot become harder as $\gamma$ increases. Formally, $\mathsf{GapSVP}_{\gamma'} \leq \mathsf{GapSVP}_\gamma$ for any $\gamma' \geq \gamma$ (again note the directionality), and similarly for $\mathsf{EstSVP}$ and $\mathsf{SVP}$.

[1] While there exist irrational lattices that cannot be 'scaled up' to integer lattices, in general it is not possible to represent them finitely (e.g., as input to an algorithm), so for simplicity we ignore them.

[2] Notice that because the lattice is integral, we can restrict to $d$ which are square roots of positive integers, and represent them by the positive integer $d^2$.

[3] If $\lambda_1(\mathcal{L}(\mathbf{B}))$ falls between $d$ and $\gamma \cdot d$, either answer is acceptable. Alternatively, this version can be considered as a "promise problem," where the input $\mathbf{B}$ is guaranteed to satisfy one of the two cases.

It is easy to check that

$$\mathsf{GapSVP}_\gamma \leq \mathsf{EstSVP}_\gamma \leq \mathsf{SVP}_\gamma,$$

i.e., being able to solve Search immediately implies being able to solve Estimation, which implies being able to solve Decision. It can also be seen that $\mathsf{EstSVP}_\gamma \leq \mathsf{GapSVP}_\gamma$, again using binary search. So these two variants are polynomial-time equivalent, and we usually deal with just $\mathsf{GapSVP}_\gamma$. However, for "interesting" $\gamma > 1$ it is currently unknown if solving decision is equivalent to solving search! The "interesting" qualifier is needed to rule out very large $\gamma \gtrsim 2^n$, for which both versions are solvable in polynomial time (as we will see shortly), and hence trivially equivalent.

**Open Problem 1.3.** *Prove or disprove that* $\mathsf{SVP}_\gamma \leq \mathsf{GapSVP}_\gamma$ *for some (or all) "interesting"* $\gamma > 1$.

In the remainder of the lecture we will develop tools that let us efficiently compute bounds on the minimum distance, and even find "somewhat short" nonzero lattice vectors.

## 2 Gram-Schmidt Orthogonalization

For linearly independent vectors $\mathbf{b}_1, \ldots, \mathbf{b}_n \in \mathbb{R}^n$, we define the *Gram-Schmidt orthogonalized* vectors $\widetilde{\mathbf{b}}_1, \ldots, \widetilde{\mathbf{b}}_n$ via an iterative process. First we define $\widetilde{\mathbf{b}}_1 = \mathbf{b}_1$, and then for $j = 2, \ldots, n$, we define $\widetilde{\mathbf{b}}_j$ to be the component of $\mathbf{b}_j$ orthogonal to $\operatorname{span}(\mathbf{b}_1, \ldots, \mathbf{b}_{j-1}) = \operatorname{span}(\widetilde{\mathbf{b}}_1, \ldots, \widetilde{\mathbf{b}}_{j-1})$, the linear span of the previous vectors. Formally,

$$\begin{aligned}
\widetilde{\mathbf{b}}_1 &:= \mathbf{b}_1, \\
\widetilde{\mathbf{b}}_2 &:= \mathbf{b}_2 - \mu_{1,2} \cdot \widetilde{\mathbf{b}}_1 && \text{where } \mu_{1,2} = \langle \mathbf{b}_2, \widetilde{\mathbf{b}}_1 \rangle / \langle \widetilde{\mathbf{b}}_1, \widetilde{\mathbf{b}}_1 \rangle, \\
&\vdots \\
\widetilde{\mathbf{b}}_j &:= \mathbf{b}_j - \sum_{i<j} \mu_{i,j} \cdot \widetilde{\mathbf{b}}_i && \text{where } \mu_{i,j} = \langle \mathbf{b}_j, \widetilde{\mathbf{b}}_i \rangle / \langle \widetilde{\mathbf{b}}_j, \widetilde{\mathbf{b}}_j \rangle.
\end{aligned}$$

We can verify that the vectors $\widetilde{\mathbf{b}}_j$ are mutually orthogonal. For example,

$$\langle \widetilde{\mathbf{b}}_2, \widetilde{\mathbf{b}}_1 \rangle = \langle \mathbf{b}_2, \widetilde{\mathbf{b}}_1 \rangle - \mu_{1,2} \cdot \langle \widetilde{\mathbf{b}}_1, \widetilde{\mathbf{b}}_1 \rangle = 0.$$

The general case can then be proved by induction.

It is often convenient to view the orthogonalization process as corresponding to the following (unique) matrix factorization:

$$\mathbf{B} = \underbrace{\begin{pmatrix} | & | & & | \\ \widetilde{\mathbf{b}}_1 & \widetilde{\mathbf{b}}_2 & \cdots & \widetilde{\mathbf{b}}_n \\ | & | & & | \end{pmatrix}}_{\widetilde{\mathbf{B}}} \cdot \underbrace{\begin{pmatrix} 1 & \mu_{1,2} & \cdots & \mu_{1,n} \\ & 1 & \cdots & \mu_{2,n} \\ & & \ddots & \vdots \\ & & & 1 \end{pmatrix}}_{\mathbf{U}}, \tag{2.1}$$

where the columns of $\widetilde{\mathbf{B}}$ are orthogonal and the matrix $\mathbf{U} \in \mathbb{R}^{n \times n}$ is upper unitriangular (i.e., upper triangular with 1s on the diagonal), and hence has determinant one. Note that $\mathbf{U}$ is typically *not unimodular*, because the $\mu_{i,j}$ are typically *not integers*. Therefore, $\widetilde{\mathbf{B}}$ is typically not a basis of the lattice generated by $\mathbf{B}$.

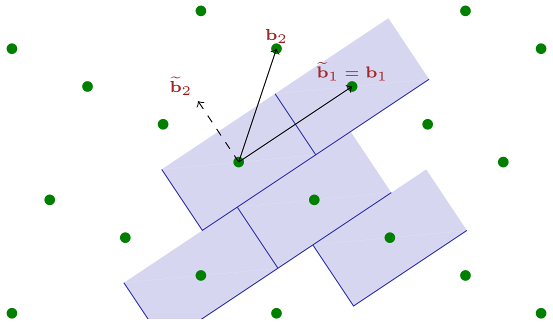
$\mathbf{b}_2$
$\widetilde{\mathbf{b}}_2$
$\widetilde{\mathbf{b}}_1 = \mathbf{b}_1$

Figure 1: An example of Gram-Schmidt orthogonalization and a (partial) tiling by the fundamental parallelepiped of the resulting vectors.

We can further factor out the lengths of the columns $\widetilde{\mathbf{b}}_i$ of $\widetilde{\mathbf{B}}$, obtaining

$$\widetilde{\mathbf{B}} = \mathbf{Q} \cdot \underbrace{\begin{pmatrix} \|\widetilde{\mathbf{b}}_1\| & & & \\ & \|\widetilde{\mathbf{b}}_2\| & & \\ & & \ddots & \\ & & & \|\widetilde{\mathbf{b}}_n\| \end{pmatrix}}_{\mathbf{D}} \tag{2.2}$$

where $\mathbf{Q}$ is an *orthogonal* matrix, i.e., $\mathbf{Q}^t\mathbf{Q} = \mathbf{I}$. This is because its columns are the mutually orthogonal *unit* vectors $\widetilde{\mathbf{b}}_i / \|\widetilde{\mathbf{b}}_i\|$. Altogether, we have the (unique) factorization

$$\mathbf{B} = \mathbf{QDU} \tag{2.3}$$

for orthogonal $\mathbf{Q}$, diagonal $\mathbf{D}$ with positive diagonal entries, and upper-unitriangular $\mathbf{U}$. This also corresponds to the so-called "QR" factorization $\mathbf{B} = \mathbf{QR}$, where $\mathbf{R} = \mathbf{DU}$ is an upper-triangular real matrix having diagonal entries $\|\widetilde{\mathbf{b}}_i\|$.

In the context of lattices, we can often ignore the orthogonal matrix $\mathbf{Q}$, taking it to be the identity matrix without loss of generality. This is because $\mathbf{Q}$ simply acts as a rigid rotation of $\mathbb{R}^n$, and therefore preserves all the relevant geometrical properties of the space, like inner products, Euclidean norms, and volumes (but not $\ell_p$ norms!). Therefore, we can usually focus on just $\mathbf{D}$ and $\mathbf{U}$.[4]

The Gram-Schmidt vectors have many important connections with the geometry of the lattice.

**Lemma 2.1.** *For any lattice* $\mathcal{L} = \mathcal{L}(\mathbf{B})$*, we have* $\det(\mathcal{L}) = \prod_{i=1}^n \|\widetilde{\mathbf{b}}_i\|$.

*Proof.* We have $\det(\mathcal{L}) = |\det(\mathbf{B})| = |\det(\mathbf{Q})\det(\mathbf{D})\det(\mathbf{U})| = \det(\mathbf{D}) = \prod_{i=1}^n \|\widetilde{\mathbf{b}}_i\|$. □

**Lemma 2.2.** *For any lattice* $\mathcal{L} = \mathcal{L}(\mathbf{B})$*, the body* $\mathcal{P}(\widetilde{\mathbf{B}}) = \widetilde{\mathbf{B}} \cdot [-\frac{1}{2}, \frac{1}{2})^n$ *is a fundamental region of* $\mathcal{L}$*.*

*Proof.* You will prove this in the homework. (Notice that the volume of $\mathcal{P}(\widetilde{\mathbf{B}})$ is $\prod_i \|\widetilde{\mathbf{b}}_i\|$, as expected.) □

[4] This all can be made formal by working with the *Gram matrix* $\mathbf{B}^t\mathbf{B} = \mathbf{U}^t\mathbf{D}(\mathbf{Q}^t\mathbf{Q})\mathbf{DU} = \mathbf{U}^t\mathbf{D}^2\mathbf{U}$ of the basis $\mathbf{B}$, which characterizes $\mathbf{B}$ and the lattice it generates up to rigid rotations. Essentially all lattice algorithms and mathematical analyses can be made to work with a Gram matrix instead of a basis.

A very useful fact is that the Gram-Schmidt vectors gives a lower bound on the lattice minimum distance.

**Lemma 2.3.** *For any lattice $\mathcal{L} = \mathcal{L}(\mathbf{B})$, we have $\lambda_1(\mathcal{L}) \geq \min_i \|\widetilde{\mathbf{b}}_i\|$.*

Before giving the full proof, let's first develop some intuition in the two-dimensional case (see Figure 2). We can partition the lattice points $\mathbf{v} = \mathbf{Bz}$ for $\mathbf{z} \in \mathbb{Z}^2$ into "slices" according to the integer coefficient $z_2$ of $\mathbf{b}_2$. If $z_2 = 0$, then $\mathbf{v}$ is in the rank-one sublattice $\mathcal{L}(\mathbf{b}_1)$, which obviously has minimum distance $\|\mathbf{b}_1\| = \|\widetilde{\mathbf{b}}_1\|$. Otherwise, $\mathbf{v}$ lies in the affine subspace $z_2\mathbf{b}_2 + \operatorname{span}(\mathbf{b}_1)$, all of whose points are at least $|z_2| \cdot \|\widetilde{\mathbf{b}}_2\| \geq \|\widetilde{\mathbf{b}}_2\|$ from the origin, and hence $\|\mathbf{v}\| \geq \|\widetilde{\mathbf{b}}_2\|$. So altogether, $\|\mathbf{v}\| \geq \min\{\|\widetilde{\mathbf{b}}_1\|, \|\widetilde{\mathbf{b}}_2\|\}$.

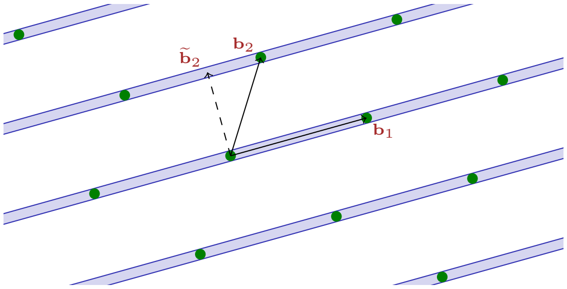

Figure 2: A two-dimensional lattice partitioned into "slices" according to the integer coefficient of $\mathbf{b}_2$.

*Proof of Lemma 2.3.* Let $\mathbf{B} = \mathbf{DU}$ be the unique factorization from Equation (2.3), where as noted above we can assume $\mathbf{Q} = \mathbf{I}$ without loss of generality. Let $\mathbf{v} = \mathbf{Bz}$ for nonzero $\mathbf{z} \in \mathbb{Z}^n$ be an arbitrary nonzero lattice point, and let $z_i$ be the last nonzero entry of $\mathbf{z}$. Then, letting $\star$s denote arbitrary real numbers, we have

$$\mathbf{v} = \mathbf{Bz} = \mathbf{D} \begin{pmatrix} 1 & \star & \star & \cdots & \star \\ & 1 & \star & \cdots & \star \\ & & 1 & \cdots & \star \\ & & & \ddots & \star \\ & & & & 1 \end{pmatrix} \begin{pmatrix} \star \\ \star \\ z_i \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \mathbf{D} \begin{pmatrix} \star \\ \star \\ z_i \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \star \\ \star \\ z_i\|\widetilde{\mathbf{b}}_i\| \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{2.4}$$

Because $|z_i| \geq 1$, we have $\|\mathbf{v}\| \geq \|\widetilde{\mathbf{b}}_i\| \geq \min_i \|\widetilde{\mathbf{b}}_i\|$, as claimed. □

Combining Minkowski's inequality with Lemmas 2.1 and 2.3, we have now have the following bounds on the minimum distance:

$$\min_i \|\widetilde{\mathbf{b}}_i\| \leq \lambda_1(\mathcal{L}(\mathbf{B})) \leq \sqrt{n} \cdot \left(\prod_{i=1}^{n} \|\widetilde{\mathbf{b}}_i\|\right)^{1/n} = \sqrt{n} \cdot \mathrm{GM}(\|\widetilde{\mathbf{b}}_i\|), \tag{2.5}$$

where GM denotes the geometric mean. While this allows us to bound $\lambda_1$ from above and below in terms of the Gram-Schmidt vectors, in the homework you will show that in general, these bounds can be arbitrarily loose (simultaneously), even in small dimensions.

## 3 Lenstra-Lenstra-Lovász (LLL) Algorithm

The LLL algorithm, due to Lenstra, Lenstra, and Lovász [LLL82], yields a polynomial-time solution to (Search-)SVP$_\gamma$ with an approximation factor $\gamma = 2^{(n-1)/2}$, which is exponential in the dimension.[5] While

[5] In fact, the algorithm can be tuned to yield an approximation factor as small as $\gamma = (2/\sqrt{3})^n$, but this is still exponential in $n$.

such a large factor might not seem so impressive at first, it is nontrivial because it depends only on the dimension $n$ of the lattice; by contrast, the bounds from Equation (2.5) depend on the lengths of the given basis vectors, which can be arbitrarily large. Also, an exponential approximation factor can be very useful when the dimension $n$ is small, or when a shortest nonzero lattice vector is much shorter than all other non-parallel lattice vectors. As we will see, at least one of these holds in many applications of LLL.

The LLL algorithm converts an arbitrary lattice basis into one that generates the same lattice, and which is "reduced" in the following sense (the notations $\mu_{i,j}$ and $\widetilde{\mathbf{b}}_i$ refer to the Gram-Schmidt orthogonalization, as in the previous section):

**Definition 3.1.** A lattice basis $\mathbf{B}$ is *LLL-reduced* if the following two conditions are met:

1. For all $i < j$, we have $|\mu_{i,j}| \leq \frac{1}{2}$. (Such a basis is said to be "size reduced.")
2. For all $i < n$, we have $\frac{3}{4}\|\widetilde{\mathbf{b}}_i\|^2 \leq \|\mu_{i,i+1}\widetilde{\mathbf{b}}_i + \widetilde{\mathbf{b}}_{i+1}\|^2$. (This is the "Lovász condition.")

The LLL conditions ensure that the lengths of the Gram-Schmidt vectors do not "decrease too quickly."

**Lemma 3.2.** *In an LLL-reduced basis* $\mathbf{B}$, *we have* $\|\widetilde{\mathbf{b}}_{i+1}\|^2 \geq \frac{1}{2}\|\widetilde{\mathbf{b}}_i\|^2$ *for all* $i < n$.

*Proof.* Since the Gram-Schmidt vectors are mutually orthogonal, by the Pythagorean theorem we have

$$\begin{aligned}\tfrac{3}{4}\|\widetilde{\mathbf{b}}_i\|^2 &\leq \|\mu_{i,i+1}\widetilde{\mathbf{b}}_i + \widetilde{\mathbf{b}}_{i+1}\|^2 \\ &= \mu_{i,i+1}^2 \cdot \|\widetilde{\mathbf{b}}_i\|^2 + \|\widetilde{\mathbf{b}}_{i+1}\|^2 \\ &\leq \tfrac{1}{4}\|\widetilde{\mathbf{b}}_i\|^2 + \|\widetilde{\mathbf{b}}_{i+1}\|^2.\end{aligned}$$

The claim follows by rearranging terms. □

Because the Gram-Schmidt vectors of a lattice basis give a lower bound on the lattice's minimum distance, it follows that the first vector in an LLL-reduced basis approximates a shortest lattice vector.

**Corollary 3.3.** *In an LLL-reduced basis* $\mathbf{B}$, *we have* $\|\mathbf{b}_1\| \leq 2^{(n-1)/2} \cdot \lambda_1(\mathcal{L}(\mathbf{B}))$.

*Proof.* Recall that $\mathbf{b}_1 = \widetilde{\mathbf{b}}_1$, so $\|\mathbf{b}_1\| = \|\widetilde{\mathbf{b}}_1\|$. By Lemma 3.2, we also have $\|\widetilde{\mathbf{b}}_{i+1}\| \geq \frac{1}{\sqrt{2}}\|\widetilde{\mathbf{b}}_i\|$ for $1 \leq i < n$. Therefore,

$$\|\mathbf{b}_1\| \leq 2^{(i-1)/2} \cdot \|\widetilde{\mathbf{b}}_i\| \leq 2^{(n-1)/2} \cdot \|\widetilde{\mathbf{b}}_i\|$$

for all $i$. From this and Lemma 2.3 we conclude that $\|\mathbf{b}_1\| \leq 2^{(n-1)/2} \cdot \min_i \|\widetilde{\mathbf{b}}_i\| \leq 2^{(n-1)/2} \cdot \lambda_1(\mathcal{L}(\mathbf{B}))$.□

We will describe the LLL algorithm itself, and its ingenious analysis, in the next lecture.

## References

[LLL82] A. K. Lenstra, H. W. Lenstra, Jr., and L. Lovász. Factoring polynomials with rational coefficients. *Mathematische Annalen*, 261(4):515–534, December 1982. Page 4.