

15.083J/6.859J Integer Optimization

Lecture 13: Lattices II

1 Outline

SLIDE 1

- Gram-Schmidt (GS) Orthogonalization.
- Reduced bases for lattices.
- Simultaneous Diophantine approximation.

2 GS orthogonalization

SLIDE 2

- **Input:** n linearly independent vectors $\mathbf{b}^1, \dots, \mathbf{b}^n \in \mathcal{Q}^n$
- **Output:** n linearly independent vectors $\tilde{\mathbf{b}}^1, \dots, \tilde{\mathbf{b}}^n$ that are orthogonal and span the same linear space.
- **Algorithm:**

1. **(Initialization)** $\tilde{\mathbf{b}}^1 = \mathbf{b}^1$.
2. **(Main iteration)** For $i = 2, \dots, n$, set:

$$\mu_{i,j} = \frac{(\mathbf{b}^i)' \tilde{\mathbf{b}}^j}{\|\tilde{\mathbf{b}}^j\|^2} \text{ for } j = 1, \dots, i-1,$$
$$\tilde{\mathbf{b}}^i = \mathbf{b}^i - \sum_{j=1}^{i-1} \mu_{i,j} \tilde{\mathbf{b}}^j.$$

2.1 Intuition

SLIDE 3

- To initialize $\tilde{\mathbf{b}}^1 = \mathbf{b}^1$.
- Decompose $\mathbf{b}^2 = \mathbf{v} + \mathbf{u}$, such that $\mathbf{v} = \lambda \mathbf{b}^1$ for some $\lambda \in \mathcal{R}$ and \mathbf{u} is orthogonal to \mathbf{b}^1 , i.e., $\mathbf{u}' \mathbf{b}^1 = 0$.
- Multiplying $\mathbf{b}^2 = \mathbf{v} + \mathbf{u}$ by \mathbf{b}^1 , $(\mathbf{b}^2)' \mathbf{b}^1 = \lambda \|\mathbf{b}^1\|^2$:

$$\lambda = \frac{(\mathbf{b}^2)' \mathbf{b}^1}{\|\mathbf{b}^1\|^2},$$

$$\tilde{\mathbf{b}}^2 = \mathbf{u} = \mathbf{b}^2 - \mathbf{v} = \mathbf{b}^2 - \lambda \mathbf{b}^1$$

- Geometrically $\tilde{\mathbf{b}}^2$ corresponds to projecting \mathbf{b}^2 to the subspace that is orthogonal to \mathbf{b}^1 .

2.2 Properties

SLIDE 4

- $(\tilde{\mathbf{b}}^i)' \tilde{\mathbf{b}}^j = 0$ for all $i \neq j$.
- $\{\mathbf{x} \in \mathcal{R}^n \mid \mathbf{x} = \sum_{i=1}^k \lambda_i \mathbf{b}^i, \lambda \in \mathcal{R}^k\} = \{\mathbf{x} \in \mathcal{R}^n \mid \mathbf{x} = \sum_{i=1}^k \lambda_i \tilde{\mathbf{b}}^i, \lambda \in \mathcal{R}^k\}$ for $k = 1, \dots, n$.
- $\det(\mathcal{L}(\mathbf{b}^1, \dots, \mathbf{b}^n)) = \prod_{j=1}^n \|\tilde{\mathbf{b}}^j\|$.
- $\|\tilde{\mathbf{b}}^j\| \leq \|\mathbf{b}^j\|$ for $j = 1, \dots, n$.

2.3 Example

SLIDE 5

- $\mathbf{b}^1 = (4, 1)'$ and $\mathbf{b}^2 = (1, 1)'$.
- The GS orthogonalization: $\tilde{\mathbf{b}}^1 = \mathbf{b}^1$ and

$$\tilde{\mathbf{b}}^2 = \mathbf{b}^2 - \mu_{2,1} \tilde{\mathbf{b}}^1 = (1, 1)' - \frac{5}{17} \tilde{\mathbf{b}}^1 = \frac{1}{17}(-3, 12)',$$

- Note that $\tilde{\mathbf{b}}^1, \tilde{\mathbf{b}}^2$ do not form a basis of \mathcal{L} .
- The GS orthogonalization depends on the order in which the vectors are processed.
- Consider $\mathbf{b}^1 = (1, 1)'$ and $\mathbf{b}^2 = (4, 1)'$. The GS orthogonalization $\tilde{\mathbf{b}}^1 = \mathbf{b}^1$, $\mu_{2,1} = 5/2$ and $\tilde{\mathbf{b}}^2 = (1/2)(3, -3)'$

2.4 Nearest vector

SLIDE 6

Given $x \in \mathcal{R}$:

$$[x] = \begin{cases} [x], & \text{if } 0 \leq x - [x] \leq \frac{1}{2}, \\ \lceil x \rceil, & \text{if } \frac{1}{2} < x - [x] \leq 1. \end{cases}$$

$[1.5] = 1$, $[3.7] = 4$ and $[5.2] = 5$.

Let $\mathbf{b}^1, \dots, \mathbf{b}^n$ be a basis of the lattice \mathcal{L} with GS $\tilde{\mathbf{b}}^1, \dots, \tilde{\mathbf{b}}^n$.

- For every $\mathbf{z} \in \mathcal{L} \setminus \{\mathbf{0}\}$,

$$\|\mathbf{z}\| \geq \min\{\|\tilde{\mathbf{b}}^1\|, \dots, \|\tilde{\mathbf{b}}^n\|\}.$$

- If $\tilde{\mathbf{b}}^1, \dots, \tilde{\mathbf{b}}^n$ is a basis of \mathcal{L} , then the nearest vector in \mathcal{L} to the vector $\mathbf{x} = \sum_{j=1}^n \lambda_j \tilde{\mathbf{b}}^j$, $\lambda \in \mathcal{R}^n$ is given by:

$$\mathbf{b}^* = \sum_{j=1}^n \mu_j \tilde{\mathbf{b}}^j, \quad \text{where } \mu_j = \lfloor \lambda_j \rfloor.$$

2.5 Proof

SLIDE 7

- $\mathbf{0} \neq \mathbf{z} = \sum_{i=1}^n \sigma_i \mathbf{b}^i$ with $\sigma_i \in \mathcal{Z}$, $i = 1, \dots, n$.
- Let k be the largest index such that $\sigma_k \neq 0$, i.e., $|\sigma_k| \geq 1$,

$$\begin{aligned} \mathbf{z} &= \sum_{i=1}^k \sigma_i \left(\tilde{\mathbf{b}}^i + \sum_{j=1}^{i-1} \mu_{i,j} \tilde{\mathbf{b}}^j \right) \\ &= \sigma_k \tilde{\mathbf{b}}^k + \sum_{j=1}^{k-1} \left(\sigma_j + \sum_{i=j+1}^k \sigma_i \mu_{i,j} \right) \tilde{\mathbf{b}}^j \\ &= \sigma_k \tilde{\mathbf{b}}^k + \sum_{j=1}^{k-1} \lambda_j \tilde{\mathbf{b}}^j, \end{aligned}$$

where $\lambda_j = \sigma_j + \sum_{i=j+1}^k \sigma_i \mu_{i,j}$.

- Since $(\tilde{\mathbf{b}}^i)' \tilde{\mathbf{b}}^j = 0$,

$$\|\mathbf{z}\|^2 = \mathbf{z}'\mathbf{z} = \sum_{j=1}^{k-1} \lambda_j^2 \|\tilde{\mathbf{b}}^j\|^2 + \sigma_k^2 \|\tilde{\mathbf{b}}^k\|^2 \geq \sigma_k^2 \|\tilde{\mathbf{b}}^k\|^2 \geq \|\tilde{\mathbf{b}}^k\|^2.$$

- $\|\mathbf{z}\| \geq \|\tilde{\mathbf{b}}^k\| \geq \min\{\|\tilde{\mathbf{b}}^1\|, \dots, \|\tilde{\mathbf{b}}^n\|\}$.

SLIDE 8

- $\mathbf{b} = \sum_{j=1}^n \nu_j \tilde{\mathbf{b}}^j$ with $\nu_j \in \mathcal{Z}$, be an arbitrary vector of the lattice \mathcal{L} .
- Let $\mathbf{x} = \sum_{i=1}^n \lambda_i \tilde{\mathbf{b}}^i$, $\lambda \in \mathcal{R}^n$. Then,

$$\|\mathbf{b} - \mathbf{x}\|^2 = \sum_{j=1}^n (\nu_j - \lambda_j)^2 \|\tilde{\mathbf{b}}^j\|^2 \geq \sum_{j=1}^n (\mu_j - \lambda_j)^2 \|\tilde{\mathbf{b}}^j\|^2 = \|\mathbf{b}^* - \mathbf{x}\|^2.$$

- For all $\mathbf{b} \in \mathcal{L}$, $\|\mathbf{b} - \mathbf{x}\| \geq \|\mathbf{b}^* - \mathbf{x}\|$.
- Importance of orthogonality.

3 Reduced Bases

3.1 Definition

SLIDE 9

Let $\mathcal{L} = \mathcal{L}(\mathbf{b}^1, \dots, \mathbf{b}^n)$ with $\mathbf{b}^1, \dots, \mathbf{b}^n \in \mathcal{Q}^n$ and with GS: $\tilde{\mathbf{b}}^1, \dots, \tilde{\mathbf{b}}^n$. The basis $\{\mathbf{b}^1, \dots, \mathbf{b}^n\}$ is called **reduced** if the following conditions hold:

- (a) $|\mu_{i,j}| \leq \frac{1}{2}$, for all i, j with $1 \leq j < i \leq n$,
- (b) $\|\tilde{\mathbf{b}}^{i+1} + \mu_{i+1,i} \tilde{\mathbf{b}}^i\|^2 \geq \frac{3}{4} \|\tilde{\mathbf{b}}^i\|^2$, for all $i = 1, \dots, n-1$.

3.2 Intuition

SLIDE 10

- Conditions (a) and (b) jointly imply that a reduced basis consists of nearly orthogonal vectors.
- $\tilde{\mathbf{b}}^1 = \mathbf{b}^1$, condition (a) for $i = 2$ implies that

$$\mu_{2,1} = \frac{(\mathbf{b}^2)' \mathbf{b}^1}{\|\mathbf{b}^1\|^2} \leq \frac{1}{2}.$$

- From GS $\mathbf{b}^2 = \tilde{\mathbf{b}}^2 + \mu_{2,1} \tilde{\mathbf{b}}^1$, and thus (b) for $i = 1$ $\|\mathbf{b}^2\|^2 \geq \frac{3}{4} \|\mathbf{b}^1\|^2$.
- Let θ be the angle between the two vectors \mathbf{b}^1 and \mathbf{b}^2 . Then

$$\cos \theta = \frac{(\mathbf{b}^2)' \mathbf{b}^1}{\|\mathbf{b}^2\| \|\mathbf{b}^1\|} = \frac{(\mathbf{b}^2)' \mathbf{b}^1 \|\mathbf{b}^1\|}{\|\mathbf{b}^1\|^2 \|\mathbf{b}^2\|} \leq \frac{1}{2} \frac{2}{\sqrt{3}} = \frac{1}{\sqrt{3}}.$$

This implies that $\theta \geq \cos^{-1}(1/\sqrt{3}) = 54.7^\circ$,

- For the purpose of achieving a bigger angle between the two vectors, that is, bringing the vectors closer to orthogonality, we would like to have as high a constant c as possible. For $c = 1$, conditions (a) and (b) imply that an angle θ would be at least $\cos^{-1}(1/2) = 60^\circ$.

3.3 Properties

SLIDE 11

For a reduced basis $\mathbf{b}^1, \dots, \mathbf{b}^n$ of a lattice \mathcal{L} and its GS $\tilde{\mathbf{b}}^1, \dots, \tilde{\mathbf{b}}^n$:

- (a) $\|\tilde{\mathbf{b}}^j\|^2 \geq 2^{i-j} \|\tilde{\mathbf{b}}^i\|^2$ for all $1 \leq i < j \leq n$.
- (b) $\|\mathbf{b}^1\| \leq 2^{(n-1)/4} \det(\mathcal{L})^{1/n}$.
- (c) $\|\mathbf{b}^1\| \leq 2^{(n-1)/2} \min\{\|\mathbf{b}\| : \mathbf{b} \in \mathcal{L} \setminus \{\mathbf{0}\}\}$.
- (d) $\|\mathbf{b}^1\| \cdots \|\mathbf{b}^n\| \leq 2^{(n(n-1))/4} \det(\mathcal{L})$.

3.4 Proof

SLIDE 12

- For all $i = 1, \dots, n-1$:

$$\begin{aligned} \frac{3}{4} \|\tilde{\mathbf{b}}^i\|^2 &\leq \|\tilde{\mathbf{b}}^{i+1} + \mu_{i+1,i} \tilde{\mathbf{b}}^i\|^2 \\ &= \|\tilde{\mathbf{b}}^{i+1}\|^2 + \mu_{i+1,i}^2 \|\tilde{\mathbf{b}}^i\|^2 \\ &\leq \|\tilde{\mathbf{b}}^{i+1}\|^2 + \frac{1}{4} \|\tilde{\mathbf{b}}^i\|^2. \end{aligned}$$

This gives

$$\|\tilde{\mathbf{b}}^{i+1}\|^2 \geq \frac{1}{2} \|\tilde{\mathbf{b}}^i\|^2, \quad \text{for all } i = 1, \dots, n-1,$$

leading to

$$\|\tilde{\mathbf{b}}^j\|^2 \geq 2^{i-j} \|\tilde{\mathbf{b}}^i\|^2, \quad \text{for all } 1 \leq i < j \leq n.$$

- Applying part (a) for $i = 1$ we obtain

$$\|\tilde{\mathbf{b}}^j\|^2 \geq 2^{1-j} \|\tilde{\mathbf{b}}^1\|^2 = 2^{1-j} \|\mathbf{b}^1\|^2, \quad \text{for all } 1 \leq j \leq n.$$

From Proposition 6.2(c), we have

$$\det(\mathcal{L})^2 = \prod_{j=1}^n \|\tilde{\mathbf{b}}^j\|^2 \geq \left(\prod_{j=1}^n 2^{1-j} \right) \|\mathbf{b}^1\|^{2n} = \left(\frac{1}{2} \right)^{(n(n-1))/2} \|\mathbf{b}^1\|^{2n},$$

proving part (b).

- From Proposition 6.3, we have that for every $\mathbf{b} \in \mathcal{L} \setminus \{\mathbf{0}\}$,

$$\|\mathbf{b}\|^2 \geq \min\{\|\tilde{\mathbf{b}}^j\|^2 : j = 1, \dots, n\} \geq 2^{1-n} \|\mathbf{b}^1\|^2,$$

proving part (c).

- From GS, Proposition 6.2 and the definition of a reduced basis we obtain

$$\begin{aligned} \|\mathbf{b}^i\|^2 &= \|\tilde{\mathbf{b}}^i\|^2 + \sum_{j=1}^{i-1} \mu_{i,j}^2 \|\tilde{\mathbf{b}}^j\|^2 \leq \|\tilde{\mathbf{b}}^i\|^2 + \frac{1}{4} \sum_{j=1}^{i-1} \|\tilde{\mathbf{b}}^j\|^2 \\ &\leq \|\tilde{\mathbf{b}}^i\|^2 + \frac{1}{4} \sum_{j=1}^{i-1} 2^{i-j} \|\tilde{\mathbf{b}}^i\|^2 \\ &= \|\tilde{\mathbf{b}}^i\|^2 \left(1 + \frac{1}{4} (2 + \dots + 2^{i-1}) \right) \\ &= \|\tilde{\mathbf{b}}^i\|^2 \left(1 + \frac{1}{4} (2^i - 2) \right) \\ &\leq \|\tilde{\mathbf{b}}^i\|^2 2^{i-1}. \end{aligned}$$

Using Proposition 6.2(c) we obtain

$$\prod_{i=1}^n \|\mathbf{b}^i\|^2 \leq 2^{(n(n-1))/2} \prod_{i=1}^n \|\tilde{\mathbf{b}}^i\|^2 = 2^{(n(n-1))/2} \det(\mathcal{L})^2,$$

proving part (d).

- From Minkowski, \mathcal{L} there exist a vector $\mathbf{u} \in \mathcal{L}$ such that $\|\mathbf{u}\|_\infty \leq \det(\mathcal{L})^{1/n}$. In contrast, $\|\mathbf{b}^1\|_\infty \leq \|\mathbf{b}^1\|_2 \leq 2^{(n-1)/4} \det(\mathcal{L})^{1/n}$ is weaker. The key difference is that we can find the vector \mathbf{b}^1 in polynomial time.

3.5 Algorithm 6.2

SLIDE 13

- **Input:** A basis $\mathbf{b}^1, \dots, \mathbf{b}^n \in \mathcal{Z}^n$ of a lattice \mathcal{L} .
- **Output:** A basis of \mathcal{L} satisfying condition (a)
- **Algorithm:**
 1. For $i = 2, \dots, n$
For $j = i - 1, \dots, 1$
 - (a) If $|\mu_{i,j}| > 1/2$, then set $\mathbf{b}^i = \mathbf{b}^i - \lfloor \mu_{i,j} \rfloor \mathbf{b}^j$.
 - (b) Compute the GS of $\mathbf{b}^1, \dots, \mathbf{b}^n$ and the corresponding multipliers $\mu_{i,j}$.
 2. Return $\mathbf{b}^1, \dots, \mathbf{b}^n$.

3.6 Correctness

SLIDE 14

- The basis returned by Algorithm 6.2 satisfies condition (a).
- Algorithm 6.2 requires $O(n^4)$ arithmetic operations.
- Algorithm 6.2 has the invariance property that after each iteration the GS of the initial basis of \mathcal{L} remains unchanged, i.e.,

$$\tilde{\mathbf{b}}^i = \tilde{\mathbf{q}}^i \text{ for all } i = 1, \dots, n.$$

3.7 Basis Reduction

SLIDE 15

- **Input:** A basis $\mathbf{b}^1, \dots, \mathbf{b}^n \in \mathcal{Z}^n$ of a lattice \mathcal{L} .
- **Output:** A basis of \mathcal{L} satisfying conditions (a) and (b).
- **Algorithm:**
 1. Compute the Gram-Schmidt orthogonalization $\tilde{\mathbf{b}}^1, \dots, \tilde{\mathbf{b}}^n$ of the vectors $\mathbf{b}^1, \dots, \mathbf{b}^n$.
 2. Apply Algorithm 6.2.
 3. For $i = 1, \dots, n$
If $\|\tilde{\mathbf{b}}^{i+1} + \mu_{i+1,i} \tilde{\mathbf{b}}^i\|^2 < 3/4 \|\tilde{\mathbf{b}}^i\|^2$, then interchange \mathbf{b}^i and \mathbf{b}^{i+1} and return to Step 1.
 4. Return $\mathbf{b}^1, \dots, \mathbf{b}^n$.

3.8 Polynomiality

SLIDE 16

Let $\mathbf{b}^1, \dots, \mathbf{b}^n \in \mathcal{Z}^n$ be a basis of the lattice \mathcal{L} . The basis reduction algorithm returns a reduced basis of \mathcal{L} by performing $O(n^6 \log_2 b_{\max})$ arithmetic operations, where b_{\max} is the largest integer (in absolute value) among the entries in $\mathbf{b}^1, \dots, \mathbf{b}^n$.

4 Simultaneous diophantine approximation

SLIDE 17

- For given numbers $\alpha_1, \dots, \alpha_n \in \mathcal{Q}$, $0 < \epsilon < 1$ and a given integer number $N > 1$, find $p_1, \dots, p_n \in \mathcal{Z}$ and $q \in \mathcal{Z}_+$ with $0 < q \leq N$ satisfying:

$$\left| \alpha_i - \frac{p_i}{q} \right| < \frac{\epsilon}{q} \quad \text{for } i \in \{1, \dots, n\}. \quad (*)$$

- If $N \geq \epsilon^{-n}$, then there exist $p_1, \dots, p_n \in \mathcal{Z}$ and $q \in \mathcal{Z}_+$ with $0 < q \leq N$ satisfying (*).
- Proof We define a lattice $\mathcal{L} = \mathcal{L}(\mathbf{b}^0, \dots, \mathbf{b}^n) \subseteq \mathcal{Q}^{n+1}$ where

$$\mathbf{b}^0 = (\alpha_1, \dots, \alpha_n, \delta)', \quad \mathbf{b}^i = -\mathbf{e}_i, \quad i = 1, \dots, n,$$

$$\delta = \epsilon^{n+1}.$$

- Since $\det(\mathcal{L}) = \delta = \epsilon^{n+1}$ and $\dim(\mathcal{L}) = n + 1$, from Convex body theorem we obtain that there exists an $\mathbf{a} \in \mathcal{L}$, $\mathbf{a} \neq \mathbf{0}$ with $\|\mathbf{a}\|_\infty \leq (\det(\mathcal{L}))^{1/(n+1)} = \epsilon$. Hence, there exist $q, p_1, \dots, p_n \in \mathcal{Z}$ such that

$$\mathbf{a} = q\mathbf{b}^0 + \sum_{i=1}^n p_i \mathbf{b}^i,$$

with $|a_i| \leq \epsilon$, or equivalently

$$|a_i| = |q\alpha_i - p_i| \leq \epsilon, \quad i = 1, \dots, n$$

$$a_n = q\delta \leq \epsilon, \quad \text{i.e., } q \leq \epsilon^{-n}.$$

- To complete the proof we need to check that $q > 0$. Note that we assume without loss of generality that $q \geq 0$, since we can always take $-\mathbf{a}$ instead of \mathbf{a} . If $q = 0$, then $|p_i| \leq \epsilon$ for all i . Since $p_i \in \mathcal{Z}$ and $0 < \epsilon < 1$, we have $p_i = 0$. This leads to $\mathbf{a} = \mathbf{0}$, which is a contradiction since $\mathbf{a} \neq \mathbf{0}$.

4.1 Using Basis Reduction

SLIDE 18

- Theorem If $N \geq 2^{n(n+1)/4} \epsilon^{-n}$, we can find in polynomial time $p_1, \dots, p_n \in \mathcal{Z}$ and $q \in \mathcal{Z}_+$ with $0 < q \leq N$ satisfying Eq. (*).
- $\delta = 2^{-n(n+1)/4} \epsilon^{n+1}$ in the basis for the lattice \mathcal{L} defined earlier.
- Applying Basis Reduction we find in polynomial time a reduced basis of \mathcal{L} . The first vector $\mathbf{c} \in \mathcal{L}$ in the reduced basis satisfies (recall that we use $n + 1$ instead of n , since $\dim(\mathcal{L}) = n + 1$)

$$\|\mathbf{c}\|_\infty \leq \|\mathbf{c}\|_2 \leq 2^{n/4} \det(\mathcal{L})^{1/(n+1)} = 2^{n/4} \delta^{1/(n+1)} = \epsilon.$$

Hence, we can find $p_1, \dots, p_n \in \mathcal{Z}$ and $q \in \mathcal{Z}_+$ such that

$$\mathbf{c} = q\mathbf{b}^0 + \sum_{i=1}^n p_i \mathbf{b}^i,$$

with $|c_i| \leq \epsilon$, or equivalently

$$|c_i| = |q\alpha_i - p_i| \leq \epsilon, \quad i = 1, \dots, n$$

$$c_n = q\delta \leq \epsilon, \text{ i.e., } q \leq 2^{n(n+1)/4} \epsilon^{-n}.$$

MIT OpenCourseWare
<http://ocw.mit.edu>

15.083J / 6.859J Integer Programming and Combinatorial Optimization
Fall 2009

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.