Sparse Vectors of Small Height Avoiding Hyperplanes

Ze, FAN

Department of Mathematics and Computer Science, Wesleyan University

zfan@wesleyan.edu

1. Siegel's Lemma

In his celebrated paper [6], Siegel proved the following lemma:

Theorem 1. Given an integer $M \times N$ matrix A, $M < N$, there exists an non-zero vector $\boldsymbol{\xi} \in \mathbb{Z}^N$ such that $A\boldsymbol{\xi} = \mathbf{0}$ and

$$|\boldsymbol{\xi}| \leq 2 + (N|A|)^{\frac{1}{N-M}},$$

where the $|\cdot|$ sign denotes the sup-norm of the vector and the matrix, respectively.

The intuition behind this result is straightforward: if a system of homogeneous linear equations has small integer coefficients, then it will also have a solution in small integers.

However, this intuition is insufficient because the bound it provided is not invariant under invertible linear transformations. It can be noticed that for any $M \times M$ invertible integer matrix B, $A\boldsymbol{\xi} = \mathbf{0}$ if and only if $(BA)\boldsymbol{\xi} = \mathbf{0}$, but $|BA|$ and $|A|$ might be quite different. Thus, it is also useful to interpret Siegel's Lemma as a result claiming the existence of an integral point with small "size" inside the $N-M$-dimensional kernel of the linear transformation provided by the matrix A. In this context (i.e. $\boldsymbol{\xi} \in \mathbb{Z}^N$), the "size" of the vector is measured by its sup-norm (that is the largest value among all the absolute values of its coordinates). However, if we want to extend the range of $\boldsymbol{\xi}$ to the ring of integers of an arbitrary number field, the "size" of $\boldsymbol{\xi}$ should be measured by a more generalized definition, height. To define the height, appropriate setup in valuation theory is required.

2. Valuation Theory

Definition 1. Let k be an number field. A valuation $|\cdot|_{\mathfrak{p}}$ on k is a non-negative real valued function defined on k satisfying the following three properties:

(1) $|a|_{\mathfrak{p}} = 0$ if and only if $a = 0$.

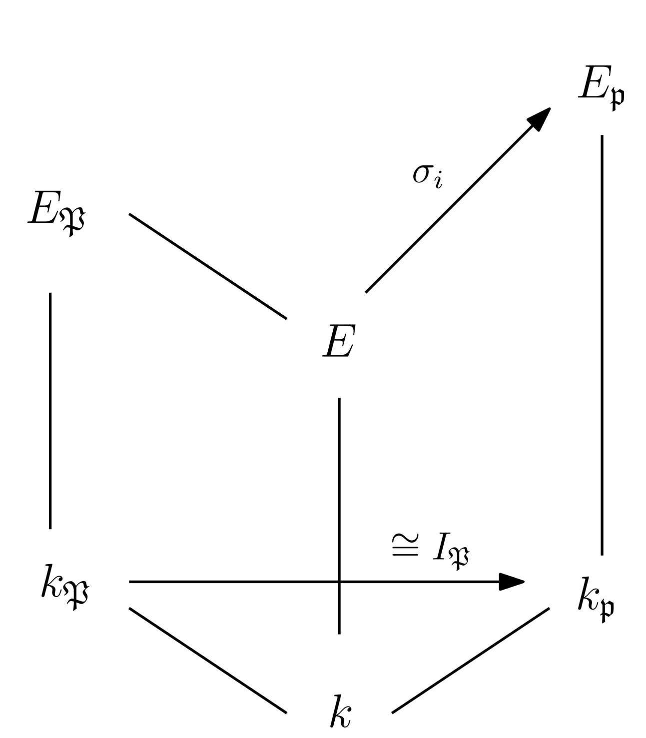
(2) $|ab|_{\mathfrak{p}} = |a|_{\mathfrak{p}}|b|_{\mathfrak{p}}$ for all $a, b \in k$.

(3) (Triangle inequality) $|a+b|_{\mathfrak{p}} \leq |a|_{\mathfrak{p}} + |b|_{\mathfrak{p}}$ for all $a, b \in k$.

If the valuation satisfies a even stronger ultra-triangle inequality (that is, $|a+b|_{\mathfrak{p}} \leq \max\{|a|_{\mathfrak{p}}, |b|_{\mathfrak{p}}\}$), we call this valuation non-Archimedean. Otherwise we call it Archimedean.

The set of valuations on $\mathbb{Q}$ is well-studied: Ostrowski has proved that the only equivalence classes of valuations (called *places*) of $\mathbb{Q}$ are the class of the classic absolute value, and the class of the p-adic valuations for each prime number p. It is also easy to extend a valuation to the smallest complete field that contains it using the canonical Cauchy sequence construction. Let $\mathfrak{p}$ be a place on a number field k with characteristic zero. We can then define a metric d on k using a valuation that represents $\mathfrak{p}$. After constructing a complete field $k_{\mathfrak{p}}$ by Cauchy sequences in k respect to the metric d, we define a valuation on $k_{\mathfrak{p}}$ by $|\alpha| = d(\alpha, 0)$. The place containing this new valuation have nice existential and uniqueness property.

To extend a place onto a non-complete finite extension field, a slightly more sophisticated construction is needed.



We begin by setting k a number field, $\mathrm{Char}(k) = 0$ and $E = k(\alpha)$ a finite extension with degree d and minimal polynomial f. For a place $\mathfrak{p}$ of k, there exists a completion $k_{\mathfrak{p}}$ of k with respect to $\mathfrak{p}$. Define $E_{\mathfrak{p}}$ to be the splitting field of f over $k_{\mathfrak{p}}$, then it can be shown that there exists a unique place $\mathfrak{P}_0$ on $E_{\mathfrak{p}}$ that divides $\mathfrak{p}$ on $k_{\mathfrak{p}}$ and that is complete over $E_{\mathfrak{p}}$. By basic field theory, there exists exactly d k-embeddings of E into $E_{\mathfrak{p}}$. We can name them by $\sigma_1, \cdots, \sigma_n$, respectively. Each $\sigma_i^{-1} : \sigma_i(E) \to E$ induces a pull-back place $\mathfrak{P}_0^{\sigma_i^{-1}}$ on E, which divides the pull-back $\mathfrak{p}^{\sigma_i^{-1}} = \mathfrak{p}$. It can be proved that all places $\mathfrak{P}$ on E that divides $\mathfrak{p}$ must equals to some $\mathfrak{P}_0^{\sigma_i^{-1}}$ for some i, and $\mathfrak{P}_0^{\sigma_i^{-1}} = \mathfrak{P}_0^{\sigma_j^{-1}}$ if and only if $\sigma_i(\alpha)$ and $\sigma_j(\alpha)$ are conjugates in $k_{\mathfrak{p}}$. Therefore, we know that it is indeed possible to extend an arbitrary place to a finite extension field, and there are at most d of them.

Another important index shall be defined. Let $\mathfrak{P}$ on E be an extension of the place $\mathfrak{p}$. Also define $E_{\mathfrak{P}}$ bo be the completion of E with respect to $\mathfrak{P}$, and $k_{\mathfrak{P}}$ be the completion of k inside $E_{\mathfrak{P}}$, or we say the closure of k. Then, the *local degree* of $\mathfrak{P}$ is defined by $d_{\mathfrak{P}} = [E_{\mathfrak{P}} : k_{\mathfrak{P}}]$, the degree of this finite extension.

3. Height of Vectors and Spaces

The definition of height relies on the valuation theory. Let k be a number field, $\mathfrak{p}$ a place on k. For an vector $\boldsymbol{x} = (x_1, \cdots, x_N) \in k^N$, we define the *local height* of $\boldsymbol{x}$ at $\mathfrak{p}$ by

$$H_{\mathfrak{p}}(\boldsymbol{x}) = \begin{cases} \left(\sum_{i=1}^N |x_i|_{\mathfrak{p}}^2\right)^{\frac{1}{2}}, & \mathfrak{p} \text{ is Archimedean}, \\ \max_{i=1}^N \{|x_i|_{\mathfrak{p}}\}, & \mathfrak{p} \text{ is non-Archimedean}. \end{cases}$$

For $\boldsymbol{x} \neq \mathbf{0}$, we define the *global height* of $\boldsymbol{x}$ by

$$H_k(\boldsymbol{x}) = \prod_{\mathfrak{p} \in M(k)} H_{\mathfrak{p}}((x))^{d_{\mathfrak{p}}}.$$

Another equally useful local height can be defined by

$$\mathcal{H}_{\mathfrak{p}}(\boldsymbol{x}) = \max_{i=1}^N \{|x_i|_{\mathfrak{p}}\}$$

for all Archimedean places $\mathfrak{p}$, and this leads us to a second global height

$$\mathcal{H}_k(\boldsymbol{x}) = \prod_{\mathfrak{p}|\infty} \mathcal{H}_{\mathfrak{p}}(\boldsymbol{x}) \cdot \prod_{\mathfrak{p}\nmid\infty} H_{\mathfrak{p}}(\boldsymbol{x}).$$

Finally, to make height an invariant under changes of k, we define two *absolute height* by

$$H(\boldsymbol{x}) = \left(H_k(\boldsymbol{x})\right)^{1/d}, \quad \mathcal{H}(\boldsymbol{x}) = \left(\mathcal{H}_k(\boldsymbol{x})\right)^{1/d}.$$

It is observed that $\mathcal{H}(\boldsymbol{x}) \leq H(\boldsymbol{x}) \leq \sqrt{N}\mathcal{H}(\boldsymbol{x})$. Also by the Product Formula, both global heights and both absolute heights are insensitive to scalar multiplications.

To define absolute heights on an arbitrary proper subspace $\mathcal{Z} \subseteq k^N$, we begin by write $\mathcal{Z}$ as $\mathcal{Z} = \{\boldsymbol{x} \in k^N \mid A\boldsymbol{x} = \mathbf{0}\}$, where A is a $M \times N$ matrix for some $M < N$. We also let $X = (\boldsymbol{x}_1\ \boldsymbol{x}_2\ \cdots\ \boldsymbol{x}_L)$ be an $N \times L$ basis matrix of $\mathcal{Z}$, where $L = N - M$. Then the height of $\mathcal{Z}$ and X is given by

$$H(\mathcal{Z}) = H(X) = H(\boldsymbol{x}_1 \wedge \cdots \wedge \boldsymbol{x}_L),$$

and by an duality principle established in [5], the height of A satisfies

$$H(A) = H(X) = H(\mathcal{Z}).$$

4. Generalized Siegel's Lemmas

The first height version of Siegel's Lemma is proposed by Bombieri and Vaaler in [1]. They stated that

Theorem 2. For an L-dimensional subspace $\mathcal{Z} \subseteq k^N$, there exists a basis $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_L$ of $\mathcal{Z}$ such that

$$\prod_{i=1}^L H(\boldsymbol{x}_i) \leq N^{L/2}\left(\left(\frac{2}{\pi}\right)^{r_2}|\mathcal{D}_k|\right)^{\frac{L}{2d}} H(\mathcal{Z}),$$

where $2r_2$ is the number of complex embeddings of k into $\mathbb{C}$.

A similar bound that does not depends on the discriminant $\mathcal{D}_k$ is also established by Vaaler in [7]:

Theorem 3. For an L-dimensional subspace $\mathcal{Z} \subseteq k^N$, there exists a basis $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_L$ of $\mathcal{Z}$ such that

$$\prod_{i=1}^L H(\boldsymbol{x}_i) \leq \gamma_k(L)^{L/2} H(\mathcal{Z}),$$

where $\gamma_k(L)$ is the generalized Hermite's constant.

Sparse vectors and integer sensing matrices are of great importance in information theory. A recent paper [2] has established a new absolute version of Siegel's Lemma on sparse vectors. They have proved

Theorem 4. Let $\mathcal{Z}$ be an L-dimensional subspace of k^N, Then there exists a basis of $(N-L+1)$-sparse vectors $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_L$ of $\mathcal{Z}$, satisfying $H(\boldsymbol{x}_i) \leq H(\mathcal{Z})$ for all $1 \leq i \leq L$.

5. Avoiding Hyperplanes

There are also lots of well-established papers (see [3], [4], and [8]) dedicated on proving another version of generalized Siegel's Lemma, which is the existence of vectors of small heights avoiding a finite collection of hyperplanes, decomposable polynomials, or algebraic varieties. Using proving techniques similar to [4], I have proved the following proposition:

Proposition. Let $\mathcal{Z}$ be an L-dimensional subspace of k^N, and let $V_1, \cdots, V_M$ also be subspaces of k^N with dimensions no greater than s. Write $a = [(N-L+s+1)/2]$. Then there exists a $(N-L+s+1)$-sparse vector $\boldsymbol{x} \in \mathcal{Z} \backslash (\cup V_i)$ such that

$$\mathcal{H}(\boldsymbol{x}) \leq \mathcal{C}_{k,N}(L,s) H(W)^d \left\{ \left(\sum_{i=1}^M \frac{1}{H(V_i)^d} \right)^{\frac{1}{(L-s)d}} + M^{\frac{1}{(L-s)d+1}} \right\},$$

where

$$\mathcal{C}_{k,N}(L,s) = 2^{L(d+3)} |\mathcal{D}_k|^{L/2} \left((Ld)^L \binom{(N-L+s+1)d}{ad}^{\frac{1}{2d}} \right)^{\frac{1}{L-s}}.$$

It should be noted that this proposition is not of its best form: one should expect getting an analogous bound that depends on $\gamma_k(L)$ but not $|\mathcal{D}_k|$.

References

[1] E. Bombieri and J. Vaaler. Addendum to: "On Siegel's lemma". *Invent. Math.*, 75(2):377, 1984.

[2] Maxwell Forst and Lenny Fukshansky. On a new absolute version of siegel's lemma. *Research in the Mathematical Sciences*, 11(1):10, 2024.

[3] Lenny Fukshansky. Integral points of small height outside of a hypersurface. *Monatsh. Math.*, 147(1):25–41, 2006.

[4] Lenny Fukshansky. Siegel's lemma with additional conditions. *J. Number Theory*, 120(1):13–25, 2006.

[5] P. Gordan. über einige Anwendungen diophantischer Approximationen. *Math, Annalen*, 7:443–448, 1873.

[6] Carl L. Siegel. über einige Anwendungen diophantischer Approximationen. In *On some applications of Diophantine approximations*, volume 2 of *Quad./Monogr.*, pages 81–138. Ed. Norm., Pisa, 2014.

[7] Jeffrey D. Vaaler. The best constant in Siegel's lemma. *Monatsh. Math.*, 140(1):71–89, 2003.

[8] Shouwu Zhang. Positive line bundles on arithmetic surfaces. *Ann. of Math. (2)*, 136(3):569–587, 1992.