Lecture III

Oliver Daisey

These notes are entirely based on [Sie13].

1. Evaluating a volume integral

Recall the result from Lecture II that

$$f_r(x) = \left(\sum_{j=1}^{n} |x_j|^r \right)^{\frac{1}{r}} \tag{1}$$

is an even gauge function on $\mathbb{R}^n$ for all $r \geq 1$. Hence f_r corresponds to a convex body $\mathcal{B}_r$ by $\mathcal{B}_r = \{x \in \mathbb{R}^n \mid f_r(x) < 1\}$. Let V_r denote the volume of $\mathcal{B}_r$. Let Γ denote the gamma function $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$.

Theorem 1. *We have*

$$V_r = \frac{2^n \Gamma\left(\frac{1}{r} + 1\right)^n}{\Gamma\left(\frac{n}{r} + 1\right)}. \tag{2}$$

In particular, the case $r = 1$ recovers $V_1 = \frac{2^n}{n!}$ as was used in Lecture II.

Proof. By definition

$$V_r = \int_{\mathcal{B}_r} dx = \underset{\sum_{j=1}^n |x_j|^r < 1}{\int \cdots \int} dx_1 \dots dx_n. \tag{3}$$

We now split the integral over different regions. Defining

$$W_{n,r} = \underset{\substack{\sum_{j=1}^n x_j^r < 1 \\ x_j \geq 0,\ j=1,..,n}}{\int \cdots \int} dx_1 \dots dx_n, \tag{4}$$

we have that

$$V_r = 2^n \cdot W_{n,r} \tag{5}$$

since $\mathcal{B}_r$ has 0 as centre. Next, we observe that for any $\lambda > 0$,

$$\underset{\substack{\sum_{j=1}^n x_j^r < \lambda \\ x_j \geq 0,\ j=1,..,n}}{\int \cdots \int} dx_1 \dots dx_n = \underset{\substack{\sum_{j=1}^n \left(x_j \lambda^{-1/r}\right)^{1/r} < 1 \\ x_j \geq 0,\ j=1,..,n}}{\int \cdots \int} dx_1 \dots dx_n \tag{6}$$

and, by making the substitution $x_j \mapsto x_j \lambda^{-1/r}$,

$$\int_{\substack{\sum_{j=1}^n \left(x_j \lambda^{-1/r}\right)^{1/r} < 1 \\ x_j \geq 0,\ j=1,..,n}} \cdots \int dx_1 \dots dx_n = \lambda^{n/r} \cdot W_{n,r}. \tag{7}$$

We may write

$$W_{n,r} = \int_{\substack{\sum_{j=1}^n x_j^r < 1 \\ x_j \geq 0,\ j=1,..,n}} \cdots \int dx_1 \dots dx_n = \int_0^1 \left(\int_{\substack{\sum_{j=1}^n x_j^r < 1 - x_n^r \\ x_j \geq 0,\ j=1,..,n-1}} \cdots \int dx_1 \dots dx_{n-1} \right) dx_n, \tag{8}$$

and applying (6)-(7) with $\lambda = 1 - x_n^r$,

$$W_{n,r} = \int_0^1 W_{n-1,r} \cdot (1 - x_n^r)^{(n-1)/r} dx_n. \tag{9}$$

and so we need only evaluate a $1-$dimensional integral. Since $W_{n-1,r}$ is constant with respect to x_n, we may factor it out of the integral, so what's left is to evaluate

$$I_{n,r} = \int_0^1 (1 - x_n^r)^{(n-1)/r} dx_n. \tag{10}$$

Let $x_n = t^{1/r}$. Then $dx_n = \frac{1}{r} \cdot t^{(1-r)/r} dt$. Substituting,

$$I_{n,r} = \frac{1}{r} \int_0^1 (1-t)^{(n-1)/r} \cdot t^{(1-r)/r} dt. \tag{11}$$

We now recognise (11) as an instance of the *beta function* $B(x,y) = \int_0^1 t^{x-1}(1-t)^{(y-1)} dt$ with $x = \frac{1}{r}$, $y = (n-1)/r + 1$. The beta function and the gamma function are related by [Art15]

$$B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}. \tag{12}$$

Hence

$$I_{n,r} = \frac{1}{r} \cdot \frac{\Gamma(\frac{n-1}{r}+1)\Gamma(\frac{1}{r})}{\Gamma(\frac{n}{r}+1)} = \frac{\Gamma(\frac{n-1}{r}+1)\frac{1}{r}\Gamma(\frac{1}{r})}{\Gamma(\frac{n}{r}+1)} = \frac{\Gamma(\frac{n-1}{r}+1)\Gamma(\frac{1}{r}+1)}{\Gamma(\frac{n}{r}+1)}, \tag{13}$$

where we used the relation $\frac{1}{r}\Gamma(\frac{1}{r}) = \Gamma(\frac{1}{r}+1)$. Now by substituting into (9) we get that

$$W_{n,r} = W_{n-1,r} \cdot \frac{\Gamma(\frac{n-1}{r}+1)\Gamma(\frac{1}{r}+1)}{\Gamma(\frac{n}{r}+1)}. \tag{14}$$

By iterating this formula repeatedly for $W_{n-1,r}, W_{n-2,r}$ so on, and noting that $W_{1,r} = 1$, we obtain

$$W_{n,r} = W_{1,r} \cdot \frac{\left(\Gamma(\frac{1}{r}+1)\right)^n}{\Gamma(\frac{n}{r}+1)} = \frac{\left(\Gamma(\frac{1}{r}+1)\right)^n}{\Gamma(\frac{n}{r}+1)}. \tag{15}$$

Hence by (5) we have

$$V_r = 2^n \cdot \frac{\Gamma(\frac{1}{r}+1)^n}{\Gamma(\frac{n}{r}+1)} \tag{16}$$

as was to be shown. □

2. Discriminant of an irreducible polynomial.

Recall that a polynomial $P(\xi) = \xi^n + a_1\xi^{n-1} + ... + a_n$ with $a_1, ..., a_n \in \mathbb{Q}$ is said to be *irreducible* (over $\mathbb{Q}$) if it cannot be written as a product of two polynomials of strictly smaller degrees with coefficients in $\mathbb{Q}$. Let $\xi_1, ..., \xi_n$ denote the zeros of P in $\mathbb{C}$. We define the *discriminant* of P by the formula

$$\Delta = \prod_{1\le j\le k\le n} (\xi_j - \xi_k)^2 = \det \begin{pmatrix} \xi_1^{n-1} & \xi_1^{n-2} & ... & 1 \\ \xi_2^{n-1} & \xi_2^{n-2} & ... & 1 \\ \vdots & \vdots & & \vdots \\ \xi_n^{n-1} & \xi_n^{n-2} & ... & 1 \end{pmatrix}^2 . \tag{17}$$

Lemma 1. *If P is irreducible, then no zero of P can be a zero of any polynomial of strictly smaller degree, not identically zero, with rational coefficients.*

Proof. See [BD22], Chapter 14. □

Lemma 2. *Let $Q(x_1, ..., x_n)$ be a polynomial with integer coefficients which is symmetric in $x_1, ..., x_n$. Then $Q(\xi_1, ..., \xi_n)$ may be expressed as a polynomial in $a_1, ..., a_n$ with integer coefficients. If $a_1, ..., a_n$ are integers, then $Q(\xi_1, ..., \xi_n)$ is an integer.*

Proof. See [BD22], Chapter 14. □

By the results of Lecture II and the lemmas above we may prove

Theorem 2. *Let $P(\xi) = \xi^n + a_1\xi^{n-1} + ... + a_n$ be an irreducible polynomial with integer coefficients $a_1, ..., a_n$. If all the zeros of P are real, and Δ denotes the discrimiannt of P, we have*

$$\Delta \ge \left(\frac{n^n}{n!}\right)^2 . \tag{18}$$

Proof. Let $x_1, ..., x_n$ be arbitrary integers not all equal to zero, $\xi_1, ..., \xi_n$ the n distinct zeros of P. Define, for $j = 1, ..., n$,

$$y_j = \sum_{k=1}^{n} \xi_j^{n-k} x_k. \tag{19}$$

Note that for any $1 \le j \le n$, $y_j \ne 0$, since if $y_j = 0$ then ξ_j is a zero of a polynomial which is not identically zero, has integer coefficients, and has degree strictly less than n, which contradicts the irreducibility of P (Lemma 1.) So the product $y_1 y_2 \dots y_n$ is not zero, and it is an integer, because it is a symmetric polynomial with integer coefficients in the zeros of P (Lemma 2.) Because it is a non-zero integer,

$$|y_1 y_2 \dots y_n| \geq 1. \tag{20}$$

Now write $y = (y_1, ..., y_n)$ and introduce the gauge function

$$f(y) = \frac{1}{n} \sum_{j=1}^{n} |y_j|. \tag{21}$$

Define $\mu = \min\{f(y) \mid y \text{ is a } g\text{-point}, y \neq (0, ..., 0)\}$. Now Theorem 13 from Lecture II states

$$V\mu^n \leq 2^n D \tag{22}$$

where $D = \sqrt{\Delta}$ is the absolute value of the determinant of the transform $(x_1, ..., x_n) \mapsto (y_1, ..., y_n)$, and V is the volume of the convex body $\mathcal{B}$ defined by the gauge function: $\mathcal{B} = \{y \in \mathbb{R}^n \mid f(y) < 1\}$. Now V is the volume of the $n-$dimensional unit octahedron, scaled by factor n. Hence by using

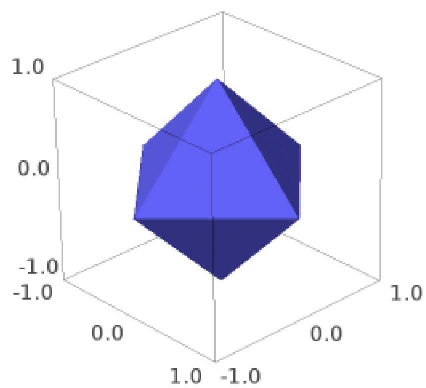

Figure 1: The convex body $\mathcal{B}$ in the case $n = 3$.

the special case $r = 1$ in Theorem 1, we obtain

$$V = \frac{(2n)^n}{n!} \tag{23}$$

and hence

$$D = \sqrt{\Delta} \geq \frac{(\mu n)^n}{n!}. \tag{24}$$

From the inequality of arithmetic and geometric means, we have

$$\frac{1}{n} \sum_{j=1}^{n} |y_j| \geq |y_1 y_2 ... y_n|^{1/n} \geq 1 \tag{25}$$

for all $y_1, ..., y_n$. Hence $\mu = \min f(y) = \min \frac{1}{n} \sum_{j=1}^{n} |y_j| \geq 1$. Then by combining this result with (24), we obtain

$$\sqrt{\Delta} \geq \frac{n^n}{n!} \tag{26}$$

as required. $\square$

Example 1. *We examine the theorem in the case $n = 2$. Let a, b be integers and let $P(\xi) = \xi^2 + a\xi + b$ be the irreducible polynomial. Then the discriminant $\Delta = a^2 - 4b$. The theorem claims that $\Delta \geq 4$. Indeed we have $\Delta > 0$ since the zeros of P are non-repeated real roots. Since a and b are integers, we must verify that $\Delta \neq 1, 2$ or 3. If $\Delta = 1$, then the polynomial is reducible, since the roots are $-\frac{a}{2} \pm \frac{1}{2}$. Since $\Delta = a^2 - 4b \equiv a^2 \pmod 4$, and the square of any integer is congruent to either 0 or 1 modulo 4, we deduce that $\Delta \neq 2$ and $\Delta \neq 3$. Hence $\Delta \geq 4$, in agreement with the theorem. In fact $\Delta \neq 4$ since in that case the polynomial is reducible; the roots are $-\frac{a}{2} \pm 1$.*

We note here that the lower bound given by the theorem is not exact. By taking $a = 1, b = -1$, in which case P is irreducible, we see that $\Delta = 5$ and this is the tightest lower bound.

In the proof of Theorem 2, we introduced the gauge function (21). We now show that the bound $\Delta \geq \left(\frac{n^n}{n!}\right)^2$ cannot be improved by choosing a gauge function of the form

$$f_r(y) = \left(\frac{1}{n}\sum_{j=1}^{n} |y_j|^r\right)^{1/r} \tag{27}$$

for $r \geq 1$. Let $V(r)$ denote the volume of $\mathcal{B}(r) = \{y \in \mathbb{R}^n \mid f_r(y) < 1\}$. Let $0 < s < r$. Recall *Hölder's inequality*

$$\sum_{j=1}^{n} a_j^p \cdot b_j^{1-p} \leq \left(\sum_{j=1}^{n} a_j\right)^p \cdot \left(\sum_{j=1}^{n} b_j\right)^{1-p} \tag{28}$$

for $a_j, b_j \geq 0$ and $0 < p < 1$, as discussed in Lecture II, (13). Use the values $p = \frac{s}{r}$, $a_j = \frac{1}{n}|y_j|^r$, $b_j = \frac{1}{n}$ for $j = 1, ..., n$. Then (28) reads

$$\frac{1}{n}\sum_{j=1}^{n} |y_j|^s \leq \left(\frac{1}{n}\sum_{j=1}^{n} |y_j|^r\right)^{s/r}. \tag{29}$$

Hence we have

$$f_s(y) = \left(\frac{1}{n}\sum_{j=1}^{n} |y_j|^s\right)^{1/s} \leq \left(\frac{1}{n}\sum_{j=1}^{n} |y_j|^r\right)^{1/r} \tag{30}$$

hence any $y \in \mathcal{B}(r)$ must also belong to $\mathcal{B}(s)$. Hence, by taking $s = 1$, we must have $V(r) \leq V(1)$ for all $r \geq 1$. So the bound cannot be improved by selecting $r > 1$.

3. Successive minima

Let f be an even gauge function on $\mathbb{R}^n$. Let $\mathcal{B}$ be the convex body $\mathcal{B} = \{x \in \mathbb{R}^n \mid f(x) < 1\}$. We define the *successive minima* of $\mathcal{B}$ as the set of real numbers μ_i with $1 \leq i \leq n$ such that $\mu_k = \inf\{\lambda \in \mathbb{R} \mid \lambda\mathcal{B}$ contains k linearly independent g-points.$\}$.

Equivalently, we may define the successive minima as follows. We define μ_1 to be the minimum of $f(g)$ over all $g-$points which are not the origin. Let $x^{(1)}$ be a vector such that $f(x^{(1)}) = \mu_1$. Then we define μ_2 to be the minimum of $f(g)$ over all $g-$points outside of the span of $x^{(1)}$. Let $x^{(2)}$ be a vector outside of the span of $x^{(1)}$, such that $f(x^{(2)}) = \mu_2$. Then we define μ_3 to be the minimum of $f(g)$ over all $g-$points outside of the span of $x^{(1)}$ and $x^{(2)}$, and so on until we have defined μ_n.

Theorem 3. *The above definitions are equivalent.*

Proof. Let ν_1 be as μ_1 in the first definition and μ_1 be as in the second definition. We will show $\mu_1 = \nu_1$; the proof for the rest of the μ_i is similar. Suppose for a contradiction that $\nu_1 < \mu_1$. Let $x^{(1)} \neq 0$ be a $g-$point on the surface of $\nu_1\mathcal{B}$, so $f(x^{(1)}) = \nu_1$. This contradicts the definition of μ_1 as the minimum of $f(g)$ for all $g-$points $g \neq 0$. On the other hand, suppose for a contradiction that $\mu_1 < \nu_1$. Then there exists a $g-$point $x^{(1)} \neq 0$ such that $f(x^{(1)}) = \mu_1 < \nu_1$, so that $x^{(1)}$ is a $g-$point in $\nu_1\mathcal{B}$, contradicting the definition of ν_1 as the value of λ such that there are no $g-$points except the origin inside $\lambda\mathcal{B}$. □

4. Minkowski's second theorem

Recall the following from Lecture II:

Theorem 4. *(Minkowski's first theorem.) Let f be an even gauge function on $\mathbb{R}^n$, V the volume of the convex body $\mathcal{B} = \{x \in \mathbb{R}^n \mid f(x) < 1\}$. Let μ_1 be the minimum of $f(x)$ as x runs through all the $g-$points different from the origin. Then we have*

$$V\mu_1^n \leq 2^n. \tag{31}$$

We may use the successive minima as defined above to generalise as follows:

Theorem 5. *(Minkowski's second theorem.) If $\mu_1, ..., \mu_n$ denote the successive minima of an even gauge function f on $\mathbb{R}^n$, then we have*

$$V\mu_1\mu_2 \ldots \mu_n \leq 2^n \tag{32}$$

where V denotes the volume of the convex body $\mathcal{B} = \{x \in \mathbb{R}^n \mid f(x) < 1\}$.

References

[Art15] Emil Artin. *The gamma function.* Courier Dover Publications, 2015.

[BD22] Maxime Bôcher and Edmund Pendleton Randolph Duval. *Introduction to higher algebra.* Macmillan, 1922.

[Sie13] Carl Ludwig Siegel. *Lectures on the Geometry of Numbers.* Springer Science & Business Media, 2013.