CPSC 664, Spring 2023 Scribe: Tim Kunisky
January 19, 2023 Lecturer: Tim Kunisky

Lecture 2: Integer partitioning II

1 First moment method by integration

Recall that we left off planning to compute the expectation of the random variable

$$Z := \#\left\{ \boldsymbol{x} \in \{\pm 1\}^n : \langle \boldsymbol{x}, \boldsymbol{a} \rangle = 0 \right\}. \tag{1}$$

Normally what we would do, and what we will do later for other problems, is to expand Z into a sum of indicators like in the last lecture and compute from there. But, in the special case of integer partitioning, it turns out that it is possible to get much more precise information using the following beautiful trick.

Define the function

$$F(t) := \prod_{j=1}^{n} (e^{ia_j t} + e^{-ia_j t}). \tag{2}$$

This is a random function, inheriting its randomness from the a_j. Consider two manipulations of this function: on the one hand, by Euler's formula $e^{ix} = \cos x + i \sin x$,

$$F(t) = 2^n \prod_{j=1}^{n} \cos(a_j t), \tag{3}$$

a real-valued function. On the other hand, expanding the product,

$$F(t) = \sum_{\boldsymbol{x} \in \{\pm 1\}^n} \exp\left(it \sum_{j=1}^{n} x_j a_j \right) = \sum_{d \in \mathbb{Z}} e^{idt} \#\left\{ \boldsymbol{x} \in \{\pm 1\}^n : \langle \boldsymbol{x}, \boldsymbol{a} \rangle = d \right\}. \tag{4}$$

Thus, $F(t)$ is a kind of generating function of the numbers of $\boldsymbol{x}$ achieving different values of the "discrepancy" or "objective function" $\langle \boldsymbol{x}, \boldsymbol{a} \rangle$. We can extract these counts by using the *Fourier transform*, which amounts to the following observation.

Proposition 1.1 (Orthogonality of Fourier modes)**.** *For any $d \in \mathbb{Z}$,*

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{idt} \, dt = \left\{ \begin{array}{ll} 1 & \text{if } d = 0 \\ 0 & \text{if } d \neq 0 \end{array} \right\}. \tag{5}$$

Proof. Use Euler's formula and calculate the integrals of sine and cosine directly. □

In particular, we may extract Z from $F(t)$ by computing

$$Z = \frac{1}{2\pi}\int_{-\pi}^{\pi} F(t)\,dt = \frac{2^n}{2\pi}\int_{-\pi}^{\pi}\prod_{j=1}^{n}\cos(a_j t)\,dt. \tag{6}$$

Taking the expectation, we may exchange the integral and expectation, and then use that the a_j are i.i.d., to observe that

$$\mathbb{E}Z = \frac{2^n}{2\pi}\int_{-\pi}^{\pi}\prod_{j=1}^{n}\mathbb{E}\cos(a_j t)\,dt = \frac{2^n}{2\pi}\int_{-\pi}^{\pi}(\mathbb{E}\cos(a_1 t))^n\,dt. \tag{7}$$

Let us define

$$\begin{aligned} f(t) &:= \mathbb{E}\cos(a_1 t) \\ &= \frac{1}{B}\sum_{a=0}^{B-1}\cos(at) \\ &= \frac{1}{2B}\sum_{a=0}^{B-1}(e^{iat} + e^{-iat}) \end{aligned}$$

and summing this as two geometric series, you may check that we arrive at

$$= \frac{1}{2B}\left(1 + \frac{\sin((B-\frac{1}{2})t)}{\sin(\frac{1}{2}t)}\right). \tag{8}$$

You may check that, extending by continuity, $f(0) = 1$, and this is the maximum of $f(t)$ over $t \in [-\pi, \pi]$; see Figure 1.

2 Laplace method

The *Laplace method* is a useful general analysis tool to understand integrals like this. The idea is that, for large n, the dominant contribution to the integral will be from a small neighborhood around $t = 0$. There, we can approximate f by a Taylor series:

$$f(0) = 1, \tag{9}$$
$$f'(0) = 0, \tag{10}$$

and you may check with a slightly more involved calculation that

$$f''(0) = -\frac{B^2}{3} + O(B). \tag{11}$$

Thus, near $t = 0$, we have

$$f(t) \approx f(0) + f'(0)t + \frac{f''(0)}{2}t^2 \approx 1 - \frac{B^2}{6}t^2. \tag{12}$$

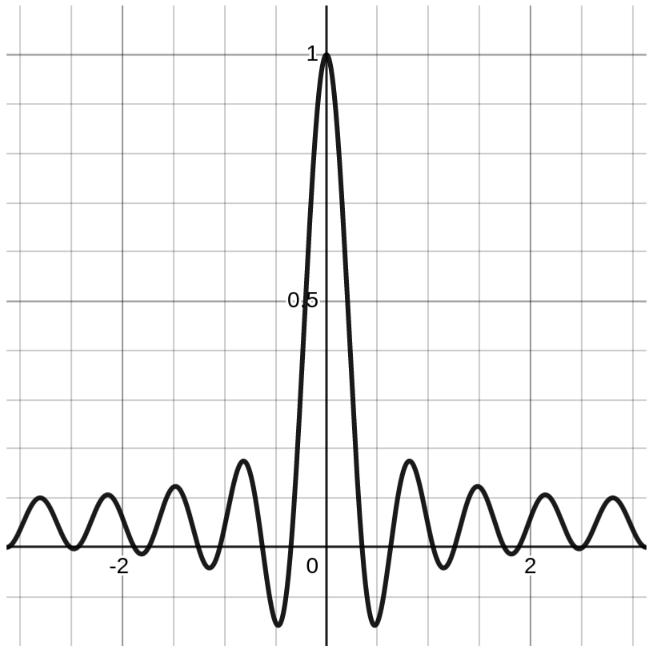

Figure 1: A plot of $f(t)$ with $B = 10$.

Now, starting with our intuition that we may restrict to a small neighborhood around $t = 0$, we approximate

$$\begin{aligned}
\int_{-\pi}^{\pi} f(t)^n \, dt &\approx \int_{-\epsilon}^{\epsilon} f(t)^n \, dt \\
&\approx \int_{-\epsilon}^{\epsilon} \left(1 - \frac{B^2}{6} t^2\right)^n \, dt \\
&\approx \int_{-\epsilon}^{\epsilon} \mathsf{exp}\left(-\frac{B^2 n}{6} t^2\right) \, dt
\end{aligned}$$

Let us view this as a Gaussian integral with "effective variance" $\sigma^2 := 3/B^2 n \ll 1$. Then, we may rewrite and, since this variance is very small, expand the range of integration without affecting the value very much,

$$\begin{aligned}
&= \int_{-\epsilon}^{\epsilon} \mathsf{exp}\left(-\frac{t^2}{2\sigma^2}\right) \, dt \\
&\approx \int_{-\infty}^{\infty} \mathsf{exp}\left(-\frac{t^2}{2\sigma^2}\right) \, dt
\end{aligned}$$

which is the normalizing constant in the Gaussian distribution,

$$\begin{aligned}
&= \sqrt{2\pi\sigma^2} \\
&= \frac{\sqrt{6\pi}}{B\sqrt{n}}. \qquad (13)
\end{aligned}$$

Substituting into the expression we wanted to calculate,

$$\mathbb{E}Z = \frac{2^n}{2\pi}\int_{-\pi}^{\pi} f(t)^n\,dt \approx \sqrt{\frac{3}{2\pi}}\cdot\frac{2^n}{B\sqrt{n}}. \tag{14}$$

If we do these calculations more carefully keeping track of error terms, we arrive at the following.

Theorem 2.1. *For any B, n,*

$$\mathbb{E}Z = (1 + O(1/B) + O(1/n))\cdot\sqrt{\frac{3}{2\pi}}\cdot\frac{2^n}{B\sqrt{n}}, \tag{15}$$

with the $O(\cdot)$ hiding absolute constants.

See Appendix A.6 of [MM11] for many more details about how to perform such calculations carefully.

Consequently, we can verify one side of our prediction from last lecture.

Corollary 2.2. *Suppose that $B = B(n)$ is such that $B/(2^n/\sqrt{n}) = B/B^* \to \infty$. Then, with high probability, there exist no perfect partitions of $a_1, \ldots, a_n$.*

3 Minimum discrepancy in the unsatisfiable regime

Essentially the same calculation can also clarify some more details of how our problem behaves in the situation described by Corollary 2.2. In that case, it is interesting to determine the scaling of the minimum achievable discrepancy, the value of the optimization problem

$$\begin{array}{ll} \text{minimize} & \left|\sum_{j=1}^n x_j a_j\right| \\ \text{subject to} & \boldsymbol{x} \in \{\pm 1\}^n. \end{array} \tag{16}$$

To that end, we calculate expectations of

$$Z_d := \#\left\{\boldsymbol{x} \in \{\pm 1\}^n : \langle \boldsymbol{x}, \boldsymbol{a}\rangle = d\right\}. \tag{17}$$

By the same argument as for $Z = Z_0$,

$$Z_d = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-idt}F(t)\,dt = \frac{2^n}{2\pi}\int_{-\pi}^{\pi}\mathsf{cos}(dt)\prod_{j=1}^n \mathsf{cos}(a_j t)\,dt, \tag{18}$$

where we may take the real part since we know Z_d is real, and taking expectations we find

$$\mathbb{E}Z_d = \frac{2^n}{2\pi}\int_{-\pi}^{\pi}\mathsf{cos}(dt) f(t)^n dt. \tag{19}$$

Near $t = 0$, we have the Taylor approximation $\mathsf{cos}(dt) \approx 1 - \frac{d^2}{2}t^2$. Since $f(t) \approx 1 - \frac{B^2}{6}t^2$, so long as $d \ll B$, we will have in the Laplace method calculation that $\mathsf{cos}(dt)$ is effectively equal to 1 in the interval $[-\epsilon, \epsilon]$ on which $f(t)^n$ is not negligibly small. Making this reasoning precise, one may confirm that

Theorem 3.1. *For any B, n, d,*

$$\mathbb{E}Z_d = (1 + O(d/B) + O(1/n)) \cdot \sqrt{\frac{3}{2\pi}} \cdot \frac{2^n}{B\sqrt{n}}, \tag{20}$$

with the $O(\cdot)$ hiding absolute constants.

What does this tell us about the minimum discrepancy? We have, for $D \ll B$,

$$\mathbb{E}[\#\{\boldsymbol{x} \in \{\pm 1\}^n : |\langle \boldsymbol{x}, \boldsymbol{a} \rangle| \leq D\}] = \mathbb{E} \sum_{d=-D}^{D} Z_d \approx 2D \cdot \sqrt{\frac{3}{2\pi}} \cdot \frac{2^n}{B\sqrt{n}}. \tag{21}$$

On the other hand, whenever $|\langle \boldsymbol{x}, \boldsymbol{a} \rangle| = d$ then $|\langle -\boldsymbol{x}, \boldsymbol{a} \rangle| = d$ as well, so we might intuit that the D for which there typically exists a $\boldsymbol{x}$ with $|\langle \boldsymbol{x}, \boldsymbol{a} \rangle| = D$ makes the above quantity equal to 2.[1] This gives

$$D = \sqrt{\frac{2\pi}{3}} \frac{B\sqrt{n}}{2^n}. \tag{22}$$

Thus we are led to the following conjecture.

Conjecture 3.2. *Suppose $B = B(n)$ is such that $B/(2^n/\sqrt{n}) = B/B^* \to \infty$. Then,*

$$\left\{ \begin{array}{ll} \textit{minimize} & \left|\sum_{j=1}^n x_j a_j\right| \\ \textit{subject to} & \boldsymbol{x} \in \{\pm 1\}^n \end{array} \right\} = (1 + o(1)) \sqrt{\frac{2\pi}{3}} \frac{B\sqrt{n}}{2^n}. \tag{23}$$

Of course, it would also be reasonable to more conservatively conjecture the above without a specific constant, but we will see in the next lecture that this constant is correct and has an intuitive interpretation in terms of the "random energy model."

4 Second moment method

Let us now turn to trying to show the converse statement to Corollary 2.2, that if $B/B^* \to 0$ then with high probability there exist perfect partitions of $a_1, \dots, a_n$.

Actually, this attempt would be doomed to fail from the outset: what we wrote above simply is not true. That is because we always have $\sum_{j=1}^n x_j a_j \equiv \sum_{j=1}^n a_j \pmod 2$, and so a perfect partition can only exist if $\sum_{j=1}^n a_j$ is even, which will only happen with probability about $\frac{1}{2}$. We should revise our notion of "perfect partition" in light of this observation: in the notation of the previous section, what is reasonable to expect is that $Z_0 + Z_1 \geq 1$ with high probability (i.e., there exists a $\boldsymbol{x}$ with $|\langle \boldsymbol{x}, \boldsymbol{a} \rangle| \in \{0, 1\}$), not $Z = Z_0 \geq 1$ with high probability.

However, let us proceed and see where our argument would break down even if we had not made this observation "from outside" of our moment calculations. In either case, the main tool for proving this kind of result is the following.

[1]This is admittedly a flimsy heuristic argument, but surprisingly gives the right result!

Proposition 4.1 (Basic second moment method)**.** *Suppose $Z \geq 0$ is a real-valued random variable. Then,*

$$\mathbb{P}[Z > 0] \geq \frac{(\mathbb{E}Z)^2}{\mathbb{E}Z^2} = 1 - \frac{\mathsf{Var}[Z]}{\mathbb{E}Z^2}. \tag{24}$$

Proof. Using the Cauchy-Schwarz inequality,

$$\mathbb{E}Z = \mathbb{E}Z\mathbb{1}\{Z > 0\} \leq \sqrt{(\mathbb{E}Z^2)(\mathbb{E}\mathbb{1}\{Z > 0\}^2)} = \sqrt{(\mathbb{E}Z^2)(\mathbb{P}[Z > 0])}, \tag{25}$$

and rearranging gives the result. □

To use this, our main task is to compute $\mathbb{E}Z^2$. Let us sketch how to do this; for further details you can see Section 14.5 of [MM11]. Squaring our integral representation of Z and rearranging,

$$Z^2 = \left(\frac{2^n}{2\pi}\right)^2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \prod_{j=1}^{n} \mathsf{cos}(a_j s)\, \mathsf{cos}(a_j t)\, ds\, dt. \tag{26}$$

Again by independence of the a_j, we have

$$\mathbb{E}Z^2 = \left(\frac{2^n}{2\pi}\right)^2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} (\underbrace{\mathbb{E}\, \mathsf{cos}(a_1 s)\, \mathsf{cos}(a_1 t)}_{g(s,t)})^n\, ds\, dt. \tag{27}$$

Though it is a little more cumbersome, you can still write a closed formula for $g(s,t)$ and verify that its maximum is at $(s,t) = (0,0)$, where $g(0,0) = 1$. We may then use a two-dimensional version of the Laplace method: we have $\nabla g(0,0) = \mathbf{0}$ since the origin is a maximum, and so the two dimensional Taylor expansion says that, near $(0,0)$,

$$g(s,t) \approx 1 + \begin{bmatrix} s \\ t \end{bmatrix}^\top \nabla^2 g(0,0) \begin{bmatrix} s \\ t \end{bmatrix}, \tag{28}$$

where $\nabla^2 g(0,0)$ is the Hessian matrix, which is negative-definite (again since the origin is a local maximum). We may then go through the previous derivation again, now working with a two-dimensional Gaussian integral; we omit the details, but ultimately you will find

$$\mathbb{E}Z^2 = 2(1 + o(1))(\mathbb{E}Z)^2. \tag{29}$$

(Here, importantly, the $o(1)$ term only satisfies that asymptotic so long as $\mathbb{E}Z \sim B^*/B \to \infty$; the actual bound we can obtain on it when doing the Laplace method carefully is $O(1/\mathbb{E}Z)$.)

It is here that we see the reflection of the parity issue mentioned above in our moment arguments: with this kind of asymptotic, the best that Proposition 4.1 can tell us is

$$\mathbb{P}[Z > 0] \geq \frac{1}{2} + o(1), \tag{30}$$

which is exactly what our parity reasoning told us to expect! To repair the situation, one may do the same calculation with Z_1 as with $Z = Z_0$, and note that only one of Z_0 and Z_1 can be positive. Combining these results, we would find the following.

Theorem 4.2. *Suppose that $B = B(n)$ is such that $B/(2^n/\sqrt{n}) = B/B^* \to 0$. Then, with high probability, there exists $\boldsymbol{x} \in \{\pm 1\}^n$ such that $\sum_{j=1}^{n} x_j a_j \in \{0, 1\}$.*

References

[MM11] Cristopher Moore and Stephan Mertens. *The nature of computation*. OUP Oxford, 2011.