A Self-contained Proof of the Gaussian Isoperimetric Inequality

BY LU XIAOYANG

Abstract

We give a mostly self-contained proof of the Gaussian isoperimetric inequality and discuss its relationship with the Kannan-Lovász-Simonovits conjecture.

1 Notations

- Vectors are in **bold Roman** font (e.g., $\mathbf{x}$).
- $\mathbf{X}_n \Rightarrow \mathbf{X}$ means the distribution of $\mathbf{X}_n$ weakly converges to the distribution of $\mathbf{X}$.
- $\mathbf{1}_i = (0, \ldots, 1, \ldots, 0)$ where only the ith coordinate is nonzero.
- $\mathbb{1}_A$ is the indicator function of set A.
- $|\mathbf{x}|$ is the L_2 norm of vector $\mathbf{x}$.
- $\liminf x_i$ (resp. $\limsup x_i$) is the smallest (resp. largest) subsequential limit of x_i.
- $B(\mathbf{x}, d)$ is the ball with center $\mathbf{x}$ and radius d.
- $f \rightarrow g$ means pointwise convergence.

2 Background

The *Gaussian isoperimetric inequality* refers to the following theorem.

Theorem 1. *Let $h > 0$. Then, among all sets $A \subseteq \mathbb{R}^n$ with fixed Gaussian measure $\gamma_n(A)$, the half spaces minimize the Gaussian measure of the neighbourhood $\gamma_n(A^h)$. Here γ_n denotes the n-dimensional Gaussian measure, and A^h is defined as*

$$A^h \;=\; \{\mathbf{x} \in \mathbb{R}^n \mid \exists \mathbf{a} \in A, |\mathbf{x} - \mathbf{a}| < h\}.$$

The theorem can be stated equivalently as follows.

Proposition 2. *For all Borel measurable sets $A \subseteq \mathbb{R}^n$ and $h > 0$, $\gamma_n(A^h) \geqslant \Phi(\Phi^{-1}(\gamma_n(A)) + h)$. Here*

$$\Phi(x) = \int_{-\infty}^{x} \varphi(t)\,\mathrm{d}t, \varphi(t) = \frac{1}{\sqrt{2\,\pi}}\, e^{-\frac{x^2}{2}}.$$

Proof. We have

$$\int_{[-\infty, t] \times \mathbb{R}^{n-1}} 1\,\mathrm{d}\gamma_n \;=\; \int_{-\infty}^{t} \varphi(x)\,\mathrm{d}x = \Phi(t).$$

Fix a measure $\gamma_n(A)$. Consider a half space H with measure $\gamma_n(A)$, say, $[-\infty, \Phi^{-1}(\gamma_n(A))] \times \mathbb{R}^{n-1}$ (according to the rotation invariance of φ_n, all half spaces have the same measure). H^h is also a half space, and for all H they are isomorphic under rotation. Using $[-\infty, \Phi^{-1}(\gamma_n(A))] \times \mathbb{R}^{n-1}$ as a representative, we can compute $\gamma_n(H^h) = \Phi(\Phi^{-1}(\gamma_n(A)) + h)$. $\square$

Sometimes we use a weaker—but in some sense still equivalent—form of the theorem.

Corollary 3. *For all Borel measurable sets $A \subseteq \mathbb{R}^n$, $\gamma_n^+(A) \geqslant I(\gamma_n(A))$. Here $\gamma_n^+(A)$ is defined as*

$$\gamma_n^+(A) \;=\; \liminf_{h \rightarrow 0^+} \frac{\gamma_n(A^h) - \gamma_n(A)}{h}.$$

I is defined as

$$I(x) = \varphi(\Phi^{-1}(x)), I(0) = I(1) = 0.$$

Figure 1 shows the graph of I.

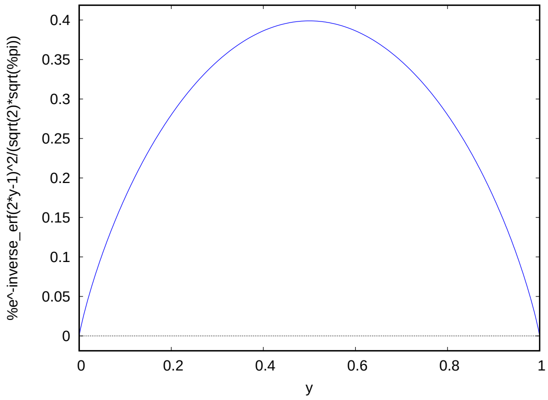

Figure 1. Graph of I, generated by Maxima.

Proof. The forward implication is clear (let $h \rightarrow 0$). We prove the converse implication. Fix a set A and $h > 0$. For all $\varepsilon > 0$ and $h \geqslant h_1 > 0$, there exists some $h_0 > 0$, such that for all $0 < h' \leqslant h_0$,

$$\frac{\gamma_n((A^{h_1})^{h'}) - \gamma_n(A^{h_1})}{h'} \;\geqslant\; I(\gamma_n(A^{h_1})) - \varepsilon. \tag{1}$$

By Taylor's Theorem (see, for example, [JCL18]) we have

$$\begin{aligned} \Phi(x+h) &= \Phi(x) + \Phi'(x)\, h + \frac{1}{2}\, \Phi''(\xi)\, h^2 \\ &= \Phi(x) + \varphi(x)\, h + \frac{1}{2}\, \varphi'(\xi)\, h^2, \end{aligned} \tag{2}$$

where $x \leqslant \xi \leqslant x + h$. Combining (1) and (2),

$$\begin{aligned} \gamma_n((A^{h_1})^{h'}) &\geqslant \varphi(\Phi^{-1}(\gamma_n(A^{h_1})))\, h' - \varepsilon\, h + \gamma_n(A^{h_1}) \\ &= \Phi\left(\Phi^{-1}(\gamma_n(A^{h_1})) + \left(1 - \frac{2\,\varepsilon}{I(\gamma_n(A^{h_1}))}\right) h'\right) - \frac{1}{2}\, \varphi'(\xi) \left(\left(1 - \frac{2\,\varepsilon}{I(\gamma_n(A^{h_1}))}\right) h'\right)^2 + \varepsilon\, h' \\ &\geqslant \Phi\left(\Phi^{-1}(\gamma_n(A^{h_1})) + \left(1 - \frac{\varepsilon}{I(\gamma_n(A^{h_1}))}\right) h'\right) \\ &\geqslant \Phi\left(\Phi^{-1}(\gamma_n(A^{h_1})) + \left(1 - \frac{\varepsilon}{\inf_{0<t\leqslant h} I(\gamma_n(A^t))}\right) h'\right) \end{aligned}$$

for all small enough h'. The last inequality holds because $|\varphi'(x)| = \left|\frac{x}{\sqrt{2\,\pi}}\, e^{-\frac{x^2}{2}}\right|$ has a bound $|\varphi'(1)|$. Let $\frac{\varepsilon}{\inf_{0<t\leqslant h} I(\gamma_n(A^t))} = \delta$. We observe that

$$\begin{aligned} \gamma_n(A^{h_1+h_2}) &= \gamma_n((A^{h_1})^{h_2}) \\ &\geqslant \Phi(\Phi^{-1}(\gamma_n(A^{h_1})) + (1-\delta)\, h_2) \\ &\geqslant \Phi(\Phi^{-1}(\Phi(\Phi^{-1}(\gamma_n(A)) + (1-\delta)\, h_1)) + (1-\delta)\, h) \\ &= \Phi(\Phi^{-1}(\gamma_n(A)) + (1-\delta)\, (h_1 + h_2)). \end{aligned}$$

for small h_1 and h_2. So we can choose a small enough h' and use the above additive property to derive

$$\gamma_n(A^h) \;\geqslant\; \Phi(\Phi^{-1}(\gamma_n(A)) + (1-\delta)\, h).$$

Now let $\varepsilon \rightarrow 0$ and take limits on both sides. $\square$

Remark 4. The previous argument implicitly assumes that h_0 can be chosen independently of h_1. This will be clear in Section 6. **[Remark end.]**

In the following sections, we first prove a discrete version on cubes in Section 3. We generalize an important inequality to smooth functions in Section 4. We further generalize it to Lipschitz functions in Section 5. Finally, in Section 6, we prove Corollary 3—and consequently, Proposition 2 and Theorem 1—by approximating indicator functions with Lipschitz functions. In Section 7 we discuss the Kannan-Lovász-Simonovits (KLS) conjecture.

The proof in this article is essentially a detailed adaption of [Bob97].

3 The discrete case

Consider the following property of function $I\colon [0,1]\to\mathbb{R}^+$.

$$I\left(\frac{a+b}{2}\right) \leqslant \frac{1}{2}\sqrt{I(a)^2+\left|\frac{a-b}{2}\right|^2}+\frac{1}{2}\sqrt{I(b)^2+\left|\frac{a-b}{2}\right|^2}. \tag{3}$$

Proposition 5. *I defined in Corollary 3 satisfies (3).*

Proof. The proof is a bit technical. Fix $a,b\in(0,1)$. Define $g(x)=I^2(c+x)+x^2, c=\frac{a+b}{2}$. (3) can be rewritten using g as

$$\begin{aligned}\sqrt{g(0)} &\leqslant \frac{1}{2}\sqrt{g\left(\frac{a-b}{2}\right)}+\sqrt{g\left(-\frac{a-b}{2}\right)}\\ &:= \frac{1}{2}\sqrt{g(t)}+\sqrt{g(-t)}.\end{aligned}$$

Or equivalently,

$$16\,g^2(0)+(g(t)-g(-t))^2 \leqslant 8\,g(0)\,(g(t)+g(-t)).$$

Introduce $h(x)=g(x)-g(0),$ we need to prove

$$(h(t)-h(-t))^2 \leqslant 8\,I(c)\,(h(t)+h(-t)).$$

Lemma 6. $I(x)\,I''(x)=-1$.

Proof. $I'(x)=\varphi'(\Phi^{-1}(x))\,\frac{1}{\varphi(\Phi^{-1}(x))}=-\Phi^{-1}(x), I''(x)=-\frac{1}{\varphi(\Phi^{-1}(x))}=-\frac{1}{I(x)}$. □

Lemma 7. *${I'}^2(x)$ is convex.*

Proof. $({I'}^2(x))'=2\,I'(x)\,I''(x)=-2\,\frac{I'(x)}{I(x)}, ({I'}^2(x))''=-2\,\frac{I''(x)\,I(x)-{I'}^2(x)}{I^2(x)}=2\,\frac{1+{I'}^2(x)}{I^2(x)}\geqslant 0.$ □

Define $R(x)=h(x)+h(-x)-2\,{I'}^2(c)\,x^2$. We claim that R is convex.

$$R'(x)=2\,I(c+x)\,I'(c+x)-2\,I(c-x)\,I'(c-x)+4\,x-4\,{I'}^2(c)\,x,$$

by Lemma 7

$$R''(x)=4\left(\frac{{I'}^2(c+x)+{I'}^2(c-x)}{2}-{I'}^2(c)\right)\geqslant 0.$$

It follows that $R(x)\geqslant R(0)=0$ because $R'(0)=0$. Now we only need to prove

$$(h(t)-h(-t))^2 \leqslant 8\,I(c)\,(2\,{I'}^2(c)\,t^2),$$

or

$$\left|\frac{h(t)-h(-t)}{t}\right|=\left|\frac{I^2(c+t)-I^2(c-t)}{t}\right|\leqslant 4\,I(c)\,|I'(c)|.$$

Because I is symmetry around $\frac{1}{2}$, without loss of generality we assume $0<c<\frac{1}{2}$. We further assume $c>t>0$. It remains to show

$$\frac{I^2(c+t)-I^2(c-t)}{t} \leqslant 4\,I(c)\,I'(c).$$

Consider $u(x)=I^2(c+x)-I^2(c-x), 0<x<c$. We have $u''(x)=2\,(I'^2(c+x)-I'^2(c-x))\leqslant 0$. Then,

$$\frac{u(x)}{x} = \int_0^1 u'(x\,t)\,\mathrm{d}t$$

is nonincreasing on $(0,c]$. So we only need to prove $\lim_{t\to 0}\frac{I^2(c+t)-I^2(c-t)}{t}\leqslant 4\,I(c)\,I'(c)$. But by Taylor's theorem,

$$I^2(c+x) = I'(c)+2\,I(c)\,I'(c)\,x+O(x^2).$$

□

Remark 8. I is the maximal continuous function satisfying (3) and $I(0)=I(1)=0$. However, the rest of the proof does not use this property; rather, it implies the property. **[Remark end.]**

We can rewrite inequality (3) as a "two point" analytic inequality. That is, for all function f: $\{-1,1\}\to[0,1]$,

$$I(\mathbb{E}\,f) \leqslant \mathbb{E}\sqrt{I^2(f)+|\nabla f|^2}. \tag{4}$$

Here expectation is with respect to the measure $p(-1)=p(1)=\frac{1}{2}$, and $\nabla f=\frac{f(1)-f(-1)}{2}$. We can generalize (4) to functions on higher dimensional cubes $f\colon\{-1,1\}^n\to[0,1]$. In that case, we define

$$|\nabla f(\mathbf{x})|^2 = \frac{1}{4}\sum_{i=1}^n |f(\mathbf{x})-f(\mathbf{s}_i(\mathbf{x}))|^2,$$

where $(\mathbf{s}_i(\mathbf{x}))_j=\begin{cases}-x_j, & i=j\\ x_j & i\neq j\end{cases}$ denotes the neighborhood of $\mathbf{x}$ in the ith coordinate.

Lemma 9. *For all function $I\colon\{-1,1\}\to\mathbb{R}^+$, if (4) holds for all $f\colon\{-1,1\}\to[0,1]$ on a measure μ, then (4) also holds for all $f\colon\{-1,1\}^n\to[0,1]$ on the product measure μ_n.*

Proof. Induction on n. Given function $f\colon\{-1,1\}^{n+1}\to[0,1]$, define $f_{-1}(\mathbf{x})=f(\mathbf{x},-1)$ and $f_1(\mathbf{x})=f(\mathbf{x},1)$. We have

$$\begin{aligned}\mathbb{E}\,f &= \sum_{\mathbf{x}\in\{-1,1\}^{n+1}} \mu_{n+1}(\mathbf{x})\,f(\mathbf{x})\\ &= \sum_{\mathbf{x}\in\{-1,1\}^n} \mu_n(\mathbf{x})\,\mu(-1)\,f(\mathbf{x},-1)+\sum_{\mathbf{x}\in\{-1,1\}^n}\mu_n(\mathbf{x})\,\mu(1)\,f(\mathbf{x},1)\\ &= \mu(-1)\,\mathbb{E}\,f_{-1}+\mu(1)\,\mathbb{E}\,f_1\end{aligned}$$

and

$$\begin{aligned}|\nabla f_i(\mathbf{x})|^2 &= \frac{1}{4}\sum_{i=1}^n |f_i(\mathbf{x})-f_1(\mathbf{s}_i(\mathbf{x}))|^2\\ &= \frac{1}{4}\sum_{i=1}^n |f(\mathbf{x},i)-f(\mathbf{s}_i(\mathbf{x}),i)|^2\\ &= |\nabla f(\mathbf{x},i)|^2-\frac{1}{4}|f_i(\mathbf{x})-f_{-i}(\mathbf{x})|^2, i\in\{-1,1\}.\end{aligned}$$

Combining the previous two equations,

$$\mathbb{E}\sqrt{I^2(f)+|\nabla f|^2} = \sum_{i=-1,1}\mu(i)\,\mathbb{E}\sqrt{I^2(f_i)+|\nabla f_i|^2+\frac{1}{4}|f_i-f_{-i}|^2}. \tag{5}$$

By the Cauchy-Schwarz inequality,

$$\begin{aligned}\left(\sum_i f_i\right)^2+\left(\sum_i g_i\right)^2 &= \sum_i f_i^2+\sum_i g_i^2+2\sum_{i\neq j}(f_i\,f_j+g_i\,g_j)\\ &\leqslant \sum_i f_i^2+\sum_i g_i^2+2\sum_{i\neq j}\sqrt{(f_i^2+f_j^2)\,(g_i^2+g_j^2)}\\ &= \left(\sum_i \sqrt{f_i^2+g_i^2}\right)^2.\end{aligned}$$

Applying the previous inequality to Equation (5) yields

$$\begin{aligned}\mathbb{E}\sqrt{I^2(f)+|\nabla f|^2} &\geqslant \sum_{i=-1,1}\mu(i)\sqrt{\left(\mathbb{E}\sqrt{I^2(f_i)+|\nabla f_i|^2}\right)^2+\left(\mathbb{E}\,\frac{f_i-f_{-i}}{2}\right)^2}\\ &= \sum_{i=-1,1}\mu(i)\sqrt{\left(\mathbb{E}\sqrt{I^2(f_i)+|\nabla f_i|^2}\right)^2+\frac{1}{4}(\mathbb{E}\,f_{-1}-\mathbb{E}\,f_1)^2}.\end{aligned}$$

By the induction hypothesis, $\mathbb{E}\sqrt{I^2(f_i)+|\nabla f_i|^2}\geqslant I(\mathbb{E}\,f_i)$. Therefore,

$$\begin{aligned}\mathbb{E}\sqrt{I^2(f)+|\nabla f|^2} &\geqslant \sum_{i=-1,1}\mu(i)\sqrt{I^2(\mathbb{E}\,f_i)+\frac{1}{4}(\mathbb{E}\,f_{-1}-\mathbb{E}\,f_1)^2}\\ &= \mathbb{E}\sqrt{I^2(\delta)+|\nabla\delta|^2}\\ &\geqslant I(\mathbb{E}\,\delta)=I(\mathbb{E}\,f),\end{aligned}$$

where $\delta(i)=\mathbb{E}\,f_i$. The last inequality uses (4) with $n=1$. □

Combining Lemma 9 and Proposition 5, we get the following.

Proposition 10. *Let $I(x)=\varphi(\Phi^{-1}(x))$. Then, for all $f\colon\{-1,1\}^n\to[0,1]$,*

$$I(\mathbb{E}\,f) \leqslant \mathbb{E}\sqrt{I^2(f)+|\nabla f|^2}.$$

Here expectation is with respect to the uniform measure.

Remark 11. Proposition 10 implies the following inequality.

Corollary 12. *For all set $A\subseteq\{-1,1\}^n$, $\mu_n^+(A):=\int|\nabla\mathbb{1}_A|\,\mathrm{d}\mu_n\geqslant I(\mu_n(A))$, where μ_n is the uniform measure.*

[Remark end.]

4 The smooth case

We extend Proposition 10 to smooth functions. Consider a bounded twice differentiable function $f\colon\mathbb{R}^n\to[0,1]$ with bounded first and second partial derivatives. We define a sequence of functions on independent uniformly sampled random vectors $\mathbf{x}_1,\ldots,\mathbf{x}_k$

$$f_k\colon\{-1,1\}^{nk}\to[0,1],\, f_k(\mathbf{x}_1,\ldots,\mathbf{x}_k) = f\left(\frac{\mathbf{x}_1+\cdots+\mathbf{x}_k}{\sqrt{k}}\right), \qquad k\in\mathbb{N}-\{0\}.$$

By the (multidimensional) central limit theorem, $\frac{\mathbf{x}_1+\cdots+\mathbf{x}_k}{\sqrt{k}}\Rightarrow\mathbf{x}\sim\varphi_n$ when $k\to\infty$. By the continuous mapping theorem (see, for example, page 261 of [Res13]), $f_k(\mathbf{x}_1,\ldots,\mathbf{x}_k)\Rightarrow f(\mathbf{x})$ and

$$\int f_k\,\mathrm{d}\mu_n \to \int f\,\mathrm{d}\gamma_n.$$

By Taylor's theorem,

$$\begin{aligned}|f_k(\mathbf{x}_1,\ldots,\mathbf{x}_k)-f_k(\mathbf{s}_{mk+i}(\mathbf{x}_1,\ldots,\mathbf{x}_k))| &= \left|f\left(\frac{\mathbf{x}_1+\cdots+\mathbf{x}_k}{\sqrt{k}}\right)-f\left(\mathbf{s}_i\left(\frac{\mathbf{x}_1+\cdots+\mathbf{x}_k}{\sqrt{k}}\right)\right)\right|\\ &= \left|\frac{\partial f\left(\frac{\mathbf{x}_1+\cdots+\mathbf{x}_k}{\sqrt{k}}\right)}{\partial x_i}\right|\frac{2}{\sqrt{k}}+\frac{1}{2}\frac{4}{k}\left|\frac{\partial^2 f\left(\frac{\mathbf{x}_1+\cdots+\mathbf{x}_k}{\sqrt{k}}+\xi\,\mathbf{1}_i\right)}{\partial x_i^2}\right|\end{aligned}$$

where $-\frac{1}{\sqrt{k}} \leqslant \xi \leqslant \frac{1}{\sqrt{k}}$. Since first and second partial derivatives are bounded, for all $\delta > 0$, when k is large enough, there is a universal (with respect to i) constant ε such that

$$|f_k(\mathbf{x}_1, \ldots, \mathbf{x}_k) - f_k(\mathbf{s}_{mk+i}(\mathbf{x}_1, \ldots, \mathbf{x}_k))|^2 \leqslant \left|\frac{\partial f\left(\frac{\mathbf{x}_1 + \cdots + \mathbf{x}_k}{\sqrt{k}}\right)}{\partial x_i}\right|^2 \frac{4}{k} + \varepsilon \frac{1}{k^{\frac{3}{2}}}.$$

Now

$$\begin{aligned} |\nabla f_k(\mathbf{x}_1, \ldots, \mathbf{x}_k)|^2 &= \frac{1}{4} \sum_{j=1}^{nk} |f_k(\mathbf{x}_1, \ldots, \mathbf{x}_k) - f_k(\mathbf{s}_j(\mathbf{x}_1, \ldots, \mathbf{x}_k))|^2 \\ &\leqslant \sum_{i=1}^{n} \left|\frac{\partial f\left(\frac{\mathbf{x}_1 + \cdots + \mathbf{x}_k}{\sqrt{k}}\right)}{\partial x_i}\right|^2 + \varepsilon' \frac{1}{\sqrt{k}} \\ &:= \mathrm{D} f\left(\frac{\mathbf{x}_1 + \cdots + \mathbf{x}_k}{\sqrt{k}}\right) + O\left(\frac{1}{\sqrt{k}}\right). \end{aligned}$$

Again by the central limit theorem and the uniform mapping theorem, $\mathrm{D} f\left(\frac{\mathbf{x}_1 + \cdots + \mathbf{x}_k}{\sqrt{k}}\right) \Rightarrow \mathrm{D} f(\mathbf{x})$ and $|\nabla f_k(\mathbf{x}_1, \ldots, \mathbf{x}_k)|^2 \Rightarrow \mathrm{D} f(\mathbf{x})$ where $\mathbf{x} \sim \varphi_n$ as $k \to \infty$. Using analogous reasoning, we conclude

$$\int \sqrt{I^2(f_k) + |\nabla f_k|^2}\, \mathrm{d}\mu_n \to \int \sqrt{I^2(f) + |\mathrm{D} f|^2}\, \mathrm{d}\gamma_n.$$

Taking limits on both sides in Proposition 10, we obtain

Proposition 13. *For all bounded twice differentiable function $f: \mathbb{R}^n \to [0, 1]$ with bounded first and second partial derivatives,*

$$I(\mathbb{E} f) \leqslant \mathbb{E} \sqrt{I^2(f) + |\mathrm{D} f|^2}.$$

Here expectation is taken over standard Gaussian measure.

Remark 14. Inspecting the proof of Proposition 13, we can weaken the conditions on f to having bounded first partial derivatives and bounded second partial derivatives of the form $\frac{\partial^2 f}{\partial x_i^2}$.

[Remark end.]

5 The Lipschitz case

We extend Proposition 13 to Lipschitz functions. We define the *mollifier function* $\phi: \mathbb{R}^n \to \mathbb{R}$ as

$$\phi(\mathbf{x}) = \begin{cases} C e^{-\frac{1}{1 - |\mathbf{x}|^2}}, & -1 < |\mathbf{x}| < 1 \\ 0, & \text{otherwise} \end{cases}.$$

The constant C is chosen such that $\int \phi(\mathbf{x})\, \mathrm{d}\mathbf{x} = 1$. Figure 2 and 3 show the graphs of one-dimensional and two-dimensional ϕ, respectively. They are indeed *smooth*.

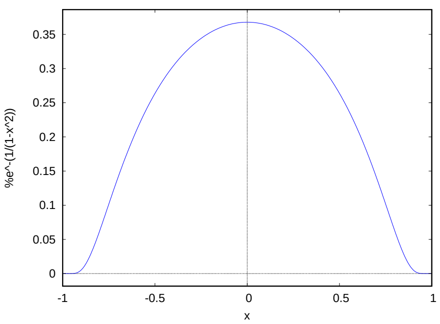

Figure 2. Graph of one-dimensional ϕ, up to a constant multiple.

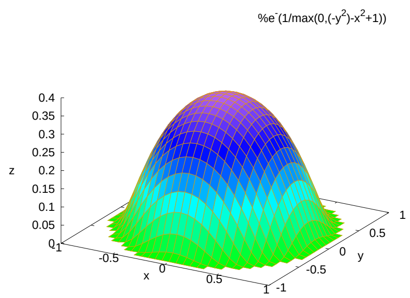

Figure 3. Graph of two-dimensional ϕ, up to a constant multiple.

As stated in [Sob38] or analysis textbooks, $\phi \in C^\infty$. In fact, partial derivatives of ϕ are bounded. For our purpose, we only show that

$$\begin{aligned}\left|\frac{\partial\phi(x_1,\ldots,x_n)}{\partial x_i}\right| &= \left|C\,\frac{\partial e^{-\frac{1}{1-x_1^2-\cdots-x_n^2}}}{\partial x_i}\right|\\ &= \left|C\,e^{-\frac{1}{1-|\mathbf{x}|^2}}\,\frac{2\,x_i}{(1-|\mathbf{x}|)^2}\right|\\ &\leqslant C\,\frac{4}{e^2}\end{aligned}$$

because $x^2\,e^{-x}\leqslant\frac{4}{e^2}$, and

$$\begin{aligned}\left|\frac{\partial^2\phi(x_1,\ldots,x_n)}{\partial x_i^2}\right| &= \left|C\,\frac{\partial e^{-\frac{1}{1-|\mathbf{x}|}}\,\frac{2\,x_i}{(1-|\mathbf{x}|)^2}}{\partial x_i}\right|\\ &= C\,e^{-\frac{1}{1-|\mathbf{x}|^2}}\left(\frac{2}{(1-|\mathbf{x}|^2)^2}+\frac{8\,x_i^2}{(1-|\mathbf{x}|^2)^3}+\frac{4\,x_i^2}{(1-|\mathbf{x}|^2)^4}\right)\\ &\leqslant C\left(\frac{8}{e^2}+\frac{216}{e^3}+\frac{1024}{e^4}\right)\end{aligned}$$

because $x^3\,e^{-x}\leqslant\frac{27}{e^3}$, $x^4\,e^{-x}\leqslant\frac{256}{e^4}$.

We define $\phi_\varepsilon(\mathbf{x})=\frac{1}{\varepsilon^n}\,\phi\left(\frac{\mathbf{x}}{\varepsilon}\right)$, $\varepsilon>0$. The integral of ϕ_ε is still 1. Fixing a bounded Lipschitz function f, we consider the smoothed function

$$f_\varepsilon := f*\phi_\varepsilon=\int f(\mathbf{y})\,\phi_\varepsilon(\mathbf{x}-\mathbf{y})\,\mathrm{d}\mathbf{y}=\int\limits_{B(\mathbf{0},1)} f(\mathbf{x}-\varepsilon\,\mathbf{y})\,\phi(\mathbf{y})\,\mathrm{d}\mathbf{y},\varepsilon>0.$$

f_ε also has arbitrary order partial derivatives, but for our purpose, we only need show f_ε has bounded first partial derivatives and bounded second partial derivatives of the form $\frac{\partial^2 f_\varepsilon}{\partial x_i^2}$ according to Remark 14. We have

$$\begin{aligned}\frac{f_\varepsilon(\mathbf{x})-f_\varepsilon(\mathbf{z})}{\varepsilon} &= \frac{1}{\varepsilon}\left(\int\limits_{B(\mathbf{0},1)}(f(\mathbf{x}-\varepsilon\,\mathbf{y})-f(\mathbf{z}-\varepsilon\,\mathbf{y}))\,\phi(\mathbf{y})\,\mathrm{d}\mathbf{y}\right),\\ |f(\mathbf{x}-\varepsilon\,\mathbf{y})-f(\mathbf{z}-\varepsilon\,\mathbf{y})|\,\phi(\mathbf{y}) &\leqslant |\mathbf{x}-\mathbf{y}|\,\phi(\mathbf{y}),\end{aligned}$$

and

$$\frac{1}{\varepsilon}\int\limits_{B(\mathbf{0},1)}|\mathbf{x}-\mathbf{y}|\,\phi(\mathbf{y})\,\mathrm{d}\mathbf{y} = |\mathbf{x}-\mathbf{y}|.$$

By the dominated convergence theorem,

$$\begin{aligned}\frac{\partial f_\varepsilon}{\partial x_i}(\mathbf{x}) &= \lim_{\delta\to 0}\frac{f_\varepsilon(\mathbf{x})-f_\varepsilon(\mathbf{x}+\delta\,\mathbf{1}_i)}{\delta}\\ &= \lim_{\delta\to 0}\int\frac{\phi_\varepsilon(\mathbf{x}-\mathbf{y})\,\mathrm{d}\mathbf{y}-\phi_\varepsilon(\mathbf{x}+\delta\,\mathbf{1}_i-\mathbf{y})\,\mathrm{d}\mathbf{y}}{\delta}\,f(\mathbf{y})\,\mathrm{d}\mathbf{y}\\ &= \int\lim_{\delta\to 0}\frac{\phi_\varepsilon(\mathbf{x}-\mathbf{y})\,\mathrm{d}\mathbf{y}-\phi_\varepsilon(\mathbf{x}+\delta\,\mathbf{1}_i-\mathbf{y})\,\mathrm{d}\mathbf{y}}{\delta}\,f(\mathbf{y})\,\mathrm{d}\mathbf{y}\\ &= \int\limits_{B(\mathbf{y},\varepsilon)}\frac{\partial\phi_\varepsilon(\mathbf{x}-\mathbf{y})}{\partial x_i}\,f(\mathbf{y})\,\mathrm{d}\mathbf{y}\end{aligned}$$

which is bounded because $\frac{\partial\phi_\varepsilon}{\partial x_i}$ and f are bounded. Using analogous argument (move $\mathbf{y}$ into f to use DCT and move $\mathbf{y}$ into $\frac{\partial\phi_\varepsilon}{\partial x_i}$ to do the derivation), we can show $\frac{\partial^2 f_\varepsilon}{\partial x_i^2}$ exists and is bounded. Let $\varepsilon\to 0$. Then $f_\varepsilon\to f$ because

$$|f_\varepsilon(\mathbf{x})-f(\mathbf{x})| \leqslant \int_{B(\mathbf{0},1)}|f(\mathbf{x}-\varepsilon\,\mathbf{y})-f(\mathbf{x})|\,\phi(\mathbf{y})\,\mathrm{d}\mathbf{y}\leqslant\varepsilon. \tag{6}$$

By Rademacher's theorem, f is differentiable almost everywhere. We can also show $\frac{\partial f_\varepsilon}{\partial x_i} \to \frac{\partial f}{\partial x_i}$ when the RHS exists analogously to Equation (6). Taking limits on both sides of Proposition 13 and use the dominated convergence theorem, we conclude

Proposition 15. *For all bounded Lipschitz* $f: \mathbb{R}^n \to [0,1]$,

$$I(\mathbb{E} f) \leqslant \mathbb{E}\sqrt{I^2(f) + |\mathrm{D} f|^2}.$$

Here expectation is taken over standard Gaussian measure, and $\mathrm{D} f(\mathbf{x}) = 0$ *if* f *is not differentiable at* $\mathbf{x}$ *(which is negligible).*

6 The differential case

We prove Corollary 3. Consider a Borel measurable set A. We approximate $\mathbb{1}_A$ using a family of functions

$$f_h(\mathbf{x}) = \max\left\{1 - \frac{1}{h} \inf_{\mathbf{a} \in A} |\mathbf{x} - \mathbf{a}|, 0\right\}, \qquad h > 0.$$

We first show f_h is $\frac{1}{h}$-Lipschitz.

Proof. For all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$\begin{aligned} |f_h(\mathbf{x}) - f_h(\mathbf{y})| &\leqslant \frac{1}{h}\left|\inf_{\mathbf{a} \in A} |\mathbf{x} - \mathbf{a}| - \inf_{\mathbf{a} \in A} |\mathbf{y} - \mathbf{a}|\right| \\ &:= \frac{1}{h}|d(\mathbf{x}, A) - d(\mathbf{y}, A)|. \end{aligned}$$

For all $\varepsilon > 0$, there exist $\mathbf{a} \in A$, $d(\mathbf{x}, A) \leqslant |\mathbf{x} - \mathbf{a}| \leqslant d(\mathbf{x}, A) + \varepsilon$. By triangular inequality,

$$d(\mathbf{y}, A) \leqslant |\mathbf{y} - \mathbf{a}| \leqslant |\mathbf{x} - \mathbf{a}| + |\mathbf{y} - \mathbf{x}|.$$

So

$$d(\mathbf{y}, A) - d(\mathbf{x}, A) \leqslant |\mathbf{y} - \mathbf{x}| + \varepsilon.$$

Similarly $d(\mathbf{x}, A) - d(\mathbf{y}, A) \leqslant |\mathbf{y} - \mathbf{x}| + \varepsilon$. To conclude, $|d(\mathbf{x}, A) - d(\mathbf{y}, A)| \leqslant |\mathbf{y} - \mathbf{x}| + \varepsilon$. Now let $\varepsilon \to 0$. $\square$

A consequence of being $\frac{1}{h}$-Lipschitz is that $|\mathrm{D} f(\mathbf{x})| \leqslant \frac{1}{h}$ whenever it exists. By Proposition 15,

$$\begin{aligned} I(\mathbb{E} f_h) &\leqslant \mathbb{E}\sqrt{I^2(f_h) + |\mathrm{D} f_h|^2} \\ &\leqslant \mathbb{E} I(f_h) + \mathbb{E}|\mathrm{D} f_h| \\ &\leqslant \mathbb{E} I(f_h) + \int_{A^h - A} \frac{1}{h} \mathrm{d}\gamma_n + \int_A |\mathrm{D} f_h| \mathrm{d}\gamma_n \int_{\mathbb{R}^n - A^h} |\mathrm{D} f_h| \mathrm{d}\gamma_n. \end{aligned}$$

But for all $\mathbf{x} \in A$, $f_h(\mathbf{x})$ is the maximum, so $f'(\mathbf{x})$ can only be $\mathbf{0}$, if it exists. Similarly for $\mathbf{x} \in \mathbb{R}^n - A^h$. We conclude

$$I(\mathbb{E} f_h) \leqslant \mathbb{E} I(f_h) + \frac{\gamma_n(A^h) - \gamma_n(A)}{h}.$$

Now let $h \to 0$. Then, $f_h \downarrow \mathbb{1}_{\bar{A}}$. If $\gamma_n(\partial A) > 0$,

$$\gamma_n^+(A) = \liminf_{h \to 0} \frac{\gamma_n(A^h) - \gamma_n(A)}{h} \geqslant \liminf_{h \to 0} \frac{\gamma_n(\bar{A}) - \gamma_n(A)}{h} = \infty.$$

So we assume $\gamma_n(\partial A) = 0$. By the monotone convergence theorem,

$$\mathbb{E} f_h \downarrow \mathbb{E}\,\mathbb{1}_{\bar{A}} = \gamma_n(\bar{A}) = \gamma_n(A).$$

Because $I(f_h) \leqslant I(0)$, by the dominated convergence theorem, $\mathbb{E} I(f_h) \to \mathbb{E} I(\mathbb{1}_{\bar{A}}) = 0$. When h is small enough that $\mathbb{E} f_h \geqslant \gamma_n(A) \geqslant \frac{1}{2}$ or $\frac{1}{2} \geqslant \mathbb{E} f_h \geqslant \gamma_n(A)$, by the monotonicity of I we have

$$\begin{aligned}\min\{I(\mathbb{E} f_h), I(\gamma_n(A))\} &= \inf_{0<x\leqslant h} I(\mathbb{E} f_x) \\ &\leqslant \inf_{0<x\leqslant h} \left(\mathbb{E} I(f_x) + \frac{\gamma_n(A^x) - \gamma_n(A)}{x}\right).\end{aligned}$$

Taking limits on both sides,

$$\begin{aligned} I(\gamma(A)) &= \liminf_{h\to 0} \min\{I(\mathbb{E} f_h), I(\gamma_n(A))\} \\ &= \liminf_{h\to 0}\left(\mathbb{E} I(f_h) + \frac{\gamma_n(A^h) - \gamma_n(A)}{h}\right) \\ &\leqslant \liminf_{h\to 0} \frac{\gamma_n(A^h) - \gamma_n(A)}{h} + \limsup_{h\to 0} \mathbb{E} I(f_h) \\ &= \liminf_{h\to 0} \frac{\gamma_n(A^h) - \gamma_n(A)}{h}.\end{aligned}$$

7 Connection with the KLS conjecture

Definition 16. *We say a probability density function f is logconcave, if*

$$f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \;\geqslant\; f^{\lambda}(\mathbf{x})\, f^{1-\lambda}(\mathbf{y}).$$

We use the version of the KLS conjecture in [LV18].

Conjecture 17. *For all logconcave density f in $\mathbb{R}^n$ and its derived measure $\mu(A) = \int_A f(\mathbf{x})\,\mathrm{d}\mathbf{x}$,*

$$\psi_p := \inf_{S \text{ measurable}} \frac{\mu^+(S)}{\min\{\mu(S), \mu(\mathbb{R}^n - S)\}}. \geqslant c \inf_{H \text{ half space}} \frac{\mu^+(H)}{\min\{\mu(H), \mu(\mathbb{R}^n - H)\}}.$$

Here c is a universal *constant, and μ^+ is defined similarly to γ_n^+.*

Notice that the Gaussian is logconcave:

$$\log \varphi_n(\mathbf{x}) = \log\left(\frac{1}{\sqrt{(2\pi)^n}}\, e^{\frac{1}{2}|\lambda \mathbf{x} + (1-\lambda)\mathbf{y}|^2}\right) = \frac{1}{2}|\mathbf{x}|^2 - \frac{n}{2}\log(2\pi).$$

We can prove the special case $f = \varphi_n$.

$$\begin{aligned}\inf_{S \text{ measurable}} \frac{\mu^+(S)}{\min\{\mu(S), \mu(\mathbb{R}^n - S)\}} &= \inf_{\gamma_n(S)\leqslant\frac{1}{2}} \frac{\gamma_n^+(S)}{\gamma_n(S)} \\ &\geqslant \inf_{\gamma_n(S)\leqslant\frac{1}{2}} \frac{I(\gamma_n(S))}{\gamma_n(S)} \\ &= \frac{I\left(\frac{1}{2}\right)}{\frac{1}{2}} = \sqrt{\frac{2}{\pi}}.\end{aligned}$$

Equality is reached when S is a half space. The inequality holds because I is concave, as we proved in Section 3.

8 Conclusion

We gave a mostly self-contained proof of the Gaussian isoperimetric inequality and briefly discussed its connection with the KLS conjecture. Some notable techniques involved were (1) using the central limit theorem to step from discrete cases to continuous ones, (2) using mollifiers to approximate a function by functions with better properties, and (3) using Lipschitz functions to approximate (regular) indicator functions.

Bibliography

[Bob97] Sergey G. Bobkov. An isoperimetric inequality on the discrete cube, and an elementary proof of the isoperimetric inequality in gauss space. *The Annals of Probability*, 25:206–214, 01 1997.

[JCL18] Chen Jixiu, Yu Chonghua, and Jin Lu. *Mathematical Analysis*. Higher Education Press, 3 edition, 2018.

[LV18] Yin Tat Lee and Santosh S. Vempala. The kannan-lovász-simonovits conjecture. 2018.

[Res13] Sidney I. Resnick. *A Probability Path*. Birkhäuser Boston, MA, 1 edition, 2013.

[Sob38] S. Soboleff. Sur un théorème d'analyse fonctionnelle. *Rec. Math. [Mat. Sbornik] N.S.*, 4(46):471–497, 1938.