November 18, 2022

Lecture 4: Bochner Identities and Curvature

Minerva mini-course on Convexity in High dimensions by Bo'az Klartag

So far we discussed two techniques in convex geometry, Convex Localization and Optimal Transport, with applications including the Poincaré inequality with the optimal Payne-Weinberger constant, Gromov's waist inequality, reverse Hölder inequalities for polynomials and the reverse Cheeger inequality. Today we discuss a method that was borrowed from Riemannian geometry: The Bochner technique that goes back to Bochner in the 1940s and also Lichnerowicz in the 1950s. In a nutshell, the idea is to make local computations involving something like curvature, as well as integrations by parts, and then dualize and obtain Poincaré-type inequalities. This may sound pretty vague, let us explain what we mean.

Suppose that μ is a log-concave probability measure in $\mathbb{R}^n$ with density $e^{-\psi}$ for a smooth, convex function $\psi : \mathbb{R}^n \to \mathbb{R}$. Therefore $\nabla^2\psi \geq 0$ everywhere in $\mathbb{R}^n$. We will measure distances using the Euclidean distances in $\mathbb{R}^n$, but we will measure volumes using the measure μ. We thus look at the *weighted Riemannian manifold* or the *metric-measure space*

$$(\mathbb{R}^n, |\cdot|, \mu).$$

Thus for instance the Dirichlet energy of a smooth function $f : \mathbb{R}^n \to \mathbb{R}$ is

$$\|f\|^2_{\dot{H}^1(\mu)} = \int_{\mathbb{R}^n} |\nabla f|^2 d\mu.$$

Indeed, we measure the length of the gradient with respect to the Euclidean metric, while we integrate with respect to the measure μ. The Laplace-type operator associated with this measure-metric space is defined, initially for $u \in C_c^\infty(\mathbb{R}^n)$, via

$$Lu = \Delta u - \nabla\psi \cdot \nabla u = e^{\psi} div(e^{-\psi}\nabla u).$$

This reason for this definition is that for any smooth functions $u, v : \mathbb{R}^n \to \mathbb{R}$, with one of them compactly-supported,

$$\int_{\mathbb{R}^n} (Lu)v d\mu = -\int_{\mathbb{R}^n} [\nabla u \cdot \nabla v] e^{-\psi}.$$

We will not use much functional analysis or operator theory today, but it is worthwhile to mention that L is essentially self adjoint in $L^2(\mu)$. In Riemannian geometry, the Ricci curvature appears when we commute the Laplacian and the gradient. Analogously, here we have the easily-verified commutation relation

$$\nabla(Lu) = L(\nabla u) - (\nabla^2\psi)(\nabla u),$$

where $L(\nabla u) = (L(\partial^1 u), \ldots, L(\partial^n u))$. Hence the matrix $\nabla^2\psi$ is a curvature term, analogous to the Ricci curvature.

Proposition 1 (Integrated Bochner's formula)**.** *For any $u \in C_c^\infty(\mathbb{R}^n)$,*

$$\int_{\mathbb{R}^n} (Lu)^2 \, d\mu = \int_{\mathbb{R}^n} \left(\nabla^2 \psi\right) \nabla u \cdot \nabla u \, d\mu + \int_{\mathbb{R}^n} \|\nabla^2 u\|_{HS}^2 d\mu,$$

where $\|\nabla^2 u\|_{HS}^2 = \sum_{i=1}^n |\nabla \partial_i u|^2$.

Proof. Integration by parts gives

$$\begin{aligned}\int_{\mathbb{R}^n} (Lu)^2 \, d\mu = -\int_{\mathbb{R}^n} \nabla(Lu) \cdot \nabla u \, d\mu &= -\int_{\mathbb{R}^n} L(\nabla u) \cdot \nabla u \, d\mu + \int_{\mathbb{R}^n} \left[(\nabla^2 \psi) \nabla u \cdot \nabla u\right] \, d\mu \\ &= \sum_{i=1}^n \int_{\mathbb{R}^n} |\nabla \partial_i u|^2 \, d\mu + \int_{\mathbb{R}^n} \left(\nabla^2 \psi\right) \nabla u \cdot \nabla u \, d\mu.\end{aligned}$$

□

We would like to use the property that the first summand on the right-hand side is non-negative, since ψ is convex. For this we would need to present a given function $f \in L^2(\mu)$ as

$$f = Lu.$$

A necessary condition, assuming that $u \in C_c^\infty(\mathbb{R}^n)$, is that

$$\int_{\mathbb{R}^n} f d\mu = 0.$$

This follows from the integration by parts above with $v \equiv 1$. This necessary condition is more-or-less sufficient:

Lemma 2. *The image of $C_c^\infty(\mathbb{R}^n)$ under the operator L is dense in*

$$H = \left\{ f \in L^2(\mu) : \int f \, d\mu = 0 \right\} \subset L^2(\mu).$$

Proof. Suppose that $f \in L^2(\mu)$ with $\|f\|_{L^2(\mu)} = 1$ is such that for each $u \in C_c^\infty(\mathbb{R}^n)$ we have that $Lu \perp f$ in $L^2(\mu)$. We will show that f is constant. Suppose first that f is smooth. We claim that

$$Lf = 0.$$

Indeed, for any $u \in C_c^\infty(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} u(Lf) d\mu = \int_{\mathbb{R}^n} (Lu) f d\mu = 0.$$

In general, f is a weak solution of $Lf = 0$ in an appropriate sense, and this is an elliptic equation, and hence f is smooth and it is a classical solution in $\mathbb{R}^n$. Next we claim that for any $u \in C_c^\infty(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} |\nabla(uf)|^2 d\mu = \int_{\mathbb{R}^n} |\nabla u|^2 f^2 d\mu.$$

The proof is just a fun exercise in integration by parts, where we use $Lf \equiv 0$ as well as the formula $L(uv) = uLv + vLu + 2\langle u, \nabla v\rangle$. If u is a bump function which equals one in some large ball $B(0, 1/\varepsilon)$ and with $|\nabla u| < \varepsilon$ everywhere, then,

$$\int_{B(0,1/\varepsilon)} |\nabla f|^2 d\mu \leq \int_{\mathbb{R}^n} |\nabla(uf)|^2 d\mu = \int_{\mathbb{R}^n} |\nabla u|^2 f^2 d\mu \leq \varepsilon^2$$

By letting ε tend to zero we see that $\nabla f \equiv 0$ and f is constant. See Cordero-Erausquin, Fradelizi and Maurey '04 for more details. □

We will use two ways for dualizing the integrated Bochner formula. Either ignore the non-negative Hessian term and obtain the Brascamp-Lieb inequality from the 1970s, or else ignore the curvature term and obtain the H^{-1}-inequality to be discussed later.

Theorem 3 (Brascamp-Lieb). *Assume $\nabla^2\psi > 0$ throughout $\mathbb{R}^n$ (strong log-concavity). Then for any C^1-smooth $f \in L^2(\mu)$,*

$$\mathrm{Var}_\mu(f) \leq \int_{\mathbb{R}^n} \left(\nabla^2\psi\right)^{-1} \nabla f \cdot \nabla f \, d\mu(x),$$

where $\mathrm{Var}_\mu(f) = \int_{\mathbb{R}^n} (f - E)^2 \, d\mu(x)$, *and* $E = \int_{\mathbb{R}^n} f d\mu$.

Proof. Assume $\int f \, d\mu = 0, \varepsilon > 0$ and pick $u \in C_c^\infty(\mathbb{R}^n)$ such that

$$\|Lu - f\|_{L^2(\mu)} < \varepsilon.$$

Then,

$$\begin{aligned}
\mathrm{Var}_\mu(f) = \|f\|_{L^2(\mu)}^2 &= \|Lu - f\|_{L^2(\mu)}^2 + 2\int fLu \, d\mu - \int (Lu)^2 \, d\mu \\
&\leq \varepsilon^2 - 2\int \nabla f \cdot \nabla u \, d\mu - \int (\nabla^2\psi)\nabla u \cdot \nabla u \, d\mu \\
&\leq \varepsilon^2 + \int (\nabla^2\psi)^{-1} \nabla f \cdot \nabla f \, d\mu,
\end{aligned}$$

where we have used the fact that

$$\int (Lu)^2 \, d\mu \geq \int (\nabla^2\psi)\nabla u \cdot \nabla u \, d\mu,$$

which follows from Bochner's formula and

$$-2x \cdot y - Ax \cdot x \leq A^{-1}y \cdot y \Longleftrightarrow |\sqrt{A}x + \sqrt{A^{-1}}y|^2 \geq 0.$$

The desired inequality follows by letting ε tend to zero. □

Remark. The Brascamp-Lieb inequality is an infinitesimal version of the Prékopa-Leindler inequality. Suppose that $f_0, f_1 : \mathbb{R}^n \to [0, \infty)$ are integrable, log-concave functions and

$$f_t(x) = \sup_{x=(1-t)y+yz} f_0(y)^{1-t} f_1(z)^t.$$

The Prékopa-Leindler inequality implies that $\log \int_{\mathbb{R}^n} f_t$ is concave in t. The second derivative in t is non-negative, and this actually amounts to the Brascamp-Lieb inequality. Thus the Brascamp-Lieb inequality is yet another incarnation of the Brunn-Minkowski theory.

Corollary 4. *If μ is more log-concave than the standard Gaussian, then its Poincaré constant (spectral gap) satisfies*

$$\lambda_\mu \geq 1.$$

Proof. Write $d\mu/dx = e^{-\psi}$. Then $\nabla^2\psi \geq \mathrm{Id}$ and hence $(\nabla^2\psi)^{-1} \leq \mathrm{Id}$. Hence for any f, from the Brascamp-Lieb inequality,

$$Var_\mu(f) \leq \int_{\mathbb{R}^n} [(\nabla^2\psi)^{-1}\nabla f \cdot \nabla f] d\mu \leq \int_{\mathbb{R}^n} |\nabla f|^2 d\mu.$$

□

This is a reasonable bound under *uniform* log-concavity assumptions. The KLS conjecture suggests that uniformity is not needed, and that knowledge about the covariance matrix suffices.

Definition 5. *Consider the orthant $\mathbb{R}^n_+$. A function $\psi : \mathbb{R}^n \to \mathbb{R}$ is invariant under coordinate reflections (a.k.a unconditional) if*

$$\psi(x_1, \ldots, x_n) = \psi(|x_1|, \ldots, |x_n|) \qquad \textit{for all } x \in \mathbb{R}^n.$$

If ψ is moreover convex, then $\psi|_{\mathbb{R}^n_+}$ is increasing in all coordinate directions.

Exercise: if ψ is convex and increasing in all of the coordinate directions, then ψ is p-convex for $p = 1/2$, i.e., $\psi(x_1^2, \ldots, x_n^2)$ is convex in the orthant.

Corollary 6. *Let μ be a probability measure in $\mathbb{R}^n_+$, set $e^{-\psi} = d\mu/dx$ and assume that ψ is p-convex for $p = 1/2$ (e.g. μ is log-concave and unconditional). Then for any C^1-smooth function $f \in L^2(\mu)$,*

$$\mathrm{Var}_\mu(f) \leq 4 \int_{\mathbb{R}^n} \sum_{i=1}^n x_i^2 |\partial_i f|^2 \, d\mu(x).$$

Proof. Change variables and use the Brascamp-Lieb inequality. Denote $\frac{d\mu}{dx} = e^{-\psi}$. Then for

$$\pi(x_1, \cdots, x_n) = (x_1^2, \cdots, x_n^2),$$

the function $\psi(\pi(x))$ is convex. Set

$$\varphi(x) = \psi(\pi(x)) - \sum_{i=1}^{n} \log(2x_i).$$

Then π^{-1} pushes-forward μ to the measure with density $e^{-\varphi}$. Moreover,

$$\nabla^2\varphi(x) \geq \nabla^2 \left(-\sum_{i=1}^{n} \log(2x_i) \right) = \begin{pmatrix} \frac{1}{x_1^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{x_2^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{x_n^2} \end{pmatrix} > 0,$$

and therefore

$$\left(\nabla^2\varphi(x)\right)^{-1} \leq \begin{pmatrix} x_1^2 & 0 & \cdots & 0 \\ 0 & x_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_n^2 \end{pmatrix}.$$

Set $g(x) = f(\pi(x))$. By the Brascamp-Lieb inequality,

$$\mathrm{Var}_{e^{-\varphi}}(g) \leq \int_{\mathbb{R}^n_+} \left[\left(\nabla^2\varphi\right)^{-1} \nabla g \cdot \nabla g \right] e^{-\varphi(x)}\, dx \leq \int_{\mathbb{R}^n_+} \sum_{i=1}^{n} x_i^2 |\partial_i g(x)|^2 e^{-\varphi(x)}\, dx.$$

The corollary follows since

$$\mathrm{Var}_{e^{-\varphi}}(g) = \mathrm{Var}_{e^{-\psi}}(f).$$

and since when $y = \pi(x) = (x_1^2, \cdots, x_n^2)$ we have

$$x_i \partial_i g(x) = 2 y_i \partial_i f(y).$$

□

Corollary 7. *Suppose that X is a random vector that is log-concave, isotropic and unconditional in $\mathbb{R}^n$. Then,*

$$Var(|X|) \leq C.$$

Proof.

$$\begin{aligned} Var(|X|) &\leq \mathbb{E}(|X| - \sqrt{n})^2 \leq \frac{1}{n}\mathbb{E}(|X|^2 - n)^2 = \frac{1}{n} Var(|X|^2) \\ &\leq \frac{4}{n}\sum_{i=1}^{n} \mathbb{E}X_i^2 (2X_i)^2 = \frac{16}{n}\sum_{i=1}^{n} \mathbb{E}X_i^4 \leq \frac{C}{n}\sum_{i=1}^{n} (\mathbb{E}X_i^2)^2 \leq C \end{aligned}$$

where we used reverse Hölder inequalities in the last passage. □

This thin shell bound is optimal. The thin shell question is thus completely solved for unconditional convex bodies (K. '09), but for the spectral gap question there is still a $\log n$ gap between the known lower and upper bounds.

The Bochner formula states that in the log-concave case, for any $u \in C_c^\infty(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} (Lu)^2 d\mu = \int_{\mathbb{R}^n} [(\nabla^2 \psi) \nabla u \cdot \nabla u] d\mu + \int_{\mathbb{R}^n} \|\nabla^2 u\|_{HS}^2 d\mu \geq \int_{\mathbb{R}^n} \|\nabla^2 u\|_{HS}^2 d\mu.$$

Let us dualize this and obtain a Poincaré type inequality. To this end, for $f \in L^2(\mu)$ we define the dual Sobolev norm

$$\|f\|_{H^{-1}(\mu)} = \sup \left\{ \int_{\mathbb{R}^n} fu d\mu \, ; \int_{\mathbb{R}^n} |\nabla u|^2 d\mu \leq 1 \, \forall u \in C_c^\infty(\mathbb{R}^n) \right\}.$$

This supremum can be finite only when $\int f d\mu = 0$. In this case, it has a geometric interpretation as infinitesimal transport cost:

$$\|f\|_{H^{-1}(\mu)} \approx \frac{1}{\varepsilon} W_2(\mu, (1 + \varepsilon f)\mu)$$

where $W_2(\mu_1, \mu_2) = \inf_{(\pi_i)_* \gamma} \sqrt{\int_{\mathbb{R}^n \times \mathbb{R}^n} |x - y|^2 d\gamma(x, y)}$ is the Wasserstein L^2-distance in Optimal Transport.

Proposition 8. *(H^{-1}-inequality) Let μ be log-concave, $f \in L^2(\mu)$. Then,*

$$Var_\mu(f) \leq \|\nabla f\|_{H^{-1}(\mu)}^2 = \sum_{i=1}^n \|\partial^i f\|_{H^{-1}(\mu)}^2.$$

Proof. We may assume that $\int f d\mu = 0$. By approximation, assume that $f = -Lu$ for $u \in C_c^\infty(\mathbb{R}^n)$. Then,

$$\int f^2 d\mu = \int [\nabla f \cdot \nabla u] d\mu \leq \|\nabla f\|_{H^{-1}(\mu)} \sqrt{\int_{\mathbb{R}^n} \|\nabla^2 u\|_{HS}^2 d\mu} \leq \|\nabla f\|_{H^{-1}(\mu)} \sqrt{\int_{\mathbb{R}^n} (Lu)^2 d\mu},$$

and the proposition follows. □