Math/Stat 523, Spring 2020

Jon A. Wellner

Lecture 17

Wednesday, June 3; Friday, June 5

Heat and Diffusion Equations & Log-Concavity

- Log-concavity: basic facts
- Diffusion & Log-concavity
- Borell's Section 2

Log-concavity: basic facts

A non-negative function f is log-concave if $x \mapsto \log f(x)$ is concave. Equivalently,

$$f(x) = \exp(-\psi(x))$$

where ψ is convex. Note that if $\theta_0 + \theta_1 = 1$ and $x_0, x_1 \in \mathbb{R}^n$, then concavity of $\log f$ can be written as

$$\theta_0 \log f(x_0) + \theta_1 \log f(x_1) \le \log f(\theta_0 x_0 + \theta_1 x_1),$$

or, equivalently

$$f^{\theta_0}(x_0) f^{\theta_1}(x_1) \le f(\theta_0 x_0 + \theta_1 x_1) \equiv f(x_\theta).$$

Now suppose that f_0, f_1 are two non-negative functions satisfying

$$f_0^{\theta_0}(x_0) f_1^{\theta_1}(x_1) \le f_\theta(\theta_0 x_0 + \theta_1 x_1) \equiv f_\theta(x_\theta)$$

for all $x_0, x_1 \in \mathbb{R}^n$. Then if A_0, A_1 are two Borel subsets of $\mathbb{R}^n$, the Prékopa - Leindler inequality says that

$$\left(\int_{A_0} f_0(x)dx\right)^{\theta_0} \cdot \left(\int_{A_1} f_1(x)dx\right)^{\theta_1} \le \int_{A_\theta} f_\theta(x)dx \tag{1}$$

where $A_\theta \equiv \theta_0 A_0 + \theta_1 A_1$. In the special case with $f_j = 1$ for $j = 0, 1, \theta$ this becomes

$$m_n(A_\theta) \ge m_n^{\theta_0}(A_0) m_n^{\theta_1}(A_1)$$

where m_n denotes Lebesgue measure on $\mathbb{R}^n$. Thus the Prékopa - Leindler inequality implies the classical Brunn-Minkowski inequality in the form:

$$m_n^{1/n}(A_0 + A_1) \ge m_n^{1/n}(A_0) + m_n^{1/n}(A_1).$$

Conversely, the classical Brunn-Minkowski inequality implies the Prékopa - Leindler inequality. See Gardner (2002), section 3.

Introduction: C. Borell

- Let $c : K \subset \mathbb{R}^n \mapsto \mathbb{R}$ where K is convex and bounded.
- Let $H_c \equiv (-1/2)\Delta + c(x)$ in K together with the Dirichlet boundary condition 0. Here Δ is the Laplace operator and c is called the potential function.
- Think of c as being convex or satisfying some other shape constraint, e.g. c is $-1/2-$concave; i.e. $c^{-1/2}$ is concave.

More background: Let $p_{\sigma,c}(t, x, y)$ denote the fundamental solution of the diffusion equation

$$\frac{\partial v}{\partial t} = \frac{\sigma^2}{2}\Delta v - \frac{1}{\sigma^2}c(x)v, \quad \text{for} \quad t > 0, \; x \in K$$

with the Dirichlet boundary condition zero on $t > 0$. Write $p_c(t, x, y) \equiv p_{1,c}(t, x, y)$

Theorem: (Brascamp and Lieb, 1975 - 1976). If c is convex, $(x,y) \mapsto p_c(t,x,y)$ is log-concave for each fixed $t > 0$.

Theorem: (Borell, 1993) If the potential function c is $-1/2$-concave, then the function

$$(s,x,y) \mapsto s \log\{s^n p_c(s^2,x,y)\}$$

is a concave function of $(s,x,y) \in (0,\infty) \times K \times K$.

Now suppose that the potential function c depends on the parameter σ as well as on the position $x \in \mathbb{R}^n$. Let $0 < \alpha \le \beta$. If $c_\sigma(x) \equiv c(x,\sigma)$, and the function

$$\frac{c(x,\sigma)}{\sigma}, \quad x \in K, \quad \alpha \le \sigma \le \beta$$

is convex, then have the following following corollary of Borell's Theorem 3.2:

Corollary: The function

$$(\sigma, x, y) \mapsto \sigma \mathsf{log}\{\sigma^n p_{c_\sigma,\sigma}(t,x,y)\}, \quad (\sigma, x, y) \mapsto [\alpha, \beta] \times K \times K$$

is concave for fixed $t > 0$.

Claim: The results of Brascamp and Lieb (1976) and of Borell (1993) are both consequences of this corollary.

C. Borell's section 2: Hamilton-Jacobi-Bellman equation

Let $\sigma > 0$ and consider the diffusion equation

$$\frac{\partial v}{\partial t} = \frac{\sigma^2}{2}\Delta v - \frac{1}{\sigma^2}c(x)v, \quad t > 0, \ x \in \mathbb{R}^n$$

with the initial condition $v(0,x) = f(x)$, $x \in \mathbb{R}^n$ where $f(x) > 0$ for all $x \in \mathbb{R}^n$. The substitutions

$$V \equiv -\sigma^2 \log v, \qquad \text{and} \qquad F \equiv -\sigma^2 \log f \quad (??)$$

reduce the above Cauchy problem to the Hamilton-Jacobi-Bellman equation

$$\frac{\partial V}{\partial t} + \frac{1}{2}|\nabla V|^2 - c(x) = \frac{\sigma^2}{2}\Delta V, \quad t > 0, \quad x \in \mathbb{R}^n$$

with the initial condition $V(0,x) = F(x)$, $x \in \mathbb{R}^n$.

For the initial calculations in this section we assume that c and F are infinitely differentiable with bounded derivatives of all orders ≥ 0. Later we follow the methods used by Fleming and Soner (1993).

Suppose that $t > 0$ is fixed and let P be Wiener measure on the Banach space Ω of all continuous functions ω of $[0, t]$ into $\mathbb{R}^n$ with $\omega(0) = 0$. If $B(\omega) = \omega = (\omega_1(s), \ldots \omega_n(s))_{0 \leq s \leq t}$, $\omega \in \Omega$, then B is a normalized Brownian motion in $\mathbb{R}^n$ relative to the probability measure P; that is, B is a centered Gaussian process in $\mathbb{R}^n$ relative to the prob. measure P with

$$E^P[B_i(s_0)B_j(s_1)] = \begin{cases} 0, & i \neq j \\ \min\{s_0, s_1\}, & i = j. \end{cases}$$

By setting

$$B_x^\sigma(s) = x + \sigma B(s), \quad s \geq 0$$

The Feynman-Kac formula yields

$$v(t, x) = E^P \left\{ \exp\left(-\frac{1}{\sigma^2} [F(B_x^\sigma(t)) + \int_0^t c(B_x^\sigma(s)) ds] \right) \right\},$$

and the assumptions on c and F imply that

$$\inf_{0 \le s \le t,\, x \in \mathbb{R}^n} v(s,x) > 0 \tag{2.2}$$

s and

$$\sup_{0 \le s \le t,\, x \in \mathbb{R}^n} |\nabla v(s,x)| < \infty. \tag{2.3}$$

Let $u(s)$, $0 \le s \le t$, be a bounded, progressively measurable process (recall Durrett (1996), Section 2.1, page 36), and set

$$h(s) = h_u(s) = \int_0^s u(r)dr$$

and

$$dQ(\omega) = \exp\left(-\frac{1}{2\sigma^2}\int_0^t |u(s)|^2 ds - \frac{1}{\sigma}\int_0^t u(s)d\omega(s)\right) dP(\omega)$$

Then by Girsanov's theorem, for any positive measurable function φ on Ω,

$$\int_\Omega \varphi\left(\omega + \frac{1}{\sigma}h\right) dQ(\omega) = \int_\Omega \varphi(\omega) dP(\omega)$$

and it follows that

$$\begin{aligned} v(t,x) &= E^P\left\{\exp\left(-\frac{1}{\sigma^2}\left[F(B_x^\sigma(t)) + \int_0^t c(B_x^\sigma(s))ds\right]\right)\right\} \\ &= E^Q\left\{\exp\left(-\frac{1}{\sigma^2}\left[F(B_x^\sigma(t)+h(t)) + \int_0^t c(B_x^\sigma(s)+h(s))ds\right]\right)\right\} \\ &= E^P\left\{\exp\left(-\frac{1}{\sigma^2}\left[F(B_x^\sigma(t)+h(t)) + \int_0^t c(B_x^\sigma(s)+h(s))ds\right]\right)\right. \\ &\qquad \left.\cdot \exp\left(-\frac{1}{2\sigma^2}\int_0^t |u(s)|^2 ds - \frac{1}{\sigma}\int_0^t u(s)d\omega(s)\right)\right\}. \end{aligned}$$

We write, for short,

$$X(s) = X_u(s) = B_x^\sigma(s) + h_u(s), \quad 0 \le s \le t$$

so that

$$\begin{aligned} v(t,x) &= E^P \Big\{ \exp\Big(-\frac{1}{\sigma^2}\Big[F(X(t)) + \int_0^t c(X(s))ds \Big]\Big) \\ &\qquad \cdot \exp\Big(-\frac{1}{2\sigma^2}\int_0^t |u(s)|^2 ds - \frac{1}{\sigma}\int_0^t u(s)d\omega(s)\Big)\Big\} \\ &\equiv E^P \exp\Big(-\frac{1}{\sigma^2} Y_u(t)\Big). \end{aligned}$$

Then it follows by Jensen's inequality (since $w \mapsto -\log w$ is convex)

$$-\log v(t,x) = -\log E^P \exp\Big(-\frac{1}{\sigma^2} Y_u(t)\Big) \le E^P \Big(\frac{1}{\sigma^2} Y_u(t)\Big),$$

or, equivalently

$$\log v(t,x) \ge -\frac{1}{\sigma^2} E^P Y_u(t). \tag{2.4}$$

where

$$Y_u(t) = F(X(t)) + \int_0^t (c(X(s)) + \frac{1}{2}|u(s)|^2)ds + \sigma \int_0^t u(s)d\omega(s).$$

Note that

$$E^P\left\{\int_0^t u(s)d\omega(s)\right\} = 0.$$

If we choose u in an appropriate way, it turns out that the random variable $Y_u(t)$ is constant with probability one, which implies that equality occurs in (2.4) for this choice of u. To find such a process u, first define

$$U(s,x) \equiv -\nabla V(t-s,x), \quad 0 \le s \le t.$$

From the assumptions on c and F we conclude that the function $U(x,s)$, $0 \le s \le t$, $x \in \mathbb{R}^n$ is bounded and continuous and, moreover, the bounds (2.2) and (2.3) imply that there exists a constant $C > 0$ such that

$$|U(s,x) - U(s,y)| \le C|x-y|, \quad 0 \le s \le t, \quad x,y \in \mathbb{R}^n.$$

Therefore the stochastic differential equation

$$dX(s) = U(s, X(s))ds + \sigma d\omega(s), \quad 0 \le s \le t$$

with the initial condition $X(0) = x$ possesses a unique solution. (For example, see Durrett (1996), Chapter 5, Theorem 2.2, page 185, and Theorem 2.9, page 190.) We set $u_0(s) \equiv U(s, X(s))$, $0 \le s \le t$, and we have

$$X(s) = x + \sigma\omega(s) + h_{u_0}(s) = B_x^\sigma(s) + h_{u_0}(s), \quad 0 \le s \le t.$$

Moreover, we claim that the random variable $Y_{u_0}(t)$ is constant with probability 1. To prove this claim, consider the process

$$\xi(s) = V(t-s, X(s)) + \int_0^s (c(X(r)) + \frac{1}{2}|u_0(r)|^2)dr + \sigma \int_0^s u_0(r)d\omega(r)$$

defined for all $0 \leq s \leq t$. Then, by Itô's lemma

$$\begin{aligned} d\xi(s) &= -V_t(t-s, X(s))ds + \nabla V(t-s, X(s)) \cdot (u_0(s)ds + \sigma d\omega(s)) \\ &+ \frac{\sigma^2}{2}\Delta V(t-s, X(s))ds + (c(X(s)) + \frac{1}{2}|u_0(s)|^2)ds \\ &+ \sigma u_0(s)d\omega(s). \end{aligned}$$

Since the function $V(t,x)$ satisfies the Hamilton-Jacobi-Bellmann equation (2.1), $d\xi(s) = 0$, and we conclude that ξ is constant with probability 1, which was to be proved.

From the above,

$$v(t,x) = \exp\left(-\inf_{u \in \mathcal{U}(t)} J(t,x,u)\right)$$

where

$$J(t,x,u) = \frac{1}{\sigma^2} E^P(Y_u(t))$$

and where $\mathcal{U}(t)$ denotes the class of all bounded, progressively measurable processes $u(s)$, $0 \leq s \leq t$.

In the following, let

$$v_{\sigma,c}^{F}(t,x) = E^{P}\left\{\exp\left(-\frac{1}{\sigma^2}\left[F(B_x^{\sigma}(t)) + \int_0^t c(B_x^{\sigma}(s))ds\right]\right)\right\}$$

and

$$J_{\sigma,c}^{F}(t,x,u)$$
$$= \frac{1}{\sigma^2}E^{P}\left\{F(B_x^{\sigma}(t)+h_u(t)) + \int_0^t \left[c(B_x^{\sigma}(s)+h_u(s)) + \frac{1}{2}|u(s)|^2\right]ds\right\}$$

for all continuous and bounded functions F and c on $\mathbb{R}^n$. Then from the above it is straightforward to conclude that

$$v_{\sigma,c}^{F}(t,x) = \exp\left(-\inf_{u\in\mathcal{U}(t)} J_{\sigma,c}^{F}(t,x,u)\right).$$

We will also use the short-hand notation

$$v_{c,\sigma}^{A,F} \equiv E^{P}\left\{1_A(B_x^{\sigma}(t))\exp\left(-\frac{1}{\sigma^2}\left[F(B_x^{\sigma}(t)) + \int_0^t c(B_x^{\sigma}(s))ds\right]\right)\right\}$$

for any Borel set A in $\mathbb{R}^n$.

C. Borell's Section 3: Application to Diffusion Equations

In the following $\theta = (\theta_0, \theta_1) \in \mathbb{R}^2$ is fixed with $\theta_0, \theta_1 > 0$. If $x_0, x_1 \in \mathbb{R}^n$, let

$$x_\theta = \theta_0 x_0 + \theta_1 x_1,$$

and, if $A_0, A_1 \subset \mathbb{R}^n$, let

$$A_\theta = \theta_0 A_0 + \theta_1 A_1 = \{x_\theta : x_0 \in A_0,\ x_1 \in A_1\}.$$

Suppose first that $\sigma_0, \sigma_1 > 0$ and let D_i, $i = 0, 1$, be sub-domains of $\mathbb{R}^n$. Below we will often consider functions $\varphi_j : D_j \mapsto \mathbb{R}$, $j = 0, 1, \theta$, which satisfy the inequality

$$\frac{1}{\sigma_\theta}\varphi_\theta(x_\theta) \ \le \ \frac{\theta_0}{\sigma_0}\varphi_0(x_0) + \frac{\theta_1}{\sigma_1}\varphi_0(x_1), \quad x_0 \in D_0,\ x_1 \in D_1.$$

Note that the inequality in the last display holds in the following cases:

Case 1: $\sigma_0 = \sigma_1$, $\theta_0 + \theta_1 = 1$ and

$$\varphi_\theta(x_\theta) \le \theta_0\varphi_0(x_0) + \theta_1\varphi_1(x_1),\ x_0 \in D_0,\ x_1 \in D_1.$$

Case 2: $\varphi_j(x) = \psi_j^2(x)$, $j = 0, 1, \theta$, where the ψ_j are nonnegative and

$$\psi_\theta(x_\theta) \le \theta_0\psi_0(x_0) + \theta_1\psi_1(x_1),\ x_0 \in D_0,\ x_1 \in D_1.$$

Case 3: $\varphi_j(x) = \sigma_j^4/\psi_j^2(x)$, $j = 0, 1, \theta$, where the ψ_j are positive and

$$\psi_\theta(x_\theta) \ge \theta_0\psi_0(x_0) + \theta_1\psi_1(x_1),\ x_0 \in D_0,\ x_1 \in D_1.$$

In connection with the last two cases it is useful to know that the function

$$\gamma_\alpha(\lambda, \sigma) = \frac{\lambda^{\alpha+1}}{\sigma^\alpha}, \quad \lambda \ge 0,\ \sigma > 0$$

is convex and positively homogeneous of degree one for $\alpha = 1$ and $\alpha = 2$ respectively.

(Convexity of γ follows from the fact the perspective of a convex (or concave) function preserves convexity (or concavity). If $\varphi(x)$ is convex then $(x, \lambda) \mapsto \lambda\varphi(x/\lambda)$ is convex. See e.g. VdV & W, *Electronic Journal of Statistics* **5** (2011).)

Discrete log-concavity: Klartag - Lehec

Klartag and Lehec (2019) reformulate the results of Borell (2000) as follows:

Let γ_n be the standard $N_n(0, I)$ Gaussian measure on $\mathbb{R}^n$. Let $(B_t)_{t\geq 0}$ be a standard $n-$dimensional Brownian motion, and let $f : \mathbb{R}^n \to \mathbb{R}$.

Theorem: (Borell (2000) reformulated) The following stochastic variational formula holds:

$$\log\left(\int_{\mathbb{R}^n} e^f d\gamma_n\right) = \sup_{u\in\mathcal{U}}\left\{E\left[f\left(B_1 + \int_0^1 u_s ds\right) - \frac{1}{2}\int_0^1 |u_s|^2 ds\right]\right\}$$

where the supremum is taken over all bounded stochastic processes u which are adapted to the Brownian filtration; i.e. u_t is measurable with respect to the $\sigma-$field generated by $\{B_s : s \leq t\}$ for all $t \in [0, 1]$.

Now we state one of the main results of Klartag and Lehec (2019):

- Let $T > 0$ be a fixed number, and let N be a Poisson point process on $[0, T] \times \mathbb{R}^+ \subset \mathbb{R}^2$ with intensity measure equal to the Lebesgue measure λ. In particular $N(A)$ is a Poisson random variable with parameter $\lambda(A)$ for any Borel set $A \subset [0, T] \times [0, T] \times \mathbb{R}^+$.
- For a Borel subset $B \subset [0, T] \times \mathbb{R}^+$ we write $\mathcal{F}_B$ for the $\sigma-$field generated by the random variables

$$\{N(C) : \ C \in \mathcal{B}_2, C \subset B\}.$$

- For $t \in [0, T]$ we set $\mathcal{F}_t = \mathcal{F}_{[0,t] \times \mathbb{R}^+}$. This defines a filtration of Ω.
- Recall that $(\lambda_t)_{0 \le t \le T}$ is *predictable* if, as a function of $t \in [0, T]$ and $\omega \in \Omega$, it is measurable with respect to the $\sigma-$field

$\mathcal{P}$ generated by the sets

$$\{(s,t] \times A : \; s \le t \le T, \; A \in \mathcal{F}_s\}.$$

It is a standard fact that if a process $(\lambda_t)_{0 \le t \le T}$is left-continuous and adapted, then it is predictable.

- Given a predictable, bounded, non-negative stochastic process $(\lambda_t)_{0 \le t \le T}$ we define the associated counting process $(X_t^\lambda)_{0 \le t \le T}$ via

$$X_t^\lambda = N(\{(s,u) \in [0,T] \times \mathbb{R}^+; \; s < t, \; u \le \lambda_s\}).$$

- Note that for every non-negative predictable process $(H_t)_{0 \le t \le T}$ we have

$$E\left[\int_0^T H_t dX^\lambda(t)\right] = E\left[\int_0^T H_t \lambda_t dt\right],$$

where the integral on the left-side is a Riemann-Stieltjes integral.

Let π_T denote the Poisson measure with parameter T; i.e.

$$\pi_T(n) = \frac{T^n}{n!} e^{-T} \quad \text{for } n \in \mathbb{N} = \{0, 1, 2, \ldots\}.$$

References:

- Gardner, R. J. (2002). The Brunn-Minkowski Inequality. *Bulletin of the Amer. Math. Soc.* **39** (3), 355 - 405.
- Klartag, B. (2019). Poisson processes and a log-concave Bernstein theorem. *Studia Mathematica* **247**(1), 85 - 107.
- Borell, C. (2000). Diffusion equations and geometric inequalities. **12**, 49 - 71.