On diffusion semigroups preserving the log-concavity

A. V. Kolesnikov

Department of Mechanics and Mathematics, Moscow State University,
119899 Moscow, Russia

Abstract

It is proved that if the semigroup with the generator $a^{ij}\partial_i\partial_j + \beta^i\partial_i$, where a^{ij} and b^i are smooth functions, sends log-concave functions to log-concave functions, then $a = (a^{ij})$ is constant and $\beta = (\beta^i)$ is an affine mapping.

AMS 2000 Subject classification:
Primary: 60G15, 52A40
Secondary: 35K10, 60H10

Key words and phrases: log-concave function, semigroup, generator of semigroup, stochastic differential equation.

It is well known that the semigroups

$$\psi \mapsto T_t^0\psi,\ T_t^0\psi(x) = \frac{1}{(2\pi t)^{n/2}} \int_{\mathbb{R}^n} \psi(y) e^{-\frac{1}{2t}(x-y)^2}\, dy$$

and

$$\psi \mapsto T_t^1\psi,\ T_t^1\psi(x) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \psi(e^{-\frac{t}{2}}x + \sqrt{1-e^{-t}}y) e^{-\frac{1}{2}(y,y)}\, dy$$

(the Ornstein-Uhlenbeck semigroup) possess the following property: for every log-concave function $\psi : \mathbb{R}^n \to \mathbb{R}$ the functions $T_t^0\psi$ and $T_t^1\psi$ are log-concave (see [6] for a more general fact). Recall that a function f is called log-concave if it has the form $f = e^{-V}$, where V is a convex function. Such functions are important in analysis, stochastics, in the theory of Gaussian measures (see [1], [2], [4]). In particular, the property of s semigroup to preserve the log-concavity occurs to be useful for proving the correlation inequality for Gaussian measures. The correlation inequality is the inequality

$$\gamma(A \cap B) \geq \gamma(A)\gamma(B),$$

where γ is a centered Gaussian measure on $\mathbb{R}^n$, A and B are convex sets, symmetric about the origin. This inequality has been proved so far only for some certain pairs of sets. For example, in [3] the above mentioned property of the Ornstein-Uhlenbeck semigroup has been used for the case, when one of the sets is an ellipsoid. It is natural to ask what semigroups preserve log-concavity. We shall show that in a certain sense these two main examples exhaust the semigroups with such a property. More precisely, it will be shown, that the above mentioned property only holds for the semigroups generated by Gaussian diffusions (these are the diffusions having the generators with constant coefficients at second derivatives and an affine drift). Recall that the generators of T_t^0 and T_t^1 are, accordingly, $\frac{1}{2}\Delta$ and $\frac{1}{2}\Delta - \frac{1}{2}(x, \nabla)$.

The upper index i will denote the i-th coordinate and the lower index i will denote the derivative along x^i. It will be assumed throughout that $\sigma(x) = (\sigma^{lk})(x)$ and $\beta(x) = (\beta^k)(x)$ are twice continuously differentiable mappings from $\mathbb{R}^n$ with values, accordingly, in the space of matrices $n \times n$ and $\mathbb{R}^n$, such that

$$|\sigma(x)| + |\beta(x)| < K(1 + |x|),$$

$$|[\sigma^{lk}(x)]_i| + |[\beta^k(x)]_i| < K,$$

$$|[\sigma^{kl}(x)]_{ij}| + |[\beta^{kl}(x)]_{ij}| < K(1 + |x|^{m_0})$$

for some constants K and m_0 and any indices i, j, k, l. Then the solution of the following stochastic equation exists:

$$\xi_t(x) = x + \int_0^t \sigma(\xi_s(x))dW_s + \int_0^t \beta(\xi_s(x))ds, \tag{1}$$

where W_t is a n-dimensional Wiener process. The corresponding semigroup is given by

$$T_t : \psi \mapsto \mathbb{E}\psi(\xi_t(x)).$$

Let $a = (a^{ij}) = \frac{1}{2}\sigma\sigma^T$. The generator of $\{T_t\}$ is

$$a^{ij}\partial_i\partial_j + \beta^i\partial_i,$$

where $\partial_i = \partial_{x_i}$ and the summation in the repeated indices is meant.

Lemma. *Let $A(x)$ be a nonnegative symmetric matrix $n \times n$ for every $x \in \mathbb{R}^n$. Suppose that $(-A(x)\omega, \omega)$ is convex for every vector ω. Then A is constant.*

Proof: Let us prove by induction on dimension n. For $n = 1$ the lemma follows from the obvious claim that an unconstant, convex on the whole line function cannot be everywhere negative. Let n=2. The condition $A \geq 0$ implies that $A^{11}(x^1, x^2) \geq 0$ and $A^{22}(x^1, x^2) \geq 0$. Taking $\omega = e^1$ and $\omega = e^2$ we get that $-A^{11}(x^1, x^2)$ and $-A^{22}(x^1, x^2)$ are convex. Consequently, fixing the coordinates x^1 and x^2, we find that A^{11} and A^{22} are constant. Taking $\omega = e^1 + e^2$, we get that the function

$-A^{12} = -A^{21}$ is convex. But the condition $\det A \geq 0$ implies that A^{12} is bounded. This means that A^{12} is constant. If $n > 2$, the claim is easily obtained from the proved one for $n-1$ if we note that matrix obtained by deleting the k-th line and k-th column and fixing an arbitrary coordinate, fulfills the conditions of the lemma. $\square$

The proof of the following theorem is given in [5, Theorem V.7.4].

Theorem 1. *Let a function f and all its derivatives approach zero at the infinity faster than $1/|x|^N$ for every $N > 0$. Then, for all $t > 0$, the function $\psi_t(x) = \mathbb{E}f(\xi_t(x))$ is well defined, where ξ_t fulfills the condition (1). Moreover, the following equalities are valid:*

$$(\psi_t)_i = \mathbb{E}(f)_k(\xi_t^k)_i,$$

$$(\psi_t)_{ij} = \mathbb{E}\left[(f)_{kl}(\xi_t^k)_i(\xi_t^l)_j + (f)_k(\xi_t^k)_{ij}\right],$$

where the processes $(\xi_t^k)_i$ and $(\xi_t^k)_{ij}$ solve the following stochastic differential equations:

$$(\xi_t^k)_i = \delta_i^k + \int_0^t (\sigma^{km})_r(\xi_s^r)_i dW_s^m + \int_0^t (\beta^k)_r(\xi_s^r)_i ds,$$

$$\begin{aligned}(\xi_t^k)_{ij} = &\int_0^t \left[(\sigma^{km})_{rl}(\xi_s^r)_i(\xi_s^l)_j + (\sigma^{km})_r(\xi_s^r)_{ij}\right] dW_s^m \\ &+ \int_0^t \left[(\beta^k)_{rl}(\xi_s^r)_i(\xi_s^l)_j + (\beta^k)_r(\xi_s^r)_{ij}\right] ds,\end{aligned}$$

(in the terms like $(\sigma^{km})_{rl}(\xi_s)$ we omit the argument ξ_s, and the summation in the repeated indices is meant).

Theorem 2. *Let the semigroup $\{T_t\}$ send log-concave functions to log-concave functions. Then a is constant and β is an affine mapping.*

Proof: For a log-concave function $\psi \in C^2(\mathbb{R}^n, \mathbb{R})$, the following matrix $\Phi(\psi)$ must be nonnegative: $(\Phi(\psi))^{ij} := (\psi)_i(\psi)_j - \psi(\psi)_{ij}$.

We prove that if the conditions of the theorem are not satisfied, then we can find a point x_0 and a log-concave function ψ such that, for small t, the matrix $\Phi(\psi_t)$, where $\psi_t = T_t\psi$, is not positive. Let $g : \mathbb{R} \mapsto \mathbb{R}$ be a function, which is smooth, nonnegative, even, convex, increases at the infinity faster than x^2 and equals zero on $[-1, 1]$. Let us consider the function $\psi = e^{(\omega,x)-\sum_{i\leq n} g(x^i - x_0^i)}$, where the vector ω will be chosen later. This function is log-concave and satisfies the condition of Theorem 1, but all the subsequent calculations will be made with the function $e^{(\omega,x)}$ instead of ψ, because the results, as it will be easily seen, will depend only on values of the function ψ in a small neighbourhood of x_0.

We take an arbitrary vector v. Then the expression $(\Phi(e^{(\omega,x)})v,v)$ equals

$$\left|\mathbb{E}e^{(\omega,\xi_t)}(\omega,(\xi_t)_i)v^i\right|^2-\mathbb{E}e^{(\omega,\xi_t)}\mathbb{E}e^{(\omega,\xi_t)}\left[(\omega,(\xi_t)_i)(\omega,(\xi_t)_j)v^iv^j+(\omega,(\xi_t)_{ij})v^iv^j\right].$$

We shall calculate the derivative of $(\Phi(e^{(\omega,x)})v,v)$ at the moment $t=0$. For this purpose we shall find the Ito-differentials of processes occuring in the formula, i. e.

$$de^{(\omega,\xi_t)}=e^{(\omega,\xi_t)}\left[(\omega,d\xi_t)+\frac{1}{2}(\sigma\sigma^T\omega,\omega)dt\right],$$

$$\begin{aligned}
&de^{(\omega,\xi_t)}(\omega,(\xi_t)_i)v^i=e^{(\omega,\xi_t)}(\omega,(\xi_t)_i)v^i\left[(\omega,d\xi_t)+\frac{1}{2}(\sigma\sigma^T\omega,\omega)dt\right]\\
&+e^{(\omega,\xi_t)}v^i\Big(\omega^k,(\sigma^{km})_r(\xi_t^r)_i dW_t^m\Big)+e^{(\omega,\xi_t)}v^i\Big(\omega^k,(\beta^k)_r(\xi_t^r)_i dt\Big)\\
&+e^{(\omega,\xi_t)}v^i\Big(\omega^k,(\sigma^{km})_r(\xi_t^r)_i\Big)(\omega^k,\sigma^{km})dt,
\end{aligned}$$

$$\begin{aligned}
&de^{(\omega,\xi_t)}\Big((\omega,(\xi_t)_i)v^i\Big)^2=e^{(\omega,\xi_t)}\Big((\omega,(\xi_t)_i)v^i\Big)^2\left[(\omega,d\xi_t)+\frac{1}{2}(\sigma\sigma^T\omega,\omega)dt\right]\\
&+2e^{(\omega,\xi_t)}v^i\Big(\omega,(\xi_t)_i\Big)v^i\Big(\omega,d(\xi_t)_i\Big)+e^{(\omega,\xi_t)}\omega^kv^i\omega^lv^j(\sigma^{km})_{r_1}(\sigma^{lm})_{r_2}(\xi_t^{r_1})_i(\xi_t^{r_2})_jdt\\
&+2e^{(\omega,\xi_t)}\Big(v^i(\omega,(\xi_t)_i)\Big)\Big(\omega^kv^i(\sigma^{km})_r\omega^j(\sigma^{jm})(\xi_t^r)_i\Big)dt,
\end{aligned}$$

$$\begin{aligned}
&de^{(\omega,\xi_t)}(\omega,(\xi_t)_{ij})v^iv^j\\
&=e^{(\omega,\xi_t)}\Big((\omega,(\xi_t)_{ij})v^iv^j\Big)\left[(\omega,d\xi_t)+\frac{1}{2}(\sigma\sigma^T\omega,\omega)dt\right]+e^{(\omega,\xi_t)}\omega^kv^iv^jd(\xi_t^k)_{ij}\\
&+e^{(\omega,\xi_t)}\omega^{k_1}(\sigma^{k_1m})(\omega^{k_2}v^iv^j)\left[(\sigma^{k_2m})_{rl}(\xi_t^r)_i(\xi_t^l)_j+(\sigma^{k_2m})_r(\xi_t^r)_{ij}\right]dt.
\end{aligned}$$

We shall use the previous calculations and the following formula:

$$\frac{d}{dt}\mathbb{E}\eta_t\Big|_{t=0}=b(x),$$

where $\eta_t=x+\int_0^t adW_s+\int_0^t bds$. We obtain

$$\frac{d}{dt}\mathbb{E}e^{(\omega,\xi_t)}\Big|_{t=0}=e^{(\omega,x)}\left[\big(\omega,\beta(x)\big)+\frac{1}{2}\big(\sigma(x)\sigma^T(x)\omega,\omega\big)\right],$$

$$\frac{d}{dt}\mathbb{E}e^{(\omega,\xi_t)}\big(\omega,(\xi_t)_i\big)v^i\Big|_{t=0}$$
$$= e^{(\omega,x)}\left\{\left[\big(\omega,\beta(x)\big)+\frac{1}{2}\big(\sigma(x)\sigma^T(x)\omega,\omega\big)\right](\omega,v)^2+2(\omega,v)\big(v^i(\beta_i(x),\omega)\big)\right\},$$

$$\frac{d}{dt}\mathbb{E}e^{(\omega,\xi_t)}\Big(\big(\omega,(\xi_t)_i\big)v^i\Big)^2\Big|_{t=0}$$
$$= e^{(\omega,x)}\left\{\left[\big(\omega,\beta(x)\big)+\frac{1}{2}\big(\sigma(x)\sigma^T(x)\omega,\omega\big)\right](\omega,v)^2+2(\omega,v)\big(v^i(\beta_i(x),\omega)\big)\right\}$$
$$+ e^{(\omega,x)}\left[v^iv^j\big(\sigma_i(x)\sigma_j^T(x)\omega,\omega\big)+2(\omega,v)\big(\sigma(x)\sigma_i^T(x)\omega,\omega\big)v^i\right],$$

$$\frac{d}{dt}\mathbb{E}e^{(\omega,\xi_t)}(\omega,(\xi_t)_{ij})v^iv^j\Big|_{t=0}= e^{(\omega,x)}v^iv^j\left[\big(\omega,\beta_{ij}(x)\big)+\big(\sigma(x)\sigma_{ij}^T(x)\omega,\omega\big)\right].$$

It follows from the previous expressions that

$$\frac{d}{dt}\big(\Phi(e^{(\omega,\xi_t)})v,v\big)\Big|_{t=0}$$
$$=2e^{2(\omega,x)}(\omega,v)\left[\big(\omega,\beta(x)\big)+\frac{1}{2}\big(\sigma(x)\sigma^T(x)\omega,\omega\big)\right](\omega,v)$$
$$+ 2e^{2(\omega,x)}(\omega,v)\left[v^i\big(\beta_i(x),\omega\big)+v^i\big(\sigma(x)\sigma_i^T(x)\omega,\omega\big)\right]$$
$$- e^{2(\omega,x)}(\omega,v)^2\left[\big(\omega,\beta(x)\big)+\frac{1}{2}\big(\sigma(x)\sigma^T(x)\omega,\omega\big)\right]$$
$$- e^{2(\omega,x)}\left\{\left[\big(\omega,\beta(x)\big)+\frac{1}{2}\big(\sigma(x)\sigma^T(x)\omega,\omega\big)\right](\omega,v)^2+2(\omega,v)\Big(v^i\big(\beta_i(x),\omega\big)\Big)\right\}$$
$$- e^{2(\omega,x)}\left[v^iv^j\big(\sigma_i(x)\sigma_j^T(x)\omega,\omega\big)+2(\omega,v)\big(\sigma(x)\sigma_i^T(x)\omega,\omega\big)v^i\right]$$
$$- e^{2(\omega,x)}v^iv^j\left[\big(\omega,\beta_{ij}(x)\big)+\big(\sigma(x)\sigma_{ij}^T(x)\omega,\omega\big)\right]$$
$$=- e^{2(\omega,x)}v^iv^j\left[\big((\beta)_{ij},\omega\big)+\Big(\big(\sigma(\sigma^T)_{ij}+(\sigma)_i(\sigma^T)_j\big)\omega,\omega\Big)\right].$$

It is easy to see that

$$\big(\Phi(\psi)v,v\big)(x_0)=\big(\Phi(e^{(\omega,x)})v,v\big)(x_0)=0$$

and

$$\frac{d}{dt}\big(\Phi(\psi_t)v,v\big)\Big|_{t=0}(x_0)=\frac{d}{dt}\big(\Phi(\mathbb{E}e^{(\omega,\xi_t)})v,v\big)\Big|_{t=0}(x_0)$$

because ψ and $e^{(\omega,x)}$ coincide in some neighbourhood of x_0. We have only to show that if the conditions of the theorem are not valid, then we can find ω, v and x_0, such that $\frac{d}{dt}(\Phi(\mathbb{E}e^{(\omega,\xi_t)})v,v)\big|_{t=0}(x_0) < 0$. Indeed, if we assume that the opposite inequality fulfilles, then taking $\omega = \pm\varepsilon e^k$ and choosing ε small enough, we get $\beta^k_{ij} = 0$. So, we see that β is affine. Therefore, we obtain that the following condition must be valid: $v^i v^j(\sigma\sigma^T\omega,\omega)_{ij} \le 0$. This means that $(-\sigma\sigma^T\omega,\omega)$ is convex. Finally, the lemma implies that $\sigma\sigma^T$ is a constant matrix. $\square$

Note that the matrix $\sigma(x)$ from the equation (1) may be non-constant. For example, $\sigma(x)$ may be unitary for every x.

Proposition. *Let a be constant and let β be an affine mapping. Then the semigroup T_t preserves log-concavity.*

Proof. Since the distribution of the solution of (1) depends only on x, β and a, we can assume without loss of generality that σ is also constant. Equation (1) takes the form:

$$\xi_t(x) = x + \sigma W_t + \int_0^t (\beta_0 + L\xi_s(x))ds,$$

where β_0 is a vector and L is some linear mapping. It is easy to see that ξ_t is a Gaussian process. Taking the expectations of both sides, we get

$$\mathbb{E}\xi_t(x) = x + \int_0^t (\beta_0 + L\mathbb{E}\xi_s(x))ds. \tag{2}$$

Let $v(t)$ be the solution of the equation: $v(t) = \int_0^t(\beta_0 + Lv(s))ds$. Since the difference $\mathbb{E}\xi_t(x) - v(t)$ is the solution of the linear differential equation with initial point x, it depends linearly on x. On the other hand, since the process $\nu_t = \xi_t - \mathbb{E}\xi$ solves the stochastic equation $\nu_t = \sigma W_t + \int_0^t L(\nu_s)ds$, it does not depend on x. So, we see that ξ_t is a Gaussian process whose mean has the form $\mathbb{E}\xi_t(x) = v(t) + M_t x$ for some linear mapping $M_t : \mathbb{R}^n \mapsto \mathbb{R}^n$ depending only on t and whose covariance matrix does not depend on x. This means that for every t we can choose a linear subspace $E_t \subset \mathbb{R}^n$ and a quadratic form Q_t on E_t such that in the suitable coordinate system

$$\mathbb{E}f(\xi_t(x)) = \int_{E_t} f(y)e^{-Q_t(y-v(t)-M_t x)}dy.$$

Obviously, for the log-concave function $f(y)$, the above integrand is a log-concave function of (x,y). By the Prékopa theorem (see [6]) for every log-concave function $g(x,y)$ the integral $\int g(x,y)dy$ is a log-concave function of x. Hence, we obtain our claim. $\square$

The work has been supported by the Russian Foundation of Fundamental Research project 00-15-99267 and the DFG Grant 436 RUS 113/343/0(R). The author thanks V.I. Bogachev for formulation of the problem.

References

[1] *Bogachev V.I.*, Gaussian Measures, Amer. Math. Soc., Rhode Island, 1998.

[2] *Borell C.*, Convex measures on locally convex spaces, Ark. Math. 12 (1974), no. 2, 239—252.

[3] *Hargé G.*, Une inégalité de décorrélation pour la mesure gaussienne, C. R. Acad. Sci. Paris, 326 (1998), Série 1, 1325—1328.

[4] *Krugova E.P.*, On shifts of convex measures, Mat. Sbornik 188 (1997), no. 2, 57—66.

[5] *Krylov N.V.*, Introduction to the Theory of Diffusion Processes, Amer. Math. Soc., Providence, Rhode Island, 1995.

[6] *Prékopa A.*, On logarithmic concave measures and functions, Acta Sci. Math. 34 (1973), 335—343.