1 Gaussian hypercontractivity

In this report we review the derivation of Nelson's hypercontractive inequality for the L^2 function space under gaussian measure. Hypercontractive inequalities are important for deriving various concentration of measure results, but we refer to the interested reader to [LeDoux] and [Chatterjee] for the motivation behind hypercontractivity.

We take material from [LeDoux] and [Chatterjee], but add details and intuition when possible.

1.1 Preliminaries

Consider the Gaussian measure,

$$\mu(A) = C_n \int_A e^{-U(x)} dt$$

for $U(x) = -||x||_2^2/2$, where C_n is a normalizing constant (explicitly, $C_n = (2\pi)^{-n/2}$). Define the operator

$$L = \Delta - \nabla U \cdot \nabla$$

Define the Dirichlet form $\mathcal{E}(f,g) = \mu(f(-Lg))$. We claim that

$$\mathcal{E}(f,g) = \mu(f(-Lg)) = \int f(-Lg) d\mu = \int \nabla f \cdot \nabla g d\mu \tag{1}$$

We will now derive (1). Recall that if Ω is an open bounded subset of $\mathbb{R}^n$ with smooth boundary Γ, $\hat{v}$ is the unit surface normal to Γ, u a function and $\mathbf{v}$ a vector-valued function, both continuously differentiable, then

$$\int_\Omega \nabla u \cdot \mathbf{v}(x) d\Omega = \int_\Gamma u(\mathbf{v} \cdot \hat{v}) d\Gamma - \int_\Omega u \nabla \cdot \mathbf{v} d\Omega \tag{2}$$

Note that

$$\begin{aligned}
\nabla \cdot ((\nabla g(x)) e^{-U(x)}) &= (\nabla \cdot \nabla g(x)) e^{-U(x)} + \nabla g(x) \cdot (\nabla e^{-U(x)}) \\
&= \Delta g(x) e^{-U(x)} + \nabla g(x) \cdot (-\nabla U(x) e^{-U(x)}) \\
&= (\Delta g(x) - \nabla g(x) \cdot \nabla U(x)) e^{-U(x)} \\
&= L g(x) e^{-U(x)}
\end{aligned}$$

Furthermore, note that $\nabla g(x)e^{-U(x)}$ vanishes as $||x||_2 \to \infty$. Therefore applying integration py parts, we can ignore the boundary term in (2) and obtain

$$\begin{aligned}\int_{\mathbb{R}^n} f(x)(-Lg(x))e^{-U(x)}dx &= -\int_{\mathbb{R}^n} f(x)\nabla \cdot ((\nabla g(x))e^{-U(x)})dx \\ &= \int_{\mathbb{R}^n} \nabla f(x) \cdot ((\nabla g(x))e^{-U(x)})dx\end{aligned}$$

which, up to normalization terms, gives $\int f(-Lg)d\mu = \int \nabla f \cdot \nabla g d\mu$.

L is an infinitesimal generator for the semigroup of operators P_t, by the *heat* equation

$$\frac{\partial}{\partial t} P_t f(x)|_{t=u} = LP_u f(x)$$

which is also written as $\partial_t P_t = LP_t$. Alternatively,

$$P_t = e^{tL} = \sum_{k=0}^{\infty} \frac{t^k}{k!} L^k$$

Here is an intuitive explanation for understanding the equivalence of the two forms. For small δ, we know that by definition, $P_{t+\delta} f \approx f + \delta LP_t f$. However, to get an accurate approximation for P_{t+u} for u large, we should first approximate

$$P_{t+(u/k)} \approx f + \frac{u}{k} LP_t f = (I + \frac{u}{k})f$$

where I is the identity operator, then approximate

$$P_{t+2(u/k)} \approx f + \frac{u}{k} LP_{t+(u/k)} f = (I + \frac{u}{k})P_{t+(u/k)} f \approx (I + \frac{u}{k} L)^2 f$$

and so on, hence

$$P_{t+u} f \approx (I + \frac{u}{k} L)^k f$$

Taking $k \to \infty$, we get the exponential form $P_{t+u} = e^{uL} P_t$. The semigroup property follows automatically from the exponential form:

$$P_{t+s} = e^{(t+s)L} = e^{tL} e^{sL} = P_t P_s$$

given the appropriate conditions for the convergence of the infinite series, etc.

1.2 The Ornstein-Uhlembeck Semigroup

In the case of the gaussian measure, where $U(x) = -||x||^2/2$, we know the explicit form of P_t:

$$P_t f(x) = C_n \int f(e^{-t}x + \sqrt{1-e^{-2t}}z)e^{-||z||^2/}dz = \mathbf{E}[f(e^{-t}x + \sqrt{1-e^{-2t}}Z)]$$

An important consequence is that

$$\lim_{t\to\infty} f(x) = \mathbf{E}[f(Z)] = \gamma^n(f). \tag{3}$$

Another consequence is that for any $t \geq 0$,

$$\mathbf{E}[P_t f(Z)] = \mathbf{E}[f(e^{-t}Z + \sqrt{1-e^{-2t}}Z')] = \mathbf{E}[f(Z)] = \gamma^n(f) \tag{4}$$

where Z, Z' are independent standard normal variates, hence $e^{-t}Z+\sqrt{1-e^{-2t}}Z'$ has the same distribution as Z.

Let us verify that P_t satisfies the heat equation. On one hand, assuming we can differentiate under the integral sign,

$$\frac{\partial}{\partial t}P_t f(x) = \mathbf{E}\left[\frac{\partial}{\partial t} f(e^{-t}x + \sqrt{1-e^{-2t}}Z)\right] \tag{5}$$

$$= \mathbf{E}\left[\nabla f(e^{-t}x + \sqrt{1-e^{-2t}}Z)\left(e^{-t} + \frac{e^{-2t}}{\sqrt{1-e^{-2t}}}Z\right)\right] \tag{6}$$

$$= e^{-t}x\mathbf{E}[\nabla f(e^{-t}x + \sqrt{1-e^{-2t}}Z)] + \frac{e^{-2t}}{\sqrt{1-e^{-2t}}}\mathbf{E}[Z\nabla f(e^{-t}x + \sqrt{1-e^{-2t}}Z)] \tag{7}$$

Using the identity $\mathbf{E}Zg(Z) = \mathbf{E}g'(Z)$

$$= e^{-t}x\mathbf{E}[\nabla f(e^{-t}x + \sqrt{1-e^{-2t}}Z)] + \frac{e^{-2t}}{\sqrt{1-e^{-2t}}}\mathbf{E}\left[\sum_{i=1}^n \frac{\partial^2}{\partial Z_i^2} f'(e^{-t}x + \sqrt{1-e^{-2t}}Z)\right] \tag{8}$$

$$= e^{-t}x\mathbf{E}[\nabla f(e^{-t}x + \sqrt{1-e^{-2t}}Z)] + \frac{e^{-2t}}{\sqrt{1-e^{-2t}}}\mathbf{E}\left[\sum_{i=1}^n \sqrt{1-e^{-2t}}\frac{\partial^2 f(w)}{\partial w_i^2}\bigg|_{w=e^{-t}x+\sqrt{1-e^{-2t}}Z}\right] \tag{9}$$

$$= e^{-t}x\mathbf{E}[\nabla f(e^{-t}x + \sqrt{1-e^{-2t}}Z)] + e^{-2t}\mathbf{E}\left[\sum_{i=1}^n \frac{\partial^2 f(w)}{\partial w_i^2}\bigg|_{w=e^{-t}x+\sqrt{1-e^{-2t}}Z}\right] \tag{10}$$

$$= e^{-t}x\mathbf{E}[\nabla f(e^{-t}x + \sqrt{1-e^{-2t}}Z)] + e^{-2t}\mathbf{E}\left[\Delta f(e^{-t}x + \sqrt{1-e^{-2t}}Z)\right] \tag{11}$$

On the other hand,

$$LP_tf(x) = \Delta P_tf(x) + x \cdot \nabla P_tf(x) \tag{12}$$
$$= \Delta \mathbf{E}[f(e^{-t}x + \sqrt{1-e^{-2t}}Z)] + x \cdot \nabla \mathbf{E}[f(e^{-t}x + \sqrt{1-e^{-2t}}Z)] \tag{13}$$
$$= \mathbf{E}\left[\sum_{i=1}^n \frac{\partial^2}{\partial x_i^2} f(e^{-t}x + \sqrt{1-e^{-2t}}Z)\right] + \sum_{i=1}^n \mathbf{E}\left[x_i \frac{\partial}{\partial x_i} f(e^{-t}x + \sqrt{1-e^{-2t}}Z)\right] \tag{14}$$
$$= \mathbf{E}\left[\sum_{i=1}^n e^{-2t} f''(e^{-t}x + \sqrt{1-e^{-2t}}Z)\right] + \sum_{i=1}^n \mathbf{E}\left[e^{-t} x_i f'(e^{-t}x + \sqrt{1-e^{-2t}}Z)\right] \tag{15}$$
$$= e^{-2t}\mathbf{E}\left[\Delta f(e^{-t}x + \sqrt{1-e^{-2t}}Z)\right] + e^{-t}x\mathbf{E}[\nabla f(e^{-t}x + \sqrt{1-e^{-2t}}Z)] \tag{16}$$

which matches (11).

1.3 Ornstein-Uhlembeck process

The OU semigroup can also be written as

$$P_tf(x) = \mathbf{E}[f(X_t)]$$

where X_t is the *Ornstein-Uhlembeck process* with $X_0 = x$.

The OU process is defined by

$$X_t = e^{-t}X_0 + e^{-t}W_{e^{2t}-1}$$

where W_t is standard Brownian motion.

An easy consequence of the representation $P_tf(x) = \mathbf{E}[f(X_t)]$ is the following. For g positive,

$$|P_tf(x)|^2 \leq (P_tg(x))(P_t\frac{f^2}{g}(x)) \tag{17}$$

which follows from the Cauchy-Schwartz inequality,

$$|P_tf(x)|^2 = \mathbf{E}^2\left[\frac{f(X_t)}{\sqrt{g}(X_t)}\sqrt{g}(X_t)\right] \tag{18}$$
$$\leq \mathbf{E}\left[\frac{f^2(X_t)}{g(X_t)}\right]\mathbf{E}[g(X_t)] = (P_t\frac{f^2}{g}(x))(P_tg(x)) \tag{19}$$

However, as we will not use any other property of the OU process in the proof of hypercontractivity, we will refer the interested reader to Karatzas and Shreve (1991) for more details.

1.4 Proof of Hypercontractivity

The hypercontractive inequality for the OU semigroup was first proved by Nelson (1973): we now state his result.
Proposition. *Let P_t be the OU semigroup. For any $p > 1$, and $t > 0$, there exists a $q = q(t,p) > p$ such that for any $f \in L^2(\mu)$, the following holds:*

$$||P_t f||_{L^q(\mu)} \leq ||f||_{L^2(\mu)}$$

Furthermore, the above holds with $q(t,p) = 1 + (p-1)e^{2t}$.

We follow the proof in Chatterjee (2014), which makes use of the logarithmic Sobolev inequality for gaussian measures,
Lemma. (Gaussian log Sobolev inequality) *Let γ^n be the standard gaussian measure in $\mathbb{R}^n$. Then if $f : \mathbb{R}^n \to \mathbb{R}$ is an absolutely continuous function, then*

$$\gamma^n \left(f^2 \log \frac{f^2}{\gamma^n(f^2)} \right) \leq \gamma^n(2||\nabla f||_2^2)$$

where $\gamma^n(\cdot)$ denotes expectation wrt the measure γ^n.

Recall also the identity (1) which implies that $\gamma^n((f)(Lg)) = -\gamma^n(\nabla f \cdot \nabla g)$.
Proof of Lemma.

Begin by defining $v = f^2$. Then

$$\gamma^n \left(f^2 \log \frac{f^2}{\gamma^n(f^2)} \right) = \gamma^n \left(v \log \frac{v}{\gamma^n(v)} \right) \tag{20}$$

$$= \gamma^n(v \log v) - \gamma^n(v) \log \gamma^n(v) \tag{21}$$

$$= \gamma^n(v \log v) - \gamma^n(v \log \gamma^n(v)) \tag{22}$$

using the property (3)

$$= \gamma^n((P_0 v)(\log P_0 v)) - \gamma^n((P_\infty v)(\log P_\infty v)) \tag{23}$$
$$= -\int_0^\infty \frac{\partial}{\partial t}\gamma^n((P_t v)(\log P_t v))dt \tag{24}$$
$$= -\int_0^\infty \gamma^n\left((\frac{\partial}{\partial t}P_t v)(\log P_t v)\right) + \gamma^n\left((P_t v)(\frac{\partial}{\partial t}\log P_t v)\right)dt \tag{25}$$
$$= -\int_0^\infty \gamma^n\left((LP_t v)(\log P_t v)\right) + \gamma^n\left((P_t v)(\frac{LP_t v}{P_t v})\right)dt \tag{26}$$
$$= -\int_0^\infty \gamma^n((LP_t v)(1+\log P_t v))dt \tag{27}$$
$$= \int_0^\infty \mathcal{E}(P_t v, 1+\log P_t v)dt \tag{28}$$
$$= \int_0^\infty \gamma^n((\nabla P_t v)\cdot(\nabla \log P_t v))dt \tag{29}$$
$$= \int_0^\infty \gamma^n(\nabla P_t v \cdot \frac{\nabla P_t}{P_t v})dt \tag{30}$$
$$= \int_0^\infty \gamma^n\left(\frac{||\nabla P_t v||_2^2}{P_t v}\right)dt \tag{31}$$
$$= \int_0^\infty \gamma^n\left(\frac{e^{-2t}||P_t \nabla v||_2^2}{P_t v}\right)dt \tag{32}$$

using (17), which was a consequence of Cauchy-Schwartz,

$$\leq \int_0^\infty \gamma^n\left(\frac{e^{-2t}(P_t v)(P_t\frac{||\nabla v||^2}{v})}{P_t v}\right)dt \tag{33}$$
$$= \int_0^\infty \gamma^n\left(e^{-2t}P_t\frac{||\nabla v||^2}{v}\right)dt \tag{34}$$
$$= \int_0^\infty e^{-2t}\gamma^n\left(P_t\frac{||\nabla v||^2}{v}\right)dt \tag{35}$$

using (4),

$$= \int_0^\infty e^{-2t}\gamma^n\left(\frac{||\nabla v||^2}{v}\right)dt \tag{36}$$

$$= \frac{1}{2}\gamma^n\left(\frac{||\nabla v||^2}{v}\right) \tag{37}$$

and since $v = f^2$

$$\tag{38}$$

$$= \frac{1}{2}\gamma^n\left(\frac{||2f\nabla f||^2}{f^2}\right) \tag{39}$$

$$\leq \frac{1}{2}\gamma^n\left(\frac{4f^2||\nabla f||^2}{f^2}\right) \tag{40}$$

$$= 2\gamma^n(||\nabla f||_2^2) \tag{41}$$

as desired. $\square$.

Proof of Proposition

Define $f_t = P_t f$ and $r(t) = \gamma^n(f_t^{q(t)})$.

Since $q(0) = p$, we have

$$||P_0 f_t||_{q(0)} = ||f_t||_p$$

so it suffices to prove that

$$\frac{\partial}{\partial t}||P_t f||_{q(t)} \leq 0$$

Begin with the special case that f is nonnegative.

The key fact we will use is that applying the logarithmic Sobolev inquality to the function $f_t^{q(t)/2}$, we get

$$\gamma^n\left(f_t^{q(t)}\log\frac{f_t^{q(t)}}{r(t)}\right) \leq 2\gamma^n(|\nabla f_t^{q(t)/2}|^2) = 2\gamma^n\left(\left(\frac{q(t)}{2}\right)^2||\nabla f_t||_2^2\right) = \frac{q(t)^2}{2}\gamma^n(||\nabla f_t||_2^2) \tag{42}$$

Note that

$$q'(t) = 2(p-1)e^{-2t} = 2(1+(p-1)e^{-2t}-1) = 2(q(t)-1) \tag{43}$$

and

$$\frac{\partial}{\partial f_t} = Lf_t$$

Now

$$\begin{aligned}
r'(t) &= \gamma^n \left(f_t^{q(t)} \frac{\partial}{\partial t}(q(t)\log f_t) \right) \\
&= q'(t)\gamma^n(f_t^{q(t)} \log f_t) + q(t)\gamma^n(f_t^{q(t)-1} L f_t)
\end{aligned}$$

using (1)

$$\begin{aligned}
&= q'(t)\gamma^n(f_t^{q(t)} \log f_t) + q(t)\gamma^n(\nabla f_t^{q(t)-1} \cdot \nabla f_t) \\
&= q'(t)\gamma^n(f_t^{q(t)} \log f_t) + q(t)\gamma^n((q(t)-1) f_t^{q(t)-2} \nabla f_t \cdot \nabla f_t) \\
&= \frac{q'(t)}{q(t)}\gamma^n(f_t^{q(t)} \log f_t^{q(t)}) - q(t)(q(t)-1)\gamma^n(f_t^{q(t)-2} ||\nabla f_t||_2^2)
\end{aligned}$$

With this, write

$$\begin{aligned}
\frac{\partial}{\partial t} \log ||f_t||_{q(t)} &= \frac{\partial}{\partial t} \frac{\log r(t)}{q(t)} \\
&= \frac{-q'(t)\log r(t)}{q(t)^2} + \frac{r'(t)}{q(t)r(t)} \\
&= \frac{-q'(t)r(t)\log r(t)}{q(t)^2 r(t)} + \frac{q'(t)}{q(t)^2 r(t)}\gamma^n(f_t^{q(t)} \log f_t^{q(t)}) - \frac{q(t)-1}{r(t)}\gamma^n(f_t^{q(t)-2}||\nabla f_t||_2^2) \\
&= \frac{-q'(t)}{q(t)^2 r(t)}\gamma^n(f_t^{q(t)} \log r(t)) + \frac{q'(t)}{q(t)^2 r(t)}\gamma^n(f_t^{q(t)} \log f_t^{q(t)}) - \frac{q(t)-1}{r(t)}\gamma^n(f_t^{q(t)-2}||\nabla f_t||_2^2) \\
&= \frac{q'(t)}{q(t)^2 r(t)}\gamma^n \left(f_t^{q(t)} \log \frac{f_t^{q(t)}}{r(t)} \right) - \frac{q(t)-1}{r(t)}\gamma^n(f_t^{q(t)-2}||\nabla f_t||_2^2)
\end{aligned}$$

using (43)

$$\begin{aligned}
&= \frac{q'(t)}{q(t)^2 r(t)}\gamma^n \left(f_t^{q(t)} \log \frac{f_t^{q(t)}}{r(t)} \right) - \frac{q'(t)}{2r(t)}\gamma^n(f_t^{q(t)-2}||\nabla f_t||_2^2) \\
&= \frac{q'(t)}{q(t)^2 r(t)} \left[\gamma^n \left(f_t^{q(t)} \log \frac{f_t^{q(t)}}{r(t)} \right) - \frac{q(t)^2}{2}\gamma^n(f_t^{q(t)-2}||\nabla f_t||_2^2) \right]
\end{aligned}$$

but from the logarithmic Sobolev inequality (42), we know that the expression in the brackets is at most zero

$$\leq 0$$

The general case follows from

$$||P_t f||_{q(t)} = |||P_t f|||_{q(t)} \leq |||P_t|f|||_{q(t)} = ||P_t|f|||_{q(t)} \leq |||f|||_p = ||f||_p$$

2 References

Chatterjee, S. (2014). *Superconcentration and Related Topics*. Springer.
Ledoux, M. (2005). *The concentration of measure phenomenon* (Vol. 89). American Mathematical Soc..