Probability in High Dimensions

Solutions to Exercises, Week 4

August 10, 2020

2.3 Poincaré Inequalities

Problem 2.8 (Exponential distribution). *Let $\mu(dx) = \mathbf{1}_{x\geq 0}e^{-x}dx$ be the one-sided exponential distribution. In this problem, we will derive two different (and not directly comparable) Poincaré inequalities for the distribution μ.*

a. *Show that*
$$\mathrm{Var}_\mu[f] \leq 2\,\mathbf{E}[\xi|f'(\xi)|^2], \qquad \xi \sim \mu.$$

b. *Show that*
$$\mathrm{Var}_\mu[f] \leq 4\,\mathbf{E}[|f'(\xi)|^2], \qquad \xi \sim \mu.$$

c. *Show that the Dirichlet form of the solution of the Cox-Ingersoll-Ross stochastic differential equation*
$$dX_t = 2(1-X_t)dt + 2\sqrt{X_t}dB_t,$$
which is a Markov process on $\mathbb{R}_+$ with generator
$$\mathcal{L}f(x) = 2(1-x)f'(x) + 2xf''(x),$$
yields the Poincaré inequality from part (a).

d. *Show that the Dirichlet form of the solution of the stochastic differential equation*
$$dX_t = -sign(X_t)dt + \sqrt{2}dB_t,$$
which is a Markov process on $\mathbb{R}$ with generator
$$\mathcal{L}f(x) = -sign(x)f'(x) + f''(x),$$
yields the Poincaré inequality from part (b).

Solution. a. By a simple polar coordinates argument the hint follows, that is $\xi \sim (X^2+Y^2)/2$ with X, Y i.i.d. $N(0,1)$.

Let
$$g(x,y) = f\left(\frac{x^2+y^2}{2}\right).$$

Then
$$\partial_x g(x,y) = f'\left(\frac{x^2+y^2}{2}\right)x, \quad \partial_y g(x,y) = f'\left(\frac{x^2+y^2}{2}\right)y.$$

Apply the Gaussian Poincare inequality we get

$$\begin{aligned}\mathrm{var}_\mu(f) &= \mathrm{var}\left[g(X,Y)\right]\\ &\le \mathbb{E}\left[\|\nabla g(X,Y)\|^2\right]\\ &= \mathbb{E}\left[|f'(\xi)|^2\left(x^2+y^2\right)\right]\\ &= 2\mathbb{E}\left[|f'(\xi)|^2\xi\right].\end{aligned}$$

b. Notice that $g(0)=0$. Then

$$\begin{aligned}\mathrm{Var}_\mu[f] &= \mathrm{Var}_\mu[f-f(0)]\\ &\le \mathbb{E}\left[\left(f(\xi)-f(0)\right)^2\right]\\ &= \mathbb{E}\left[2f'(\xi)\left(f(\xi)-f(0)\right)\right]\end{aligned}$$

and by Cauchy-Schwarz

$$\le 2\mathbb{E}\left[|f'(\xi)|^2\right]^{1/2}\mathbb{E}\left[\left(f(\xi)-f(0)\right)^2\right]^{1/2}.$$

Therefore

$$\mathrm{Var}_\mu[f] \le \mathbb{E}\left[\left(f-f(0)\right)^2\right] \le 4\mathbb{E}\left[|f'(\xi)|^2\right],$$

and the result follows.

c. Let $g(x)=xf(x)f'(x)$. Then

$$g'(x)=f(x)f'(x)+xf'(x)^2+xf(x)f''(x).$$

Thus

$$\begin{aligned}\int \mathrm{e}^{-x}xf(x)f'(x)\mathrm{d}x &= \int[-\mathrm{e}^{-x}]'xf(x)f'(x)\mathrm{d}x\\ &= \int \mathrm{e}^{-x}[xf(x)f'(x)]'\mathrm{d}x\\ &= \int \mathrm{e}^{-x}[ff'+xf'f'+xff'']\mathrm{d}x.\end{aligned}$$

Therefore

$$\begin{aligned}\langle f,\mathcal{L}f\rangle &= 2\int \mathrm{e}^{-x}f(x)\left[f'(x)-xf'(x)+xf''(x)\right]\mathrm{d}x\\ &= 2\int \mathrm{e}^{-x}\left[f(x)f'(x)-xf(x)f'(x)+xf(x)f''(x)\right]\mathrm{d}x\\ &= 2\int \mathrm{e}^{-x}\left[xf(x)f'(x)-xf'(x)f'(x)-xf(x)f''(x)-xf(x)f'(x)+xf(x)f''(x)\right]\mathrm{d}x\\ &= 2\int \mathrm{e}^{-x}\left[-xf'(x)f'(x)\right]\mathrm{d}x\\ &= -2\mathbb{E}[\xi|f'(\xi)|^2].\end{aligned}$$

d. Let $f : \mathbb{R} \to \mathbb{R}$. Then

$$\langle f, \mathcal{L} f \rangle = \frac{1}{2} \int_{x=-\infty}^{\infty} \mathrm{e}^{-|x|} f(x)[-\operatorname{sign}(x) f'(x) + f''(x)] \mathrm{d}x.$$

Then

$$\begin{aligned} &\frac{1}{2} \int_{x=-\infty}^{\infty} \mathrm{e}^{-|x|} f(x) f''(x) \mathrm{d}x \\ &= -\frac{1}{2} \int_{x=-\infty}^{\infty} \left[\mathrm{e}^{-|x|} f(x) \right]' f'(x) \mathrm{d}x \\ &= -\frac{1}{2} \int_{x=-\infty}^{\infty} \mathrm{e}^{-|x|} \left[-\operatorname{sign}(x) f(x) + f'(x) \right] f'(x) \mathrm{d}x, \end{aligned}$$

and thus

$$\begin{aligned} \langle f, \mathcal{L} f \rangle &= -\frac{1}{2} \int_{x=-\infty}^{\infty} \mathrm{e}^{-|x|} f(x) \operatorname{sign}(x) f'(x) \mathrm{d}x + \frac{1}{2} \int_{x=-\infty}^{\infty} \mathrm{e}^{-|x|} \operatorname{sign}(x) f(x) f'(x) \mathrm{d}x \\ &\quad - \frac{1}{2} \int_{x=-\infty}^{\infty} \mathrm{e}^{-|x|} [f'(x)]^2 \mathrm{d}x \\ &= -\frac{1}{2} \int_{x=-\infty}^{\infty} \mathrm{e}^{-|x|} [f'(x)]^2 \mathrm{d}x. \end{aligned}$$

Problem 2.9 (Dependent random signs)**.** *Question is too long and is written in the notes.*

Solution. a.

b. For an n-vector $x = (x_1, \ldots, x_n)$, write x^{-i} for the $n-1$-vector $(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$. For small enough $h > 0$ and bounded continuous f we have

$$\begin{aligned} P_h f(x) &= \mathbb{E}\left[f(X_h) \mid X_0 = x \right] \\ &= \mathbb{E}\left[f(X_h) \mathbb{1}\{\text{no jump}\} \mid X_0 = x \right] \\ &\quad + \mathbb{E}\left[f(X_h) \mathbb{1}\{\text{one jump}\} \mid X_0 = x \right] + \mathbb{E}\left[f(X_h) \mathbb{1}\{\text{more than one jump}\} \mid X_0 = x \right] \\ &= f(x) \mathrm{e}^{-nh} + \mathrm{e}^{-nh} nh \frac{1}{n} \sum_{i=1}^{n} \int f\left(x_1, \ldots, x_{i-1}, z_i, x_{i+1}, \ldots, x_n \right) \mu(\mathrm{d}z_i \mid x^{-i}) + o(h). \end{aligned}$$

Therefore

$$\begin{aligned} \frac{1}{h} \left[P_h f(x) - f(x) \right] &= f(x) \frac{\mathrm{e}^{-nh} - 1}{h} + \frac{1}{n} \mathrm{e}^{-nh} \sum_{i=1}^{n} \int f(x) \mu(\mathrm{d}x_i \mid x) + o(1) \\ &= \sum \delta_i f + o(1). \end{aligned}$$

For the Dirichlet form notice that

$$\begin{aligned} &\langle f, \delta_i g \rangle \\ &= \int \mu(\mathrm{d}x_1, \ldots, \mathrm{d}x_n) f(x_1, \ldots, x_n) \int g(x_1, \ldots, x_{i-1}, z_i, x_i, \ldots, x_n) \mu(\mathrm{d}z_i \mid x^{-i}) \\ &= \int \mu(\mathrm{d}x^{-i}) \int \mu(\mathrm{d}y_i \mid x^{-i}) \int \mu(\mathrm{d}z_i \mid x^{-i}) f(x_1, \ldots, x_{i-1}, y_i, x_i, \ldots, x_n) g(x_1, \ldots, x_{i-1}, z_i, x_i, \ldots, x_n) \\ &= \langle \delta_i f, \delta_i g \rangle = \langle \delta_i f, g \rangle. \end{aligned}$$

c. Say first that $i < j$. Given a vector x, let $x^{j,\pm 1}$ be the vector that agrees with x at all entries except perhaps at the j-th entry where $(x^{j,\pm 1})_j := \pm 1$.

$$
\begin{aligned}
&\Delta_j \int f \mathrm{d}\mu_i \\
&= \max_{x\in\{\pm 1\}^n} \Big| \int f(x_1,\dots,x_{i-1},z_i,x_{i+1},\dots,x_{j-1},1,x_{j+1}\dots)\mu(\mathrm{d}z_i \mid x^{j,+1}) \\
&\qquad - \int f(x_1,\dots,x_{i-1},z_i,x_{i+1},\dots,x_{j-1},-1,x_{j+1}\dots)\mu(\mathrm{d}z_i \mid x^{j,-1}) \Big| \\
&= \max_{x\in\{\pm 1\}^n} \Big| \Big[\int f(x_1,\dots,x_{i-1},z_i,x_{i+1},\dots,x_{j-1},1,x_{j+1}\dots)\mu(\mathrm{d}z_i \mid x^{j,+1}) \\
&\qquad\qquad - \int f(x_1,\dots,x_{i-1},z_i,x_{i+1},\dots,x_{j-1},-1,x_{j+1}\dots)\mu(\mathrm{d}z_i \mid x^{j,+1}) \Big] \\
&\qquad - \Big[\int f(x_1,\dots,x_{i-1},z_i,x_{i+1},\dots,x_{j-1},-1,x_{j+1}\dots)\mu(\mathrm{d}z_i \mid x^{j,-1}) \Big| \\
&\qquad\qquad - \int f(x_1,\dots,x_{i-1},z_i,x_{i+1},\dots,x_{j-1},-1,x_{j+1}\dots)\mu(\mathrm{d}z_i \mid x^{j,+1}) \Big] \\
&\le \int \mu(\mathrm{d}z_i \mid x^{j,+1})\, |\Delta_j f| \\
&+ \max_{x\in\{\pm 1\}^n} \Big| f(x_1,\dots,x_{i-1},+1,x_{i+1},\dots,x_{j-1},-1,x_{j+1}\dots)\left[\mu(\mathrm{d}z_i = +1 \mid x^{j,+1}) - \mu(\mathrm{d}z_i = +1 \mid x^{j,-1})\right] \\
&\qquad - f(x_1,\dots,x_{i-1},-1,x_{i+1},\dots,x_{j-1},-1,x_{j+1}\dots) \times \left[\mu(\mathrm{d}z_i = +1 \mid x^{j,+1}) - \mu(\mathrm{d}z_i = +1 \mid x^{j,-1})\right] \Big| \\
&\le \Delta_j f \\
&+ \max_{x\in\{\pm 1\}^n} \Big| [f(x_1,\dots,x_{i-1},+1,x_{i+1},\dots,x_{j-1},-1,x_{j+1}\dots) - f(x_1,\dots,x_{i-1},-1,x_{i+1},\dots,x_{j-1},-1,x_{j+1}\dots)] \\
&\qquad \max_{x\in\{\pm 1\}^n} \left[\mu(\mathrm{d}z_i = +1 \mid x^{j,+1}) - \mu(\mathrm{d}z_i = +1 \mid x^{j,-1})\right] \\
&\le \Delta_j f + \Delta_i f C_{ij}.
\end{aligned}
$$

d. Since $\mathcal{L}f = \lim h^{-1}(P_h f - f)$ we roughly have that $P_h f = f + h\mathcal{L}f + o(h)$. Then applying

$$
\begin{aligned}
\Delta_j \left(f + \frac{t}{m}\mathcal{L}f \right) &= \Delta_j f + \frac{t}{m}\Delta_j \frac{P_{t/m}f - f}{t/m} + o\left(\frac{t}{m}\right) \\
&= \Delta_j P_{t/m} f + o\left(\frac{t}{m}\right) \\
&= \mathrm{e}^{-t/m}\Delta_j f + \frac{t}{m}\sum_{i=1}^n \Delta_j \left(\int f \mathrm{d}\mu_i - f \right) + o\left(\frac{t}{m}\right)
\end{aligned}
$$

and using the inequality from the last part

$$
\le \left(1 - \frac{t}{m}\right)\Delta_j f + \sum_{i=1}^n \Delta_i f C_{ij} + o\left(\frac{t}{m}\right).
$$

e. Then using the previous result in vector notation

$$
\Delta P_t f = \lim_{m\to\infty} \Delta(I + t\mathcal{L}t/m)^m f
$$

$$
\begin{aligned}
&= \lim_{m\to\infty} \Delta(I+t\mathcal{L}t/m)\left[(I+t\mathcal{L}t/m)^{m-1}f\right] \\
&\le \lim_{m\to\infty} \Delta\left[(I+t\mathcal{L}t/m)^{m-1}f\right][I-t(I-C)/m] \\
&\le \lim_{m\to\infty} \Delta\left[(I+t\mathcal{L}t/m)(I+t\mathcal{L}t/m)^{m-2}f\right][I-t(I-C)/m] \\
&\le \lim_{m\to\infty} \Delta\left[(I+t\mathcal{L}t/m)^{m-2}f\right][I-t(I-C)/m]^2 \\
&\le \cdots \le \lim_{m\to\infty} \Delta f\,[I-t(I-C)/m]^m = \Delta f \mathrm{e}^{-t(I-C)}.
\end{aligned}
$$

f. It is sufficient to get

$$\mathcal{E}(P_tf,P_tf) \le \kappa(f)e^{2t(1-\|C\|_{sp})}$$

with $\kappa(f)=\Delta f\cdot(\Delta f)^T$, and then use the implication suggested in the question.

Using (b) we have $\mathcal{E}(P_tf,P_tf)=\sum_i \mu\left((\delta_i P_t f)^2\right)$ which is less than $\Delta P_t f(\Delta P_t f)^T$; this is in turn less than $\Delta f e^{-t(I-C)}$ by (e). Hence we have

$$
\begin{aligned}
\mathcal{E}(P_tf,P_tf) \le \Delta f e^{-t(I-C)}e^{-t(I-C)^T}(\Delta f)^T &\le \Delta f e^{-t(I-C)}e^{-t(I-C)^T}(\Delta f)^T \\
&\le \Delta f(\Delta f)^T\|e^{-t(I-C)}e^{-t(I-C)^T}\|_{sp}.
\end{aligned}
$$

First note that the eigenvalues of e^A are the exponentials of the eigenvalues of A, and that the eigenvalues of e^{-A} are the multiplicative inverses of the eigenvalues of e^A. Hence

$$\lambda_{\max}(e^{-A}) = \lambda_{\min}(e^A) = e^{\lambda_{\min}A} = e^{-\lambda_{\max}(-A)}.$$

In particular, $\lambda_{\max}(e^{-t(I-C)}) = \exp\{-t\lambda_{\max}(C-I)\} = \exp\{-t(\|C\|_{sp}-1)\} = e^{t(1-\|C\|_{sp})}$. Then using that $\|AB\|_{sp}\le\|A\|_{sp}\|B\|_{sp}$ (which can be seen directly from the definition of eigenvalues), we have the required result.