Paco Goze
TP n°1

Reminder on Markov chains – Stochastic Gradient Descent

Exercise 1.

① Let $h : \mathbb{R}^2 \to \mathbb{R}$ a continuous bounded function

$$\mathbb{E}(h(X,Y)) = \mathbb{E}(h(R\cos(\Theta), R\sin(\Theta)))$$
$$= \int_{\mathbb{R}_+} \int_0^{2\pi} h(r\cos(\theta), r\sin(\theta))\, r e^{-\frac{r^2}{2}} \frac{1}{2\pi}\, dr\, d\theta$$

Let $x = r\cos(\theta)$, $y = r\sin(\theta)$, $J_{\theta,r} = \begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix}$

We have $|J_{\theta,r}|^{-1} = \frac{1}{r}$

Therefore: $$\mathbb{E}(h(X,Y)) = \int_{\mathbb{R}} \int_{\mathbb{R}} h(x,y)\, e^{-\frac{x^2+y^2}{2}} \frac{1}{2\pi}\, dx\, dy$$
$$= \int_{\mathbb{R}} \int_{\mathbb{R}} \underbrace{\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}}_{f_X(x)} \underbrace{\frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}}_{f_Y(y)}\, dx\, dy$$

We have: X and Y are independent, each follow a $\mathcal{N}(0,1)$ distribution.

② On peut donc simuler deux variables aléatoires indépendantes $\mathcal{N}(0,1)$ comme suit :

1- Simuler une loi de Rayleigh (1), R
Simuler une loi $\mathcal{U}[0,2\pi]$, Θ

2- Poser $X = R\cos(\Theta)$ et $Y = R\sin(\Theta)$

D'après ①, on obtient bien deux normales centrées réduites indépendantes.

En particulier à l'étape 1, on simule une Rayleigh(1) en utilisant l'inverse de la cdf, et une $U([0,1])$:

$F(r) = \mathbb{P}(R \leq r) = \int_0^r u e^{-u^2/2} du = 1 - e^{-r^2/2}$

Donc :

$F^{\leftarrow} : r \mapsto \sqrt{-2\log(1-r)}, \quad r \in]0,1[$

Si $U \sim \mathcal{U}([0,1])$, $F^{\leftarrow}(U) \sim \mathcal{R}(1)$ (Rayleigh(1)).

L'algorithme est donc le suivant :

Sampler $U_1 \sim U([0,1])$, $U_2 \sim U([0,1])$

Poser $\Theta = 2\pi U_1$ et $R = F^{\leftarrow}(U_2)$

Renvoyer $X = R\cos(\Theta)$, $Y = R\sin(\Theta)$

③ a) À la fin de chaque itération dans la boucle while, on a :

$(V_1, V_2) \sim \mathcal{U}([-1,1] \times [-1,1])$

C'est à dire $(V_1, V_2) \sim \mathcal{U}(B_{0,1}^{\|\cdot\|_\infty})$ avec (*)

$B_{0,1}^{\|\cdot\|_\infty}$ la boule centrée en O de rayon 1 pour la norme $\|\cdot\|_\infty$

On sort de la boucle while lorsque $\|(V_1,V_2)\|_2 \leq 1$

D'après (*),

$(V_1, V_2) \sim \mathcal{U}(B_{0,1}^{\|\cdot\|_2})$

b) À chaque passage dans la boucle, la probabilité de sortie est : $p = \mathcal{A}(B_{0,1}^{\|\cdot\|_2}) / \mathcal{A}(B_{0,1}^{\|\cdot\|_\infty})$

$= \pi/4$

En posant G la variable aléatoire qui compte le nombre de passages dans la boucle while : $G \sim \mathcal{G}(p)$

Donc : $\mathbb{E}(G) = 1/p = 4/\pi$; le nombre moyen de passages dans la boucle while est $\frac{4}{\pi}$

c) Ainsi, $(V_1, V_2) \sim \mathcal{U}(D)$, D étant le disque unité pour $\|\cdot\|_2$

$T_1 = \frac{V_1}{\sqrt{V_1^2+V_2^2}}$, $V = V_1^2 + V_2^2$

Soit $\varphi : D \to \mathbb{R}^2$
$$(v_1, v_2) \mapsto \left(\frac{v_1}{\sqrt{v_1^2+v_2^2}}, v_1^2+v_2^2 \right)$$

$\varphi(v_1, v_2) = (t_1, v) \Leftrightarrow \begin{cases} v_1 = t_1\sqrt{v} \\ v_2 = \sqrt{v(1-t_1^2)} \end{cases}$ est bijective

Donc :
$\varphi^{\leftarrow} : (t_1, v) \mapsto (t_1\sqrt{v}, \sqrt{v(1-t_1^2)})$ est bien bijective

Soit h une fonction continue bornée.
$\mathbb{E}(h(T_1, V)) = \mathbb{E}(h \circ \varphi(V_1, V_2))$

$$= \iint_D \frac{1}{\pi} h \circ \varphi(v_1, v_2)\, dv_1\, dv_2 \quad (*)$$

On a $J_{\varphi^{\leftarrow}}(t_1, v) = \begin{pmatrix} \sqrt{v} & t_1/(2\sqrt{v}) \\ -t_1\sqrt{\frac{v}{1-t_1^2}} & \frac{1}{2}\sqrt{\frac{1-t_1^2}{v}} \end{pmatrix}$

Donc $|J_{\varphi^{\leftarrow}}(t_1, v)| = \frac{1}{2}(1-t_1^2)^{-1/2}$

$$(*) = \iint_{[-1,1]\times[0,1]} \frac{1}{\pi} h(t_1, v) \frac{1}{2}(1-t_1^2)^{-1/2}\, dt_1\, dv$$

Donc :

$$\mathbb{E}(h(T_1,V)) = \iint_{[-1,1]\times[0,1]} h(t_1,v)\, \frac{1}{2\pi}\, (1-t_1^2)^{-1/2}\, dt_1\, dv$$

On a bien T_1 et V indépendantes

On a $V \sim U([0,1])$ et $f_{T_1}(t_1) = \frac{1}{2\pi}\, \frac{1}{\sqrt{1-t_1^2}}$

$$\mathbb{E}(h(\arccos(T_1))) = \int_{-1}^{1} h(\arccos(t_1))\, \frac{1}{2\pi}\, \frac{1}{\sqrt{1-t_1^2}}\, dt_1$$

$$= \int_0^{2\pi} \frac{h(\theta)}{2\pi}\, d\theta \quad \text{en posant } t_1 = \cos(\theta)$$

Donc : $\arccos(T_1) \sim U([0,2\pi])$

Donc : $T_1 \sim \cos(\Theta)$, avec $\Theta \sim U([0,2\pi])$

Ainsi : $\boxed{T_1 \perp\!\!\!\perp V, \quad (T_1, V) \sim (\cos(\Theta), U([0,1])) \text{ avec } \Theta \sim U([0,2\pi])}$

d) $V \sim U([0,1])$ donc $S \sim \text{Rayleigh}(1)$

On montre par un raisonnement analogue que

$T_2 = \frac{V_2}{\sqrt{V_1^2+V_2^2}} \sim \sin(\Theta)$ avec $\Theta \sim U([0,2\pi])$

et que $T_2 \perp\!\!\!\perp V$.

Donc $ST_1 \sim R\cos(\Theta)$, avec $R \sim \text{Rayleigh}(1)$, $\Theta \sim U([0,2\pi])$

$ST_2 \sim R\sin(\Theta)$,

D'après (1) : $X \sim \mathcal{N}(0,1)$, $Y \sim \mathcal{N}(0,1)$ et $X \perp\!\!\!\perp Y$.

Exercice 2.

① On a $P(x,A) = \mathbb{P}(X_{n+1} \in A \mid X_n = x)$

Si $x \neq 1/k$, pour $k \in \mathbb{N}^*$,
Alors $X_{n+1} \sim \mathcal{U}([0,1])$
Donc $P(x,A) = \mu(A) = \int_{A \cap [0,1]} dt$

Si $x = 1/k$, pour $k \in \mathbb{N}^*$
$\mathbb{P}(X_{n+1} \in A \mid X_n = x) = \mathbb{P}(X_{n+1} \in A \mid X_n = x, X_{n+1} = \frac{1}{k+1})\mathbb{P}(X_{n+1} = \frac{1}{k+1})$
$+ \mathbb{P}(X_{n+1} \in A \mid X_n = x, X_{n+1} \sim \mathcal{U}([0,1]))\mathbb{P}(X_{n+1} \sim \mathcal{U}([0,1]))$
$= \delta_{\frac{1}{k+1}}(A)(1-x^2) + \int_{A \cap [0,1]} dt \, x^2$

Donc: $\boxed{P(x,A) = \left(x^2 \int_{A \cap [0,1]} dt + (1-x^2)\delta_{\frac{1}{k+1}}(A)\right)\mathbb{1}_{x \in \frac{1}{\mathbb{N}^*}} + \int_{A \cap [0,1]} dt \, \mathbb{1}_{x \notin \frac{1}{\mathbb{N}^*}}}$

② Soit A un borélien de $[0,1]$

$$\pi P(A) = \int_{[0,1]} P(x,A)\,dx$$
$$= \int_{[0,1]} \left(x^2 \int_{A \cap [0,1]} dt + (1-x^2)\delta_{\frac{1}{k+1}}(A)\right)\mathbb{1}_{x \in \frac{1}{\mathbb{N}^*}} dx$$
$$+ \int_{[0,1]} \left(\int_{A \cap [0,1]} dt\right)\mathbb{1}_{x \notin 1/\mathbb{N}^*} dx$$
$$= \underbrace{\int_{[0,1] \cap \frac{1}{\mathbb{N}^*}} \left(x^2 \int_{A \cap [0,1]} dt + (1-x^2)\delta_{\frac{1}{1+k}}(A)\right) dx}_{=0}$$
$$+ \int_{[0,1] \setminus \frac{1}{\mathbb{N}^*}} \mu(A)\,dx$$
$$= \mu(A)$$
$$= \pi(A)$$

Ainsi : $\Pi P(A) = \Pi(A)$; la distribution uniforme sur $[0,1]$ est invariante par P.

③ $Pf(x) = \mathbb{E}(f(X_1) \mid X_0 = x) = \mathbb{E}(f(U) \mid U \sim \mathcal{U}([0,1]))$
$$= \int_0^1 f(t)\,\pi(t)\,dt$$

On a par récurrence immédiate que si $x_0 \notin 1/\mathbb{N}^*$, alors pour tout $n \geq 1$, $X_n \sim \mathcal{U}([0,1])$.

Donc $P^n f(x) = \mathbb{E}(f(X_n) \mid X_0 = x) = \mathbb{E}(f(X_n) \mid X_n \sim \mathcal{U}([0,1]))$
$$= \int_0^1 f(t)\,\pi(t)\,dt.$$

Donc, $\lim\limits_{n \to +\infty} P^n f(x) = \int_0^1 f(t)\,\pi(t)\,dt$

④ a) $P^n\left(x, \frac{1}{n+k}\right) = \mathbb{P}\left(X_n = \frac{1}{n+k} \mid X_0 = \frac{1}{k}\right)$

$$= \mathbb{P}\left(X_n = \frac{1}{n+k} \mid X_{n-1} = \frac{1}{n-1+k}\right) \cdots \mathbb{P}\left(X_1 = \frac{1}{k+1} \mid X_0 = \frac{1}{k}\right)$$

$$= \prod_{i=0}^{n-1} \left(1 - \frac{1}{(k+i)^2}\right)$$

Si $n \neq m$, $P^n\left(x, \frac{1}{m+k}\right) = 0$.

b) On pose $A = \bigcup\limits_{q \in \mathbb{N}} \left\{\frac{1}{k+1+q}\right\}$: $\Pi(A) = 0$

$$P^n(x, A) = \sum_{q \in \mathbb{N}} P^n(x, 1/k+1+q) = P^n(x, 1/n+k)$$
$$= \prod_{i=0}^{n-1} \left[1 - \left(\frac{1}{k+i}\right)^2\right]$$

La série des $\frac{1}{n^2}$ converge, ce produit ne tend donc pas vers 0 quand n tend vers $+\infty$ (passer au log).

Ainsi, $\boxed{\lim\limits_{n \to +\infty} P^n(x, A) > \Pi(A)}$