Committor function-report 1

September 17, 2020

1 Thread

Let us consider a Markov process $\{x_t\}$ with transition density $p(x, y) = \mathbb{P}(x_{t+\tau} = y \mid x_t = x)$. For two sets A, B, the committor function $q(x)$ is defined as the probability of hitting region B before region A with the process starting at x. By considering the timescale $\tau, q(x)$ satisfies

$$q(x) = \begin{cases} 0, & x \in A \\ 1, & x \in B \\ \tilde{q}(x), & \text{otherwise} \end{cases}$$

where $\tilde{q}(x) = \int p(x, y) q(y) \mathrm{d}y$

$$\begin{aligned} & A, B \subset \Omega, A \cap B = \emptyset \\ & \tau_A(x) = \inf\{t \geq 0 : X_t \in A, X_0 = x\} \\ & \tau_B(x) = \inf\{t \geq 0 : X_t \in B, X_0 = x\} \\ & q(x) = \mathbb{P}\{\tau_B(x) < \tau_A(x)\},\ \Omega \to [0, 1] \\ & q(x) = 0, x \in \partial A \cup A,\ q(x) = 1, x \in \partial B \cup B \end{aligned}$$

Aim: Find an algorithm of Computing $q(x)$.
Consider

$$\begin{aligned} & q(x) = \int p(x, y) q(y) dy,\ x \in \Omega\ (A \cup B), y \in \Omega\ (A \cup B) \\ & \rho_0(x) q(x) = \int \rho_0(x) p(x, y) q(y) dy \\ & \rho_0(x) > 0, x \in \Omega, x \sim \rho_0(x) \end{aligned}$$

1.1 Sampling

$$\rho_0(x) q(x) = \int \rho_0(x) p(x, y) q(y) dy \tag{1}$$

Assume that there are N transition pairs $\{(x_n, y_n)\}_{n=1}^N$ drawn from $\rho_0(x)p(x,y)$ and $\rho_0(x) = 0$ for $x \in A \cup B$ without lack of generality [1]

1.2 WGAN

Here we introduce an instrumental variable $z \in [0,1]$, and define

$$\mathbb{Q}(x,z) = 2\rho_0(x)(z \cdot q(x) + (1-z) \cdot (1-q(x)))$$
$$\tilde{\mathbb{Q}}(x,z) = 2\rho_0(x)(z \cdot \tilde{q}(x) + (1-z) \cdot (1-\tilde{q}(x)))$$

It can be seen that $\mathbb{Q}(x,z) = \tilde{\mathbb{Q}}(x,z)$ if $q(x)$ is exactly the committor function and

$$\iint \mathbb{Q}(x,z)\mathrm{d}x\mathrm{d}z = \iint \tilde{\mathbb{Q}}(x,z)\mathrm{d}x\mathrm{d}z = 1$$

Hence, we can optimize the parametric model of q by minimizing the Wasserstein1 distance

$$W(\mathbb{Q}, \tilde{\mathbb{Q}}) = \sup_{\|f\|_L \le 1} \mathbb{E}_{\mathbb{Q}}[f(x,z)] - \mathbb{E}_{\tilde{\mathbb{Q}}}[f(x,z)]$$

where

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}[f(x,z)] &= \mathbb{E}_{x\sim\rho_0(x), z\sim\mathcal{U}_{[0,1]}}[2(z \cdot q(x) + (1-z) \cdot (1-q(x)))f(x,z)] \\
&\approx \frac{1}{N}\sum_n \mathbb{E}_{z\sim\mathcal{U}_{[0,1]}}\left[2\left(z \cdot q\left(x_n\right) + (1-z) \cdot \left(1-q\left(x_n\right)\right)\right) f\left(x_n, z\right)\right] \\
\mathbb{E}_{\tilde{\mathbb{Q}}}[f(x,z)] &= \mathbb{E}_{x\sim\rho_0(x), z\sim\mathcal{U}_{[0,1]}}[2(z \cdot \tilde{q}(x) + (1-z) \cdot (1-\tilde{q}(x)))f(x,z)] \\
&= \iiint 2\rho_0(x)p(x,y)(z \cdot q(y) + (1-z) \cdot (1-q(y)))f(x,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z \\
&\approx \frac{1}{N}\sum_n \mathbb{E}_{z\sim\mathcal{U}_{[0,1]}}\left[2\left(z \cdot q\left(y_n\right) + (1-z) \cdot \left(1-q\left(y_n\right)\right)\right) f\left(x_n, z\right)\right]
\end{aligned}$$

[1] In applications, the latter assumption can be satisfied by simply removing all transition pairs (x_n, y_n) with $x_n \in A \cup B$ from the data set.

1.3 WGAN-GP

Consider

$$\min_q \max_{\|f\|_{Lip}\le 1} \mathbb{E}_{x,z}[2(z\cdot q(x)+(1-z)\dot{(}1-q(x)))f(x,z)] - \mathbb{E}_{x,z}[2(z\cdot \tilde{q}(x)+(1-z)\dot{(}1-\tilde{q}(x)))f(x,z)]+$$
$$\lambda \mathbb{E}_{x,z}\left[(\|\nabla_x f(x,z)\|_2 - 1\|)^2\right]$$
$$\text{sample}\{x^{(n)}, y^{(n)}\}_1^N \sim p(x,y)$$
$$\min_q \max_{\|f\|_{Lip}\le 1} \frac{1}{N}\sum_n \mathbb{E}_{z\sim U[0,1]}[2(z\cdot q(x_n)+(1-z)\dot{(}1-q(x_n)))f(x_n,z)] -$$
$$\frac{1}{N}\sum_n \mathbb{E}_{z\sim U[0,1]}[2(z\cdot q(y_n)+(1-z)\dot{(}1-q(y_n)))f(x_n,z)] + \lambda\frac{1}{N}\mathbb{E}_{z\sim U[0,1]}\left[(\|\nabla_{x_n} f(x_n,z)\|_2 - 1\|)^2\right]$$
$$\text{sample}\{z^{(m)}\}_1^M \sim U[0,1]$$
$$\min_q \max_{\|f\|_{Lip}\le 1} \frac{2}{MN}\sum_N\sum_M [(2z_m-1)(q(x_n)-q(y_n))f(x_n,z_m)$$
$$\lambda\left[(\|\nabla_x f(x_n,z_m)\|_2 - 1\|)^2\right]$$

1.4 Algorithm

Aim: $\theta : q_\theta, w : f_w$

Algorithm 1 Committor Function

Input: gradient penalty coefficient λ; Number of critic iterations per generator iteration n_{critic}; batch size N; Adam hyperparameters α, β_1, β_2; initial critic parameters ω_0; initial generator parameters θ_0
while θ has not converged **do**
    for $t = 1, \cdots, n_{critic}$ **do**
        for $n = 1, \cdots, N$ **do**
            Sample real date $(x_n, y_n) \sim \rho_0(x) * p(x,y)$
            for $m = 1, \cdots, M$ **do**
                Sample $z_m \sim U[0,1]$
                $L^{(n,m)} \leftarrow (2z_m-1)(q(x_n)-q(y_n))f(x_n,z_m)+\lambda(\|\nabla_x f(x,z)\|_2-1\|)^2$
            end for
        end for
        $\omega \leftarrow Adam(\nabla_\omega \frac{2}{N*M}\sum_n\sum_m L^{(n,m)}, \omega, \alpha, \beta_1, \beta_2)$
    end for
    Sample $(x_n, y_n) \sim \rho_0(x) * p(x,y), z_m \sim U[0,1]$
    $\theta \leftarrow Adam(\nabla_\theta \frac{2}{NM}\sum_n\sum_m (2z_m-1)(q(x_n)-q(y_n))f(x_n,z_m), \omega, \alpha, \beta_1, \beta_2)$
end while