Wahrscheinlichkeitstheorie und Statistik

Serie 11

Aufgabe 11-1 (Lösungsbegriff der Bayesianische Optimierung). Für eine nicht-negative Zufallsvariable L auf einem Wahrscheinlichkeitsraum $(\Theta, \mathcal{A}, \pi_0)$ definieren wir rekursiv die Folge von Wahrscheinlichkeitsmassen

$$\pi_n(d\theta) = \frac{\exp(-L(\theta))\pi_{n-1}(d\theta)}{\int_\Theta \exp(-L(\theta))\pi_{n-1}(d\theta)}, \qquad n = 1, 2, \ldots$$

Zeigen Sie, dass $\lim_{n\to\infty} \pi_n(A) = 0$ für jedes Ereignis $A \in \mathcal{A}$, sodass

$$\exists \delta > 0 : \qquad \operatorname{ess\,inf}_{\theta \in A} L(\theta) \geq \delta + \operatorname{ess\,inf}_{\theta\in\Theta} L(\theta).$$

Die Interpretation dieser Aussage ist, dass sich π_n für grosses n an Minima von L konzentriert.

Aufgabe 11-2 (Bayesianische Optimierung mit quadratischer Verlustfunktion). Berechnen Sie π_n in folgendem Spezialfall von Aufgabe 4:

$$\pi_0 \sim \mathcal{N}(\mu_0, \sigma_0^2), \qquad L(\theta) = \frac{(\theta - \mu_1)^2}{2\sigma_1^2}.$$

Bonusfrage: Konvergiert π_n in einem geeigneten Sinn gegen das globale Minimum μ_1? **MC 11-3 (Bayesianische Optimierung für Maximum-Likelihood Schätzung).** Wenden Sie den Algorithmus für Bayesianische Optimierung aus Aufgabe 1 auf das Problem der Maximum-Likelihood Schätzung an. Welche Aussagen sind korrekt? (Mehrere richtige Antworten sind möglich.)

(a) Man muss eine a-priori Verteilung π_0 wählen.

(b) Die Verlustfunktion ist die log-Likelihood $L(\theta) = \log f(x|\theta)$.

(c) Man erhält als π_1 die a-posteriori Verteilung für eine Beobachtung x.

(d) Man erhält als π_2 die a-posteriori Verteilung, indem man eine neue Beobachtung x unter dem vorherigen Prior π_1 berücksichtigt.

MC 11-4 (Gradientenfluss). Der Gradientenfluss $\mathrm{Fl}_t(x_0)$ einer Funktion f mit Anfangswert x_0 ist definiert als die Lösung der Differentialgleichung

$$\frac{d}{dt}\mathrm{Fl}_t(x_0) = -\nabla f\big(\mathrm{Fl}_t(x_0)\big), \qquad x(0) = x_0.$$

Welche Aussagen gelten im Limes $t \to \infty$?

(a) Für $f(x) = \exp(-x)$ konvergiert $\mathrm{Fl}_t(0)$ gegen einen Minimierer von f.

(b) Für $f(x) = \exp(-x)$ konvergiert $f\big(\mathrm{Fl}_t(0)\big)$ gegen das Infimum von f.

(c) Für $f(x) = x^2/2$ konvergiert $\mathrm{Fl}_t(1)$ gegen einen Minimierer von f.

(d) Für $f(x) = x^2/2$ konvergiert $f\big(\mathrm{Fl}_t(1)\big)$ gegen das Infimum von f.

Aufgabe 11-5 (Langevin Dynamik). Implementieren Sie die diskrete Langevin-Dynamik

$$\theta_{n+1} = \theta_n - \eta \nabla L(\theta_n) + \sqrt{\eta}\sigma\epsilon_n, \qquad \epsilon_n \sim \mathcal{N}\big(\left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)\big),$$

mit unabhängigen Störungen (ϵ_n), um folgende Verlustfunktion Bayesianisch zu minimieren:

$$L : \mathbb{R}^2 \to \mathbb{R}, \qquad L(x, y) = (x^2 - 1)^2 + y^2.$$

Simulieren Sie mehrere Lösungen θ_{500} mit Anfangswert $\theta_0 = \left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)$, Schrittweite $\eta = 0.1$ und Standardabweichung $\sigma = 0.1$. Plotten Sie die Lösungen gemeinsam mit den Kontouren der Verlustfunktion.