Cross-Entropy Method

Initialize $\mu \in \mathbb{R}^d, \sigma \in \mathbb{R}^d$
for iteration $= 1, 2, \ldots$ **do**
  Collect n samples of $\theta_i \sim N(\mu, \text{diag}(\sigma))$
  Perform a noisy evaluation $R_i \sim \theta_i$
  Select the top $p\%$ of samples (e.g. $p = 20$), which we'll call the **elite set**
  Fit a Gaussian distribution, with diagonal covariance, to the elite set, obtaining a new μ, σ.
end for
Return the final μ.

2. CEM Analysis

Let's introduce some notation to help with the analysis of CEM. The optimization problem solved by CEM can be written as follows:

$$\underset{\theta}{\text{maximize}}\, E_\zeta[f(\theta, \zeta)]$$

Let $\psi \in \mathbb{R}^d$ parameterize the distribution over θ (e.g., via mean and variance of a Gaussian), so in each iteration of CEM, we collect N samples $\theta_i \sim p_\psi(\cdot)$ and function values $f_1, f_2, \ldots, f_N$. Each iteration, we update ψ by solving the subproblem

$$\text{maximize}_\psi \sum_{i \in \text{elite}} \log p_\psi(\theta_i).$$

(1) Explain why CEM does not always reach a local maximum of $E_\zeta[f(\theta, \zeta)]$, even in the limit of infinite samples ($N \to \infty$). (Hint: CEM seeks θ where the $\text{Var}_\zeta[f(\theta, \zeta)]$ is high.)

(2) Let's modify CEM so that each iteration, it solves a subproblem of the form

$$\text{maximize}_\psi \sum_{i=1}^{N} f_i \log p_\psi(\theta_i)$$

I.e., samples θ_i are weighted by the return f_i. (CEM corresponds to weighting by 1 or 0, depending on whether the sample is in the elite set). Show that the resulting algorithm does reach a local maximum of the objective, in an appropriate limit as $N \to \infty$.

3. Score Function Gradient Estimation

It's sometimes illuminating to see the functional form of the score function gradient estimators, for different distributions and parameters in those distributions.

This exercise asks you to determine if the score function gradient estimator can be used for several parameterized probability distributions, and if it can, compute $\nabla_\theta \log p(x|\theta)$.

Here's some notation:

$\mathrm{Ber}(p; a, b)$ is the distribution that takes values a, b with probility $(p, 1-p)$, respectively.
$\mathrm{Poi}(\lambda)$ is the Poisson distribution with mean λ $p(k|\lambda) = \lambda^k e^{-k}/k!,\ k \in \mathbb{N}$
$U(a, b)$ is the uniform distribution between a, b
$\mathrm{Exp}(\lambda)$ is the exponential distribution, $p(x|\lambda) = \lambda e^{-\lambda x}$

Distributions:

1. $\mathrm{Ber}(\theta, a, b)$
2. $\mathrm{Ber}(p, \theta, b)$ ×
3. $\mathrm{Poi}(\theta)$ ✓
4. $N(\theta, \sigma^2)$ ✓
5. $N(\mu, \theta^2)$ ✓
6. $U(0, \theta)$ ✓
7. $\mathrm{Exp}(\theta)$ ✓

1. $x \sim \mathrm{Ber}(\theta, a, b)$ $\quad p(x;\theta) = \begin{cases} \theta & \text{if } x = a \\ 1-\theta & \text{if } x = b \end{cases}$

$$\nabla_\theta \log p(x;\theta) = \begin{cases} 1/\theta & \text{if } x = a \\ 1/(\theta - 1) & \text{if } x = b \end{cases}$$

2. $x \sim \mathrm{Ber}(p, \theta, b)$ $\quad p(x;\theta) = \begin{cases} p & \text{if } x = \theta \\ 1-p & \text{if } x = b \end{cases}$

$$\nabla_\theta \log p(x;\theta) = \begin{cases} 0 & \text{if } x = \theta \\ 0 & \text{if } x = b \end{cases}$$

3. $x \sim \mathrm{Poi}(\theta)$ $\quad p(x;\theta) = \dfrac{\theta^x e^{-x}}{x!}$

$$\nabla_\theta \log p(x;\theta) = \nabla_\theta \log \frac{\theta^x e^{-x}}{x!} = \frac{x!}{\theta^x e^{-x}} \cdot \frac{x\,\theta^{x-1} e^{-x}}{x!} = \frac{x}{\theta}$$

4. $x \sim \mathcal{N}(\theta, \sigma^2)$ $\quad p(x;\theta) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(\frac{(x-\theta)^2}{2\sigma^2}\right)$

$$\nabla_\theta \log p(x;\theta) = \nabla_\theta (x-\theta)^2/2\sigma^2 = -(x-\theta)/\sigma^2$$

5. $x \sim \mathcal{N}(\mu; \theta^2)$ $\quad p(x;\theta) = \frac{1}{\sqrt{2\pi\theta}} \exp\left(\frac{(x-\mu)^2}{2\theta^2}\right)$

$$\nabla_\theta \log p(x;\theta) = -\frac{1}{2}(2\pi\theta)^{-3/2} - (x-\mu)^2/\theta^3$$

6. $x \sim U(0,\theta)$ $\quad p(x;\theta) = 1/\theta$

$$\nabla_\theta \log p(x;\theta) = \nabla_\theta \log 1/\theta = -1/\theta$$

7. $x \sim \mathrm{Exp}(\theta)$ $\quad p(x;\theta) = \theta e^{-\theta x}$

$$\nabla_\theta \log p(x;\theta) = \nabla_\theta [\log\theta - \theta x] = 1/\theta - x$$