Mathematical Algorithms II, M1407.000500
E. Ryu
Fall 2022

Homework 2
Due 5pm, Wednesday, November 02, 2022

Problem 1: Let $\pi^\star$ be an optimal policy. Show that for any policy π,

$$Q^{\pi^\star}(s,a) \ge Q^\pi(s,a), \qquad \forall\, s \in \mathcal{S},\; a \in \mathcal{A}.$$

Problem 2: *Bellman operators for Q are contractions.* Show that the $\mathcal{B}^\pi$ and $\mathcal{B}^\star$ for Q are γ-contractions.

Problem 3: Let π be a policy, not necessarily optimal. Let $\mathcal{B}^\pi$ be the Bellman operator for π and $\mathcal{B}^\star$ the Bellman optimality operator. Show that for any $V\colon \mathcal{S} \to \mathbb{R}$,

$$\mathcal{B}^\pi[V] \le \mathcal{B}^\star[V].$$

Also show that for any $U\colon \mathcal{S} \to \mathbb{R}$ and $V\colon \mathcal{S} \to \mathbb{R}$ such that $U \le V$,

$$\mathcal{B}^\star[U] \le \mathcal{B}^\star[V].$$

Problem 4: *Gaussian calculations for DDPM.* Let $\{\beta_t\}_{t=1,\dots,T} \subset (0,1)$, $X_0 \sim p_0$, and

$$X_t \,|\, X_{t-1} \sim \mathcal{N}\left(\sqrt{1-\beta_t} X_{t-1}, \beta_t I\right), \qquad \text{for } t = 1,\dots,T.$$

Show that

$$X_t \,|\, X_0 \sim \mathcal{N}\left(\sqrt{\bar{\alpha}_t} X_0, (1-\bar{\alpha}_t) I\right), \qquad \bar{\alpha}_t = \prod_{s=1}^{t} (1-\beta_s).$$

Also show that

$$\mathcal{P}(X_{t-1} \,|\, X_t) \approx \mathcal{N}(\mu(X_t,t), \beta_t I), \qquad \mu(X_t,t) = \frac{1}{\sqrt{1-\beta_t}}(X_t + \beta_t \log \nabla p_t(X_t))$$

for small β_t.

Hint. For small β_t, use the approximation

$$p_t(x) = p_{t-1}(x) + \text{h.o.t.}, \qquad \forall\, x \in \mathbb{R}^d.$$

Problem 5: *Fixed-point of Langevin SDE.* Let $p(x)$ is a probability density function that is smooth and strictly positive for all $x \in \mathbb{R}^d$. Let $\{p_t\}_{t\in[0,T]}$ be the marginal density functions of the Langevin SDE

$$dX_t = \frac{1}{2}\nabla_{X_t} \log p(X_t) dt + dW_t.$$

Show that if $p_0 = p$, then $p_t = p$ for all $t > 0$.

Problem 6: *Reverse conditional distribution conditioned on* X_0. DDPM considers the forward process

$$\mathcal{P}(X_t \mid X_{t-1}) \sim \mathcal{N}(\sqrt{1-\beta_t}X_t, \beta_t I)$$

for $t = 1, 2, \dots$ with $X_0 \sim p_{\text{data}}$. In class, we argued that

$$\mathcal{P}(X_{t-1}|X_t) \approx \mathcal{N}\left(\mu_t(X_t), \beta_t I\right), \qquad \mu_t(X_t) = \frac{1}{\sqrt{1-\beta_t}}(X_t + \beta_t \nabla \log p_t(X_t))$$

for $t = 1, 2, \dots$ when $\beta_t \approx 0$. In this problem, show that

$$\mathcal{P}(X_{t-1}|X_t, X_0) = \mathcal{N}\left(\mu_t(X_t \mid X_0), \tilde{\beta}_t I\right),$$

$$\mu_t(X_t \mid X_0) = \frac{1}{\sqrt{1-\beta_t}}(X_t + \beta_t \nabla_{X_t} \log p_{t\,|\,0}(X_t \mid X_0)), \qquad \tilde{\beta}_t = \frac{1 - \prod_{s=1}^{t-1}(1-\beta_s)}{1 - \prod_{s=1}^{t}(1-\beta_s)}\beta_t$$

for $t = 1, 2, \dots$. Do not assume $\beta_t \approx 0$.

Problem 7: D_{KL} *of Gaussian random variables.* Show that

$$D_{\text{KL}}\left(\mathcal{N}(\mu_0, \sigma_0^2 I) \| \mathcal{N}(\mu_1, \sigma_1^2 I)\right) = \frac{1}{2\sigma_1^2}\|\mu_1 - \mu_0\|^2 + \frac{(\sigma_0^2/\sigma_1^2 - 1)d}{2} + d\log\left(\frac{\sigma_1}{\sigma_0}\right),$$

where d is the underlying dimension of the random variables, $\mu_0, \mu_1 \in \mathbb{R}^d$, $\sigma_0 > 0$, and $\sigma_1 > 0$.

Remark. In the context of deep learning, if σ_0 and σ_1 are not trainable parameters, then we can write

$$D_{\text{KL}}\left(\mathcal{N}(\mu_0, \sigma_0^2 I) \| \mathcal{N}(\mu_1, \sigma_1^2 I)\right) = \frac{1}{2\sigma_1^2}\|\mu_1 - \mu_0\|^2 + C.$$

Problem 8: *Single-Q overestimates and double-Q underestimates.* Let $X_1, \dots, X_N \in \mathbb{R}$ be independent (but not necessarily identically distributed) continuous random variables. Assume $\mathbb{E}[|X_i|] < \infty$ for $i = 1, \dots, N$. Write μ_i, f_i, and F_i to respectively denote the mean, PDF, and CDF of X_i for $i = 1, \dots, N$. Consider the goal of estimating $\max_{i=1,\dots,N} \mu_i$.

(a) Using Jensen's inequality, show

$$\max_{i=1,\dots,N} \mathbb{E}[X_i] \le \mathbb{E}\left[\max_{i=1,\dots,N} X_i\right].$$

(b) Show

$$\mathbb{E}\left[\max_{i=1,\dots,N} X_i\right] = \sum_{i=1}^{N} \int_{-\infty}^{\infty} x f_i(x) \prod_{j\neq i} F_j(x)\, dx.$$

(c) Assume we have another set of independent random variables $X_1', X_2', \dots, X_N'$ such that $X_i' \overset{\mathcal{D}}{=} X_i$. To clarify, X_i and X_j' are independent for all $i, j \in \{1, \dots, N\}$, including the case $i = j$. Consider the estimator

$$X_I', \qquad I \in \underset{i=1,\dots,N}{\operatorname{argmax}}\, X_i.$$

Show

$$\mathbb{E}\left[X_I'\right] = \sum_{j=1}^{N} \mathbb{E}[X_i']\mathbb{P}(I = i) = \sum_{i=1}^{N} \int_{-\infty}^{\infty} \mu_i f_i(x) \prod_{j\neq i} F_j(x)\, dx.$$

(d) Show that

$$\mathbb{E}\left[X_I'\right] \le \max_{i=1,\dots,N} \mu_i.$$

Hint. Note, $\max_{i=1,\dots,N} X_i$ has CDF $\prod_{i=1}^{N} F_i$. Once (c) is established, (d) is immediate.