CS229 Machine Learning Problem Set #4 Solution

Roy C.K. Chan

Problem 1. See the file "Q1.ipynb", which is located in the directory $/ps4/Q1/MNIST$.

Problem 2. In the problem, we assume all the regularity conditions, under which we can interchange the order of differentiation, summation and expectation. I only show the case when $z^{(i)}$'s are discrete random variables, but the following arguments can be easily extended to general RVs.

When EM converges, the M-step would have reached a fixed point θ^*, i.e.,

$$\theta^* = \theta^{(T)} = \theta^{(T+1)} = \theta^{(T+2)} = \cdots \,, \tag{1}$$

for some T. Suppose the EM algorithm is now at iteration T and the parameters start out as $\theta^{(T)} = \theta^*$. In the E-step,

$$Q_i(z^{(i)}) = p(z^{(i)}|x^{(i)}; \theta^{(T)}) = p(z^{(i)}|x^{(i)}; \theta^*) \,. \tag{2}$$

Then, in the M-step,

$$\begin{aligned} \theta^{(T+1)} &= \arg\max_\theta \sum_i \sum_{z^{(i)}} Q_i(z^{(i)}) \log \frac{p(x^{(i)}, z^{(i)}; \theta)}{Q_i(z^{(i)})} \\ \implies \theta^* &= \arg\max_\theta \sum_i \sum_{z^{(i)}} Q_i(z^{(i)}) \log p(x^{(i)}, z^{(i)}; \theta) \,, \end{aligned}$$

because of (1). Hence,

$$\nabla_\theta \sum_i \sum_{z^{(i)}} Q_i(z^{(i)}) \log p(x^{(i)}, z^{(i)}; \theta)\Big|_{\theta=\theta^*} = 0 \,. \tag{3}$$

Now consider the LHS of equation (3),

$$
\begin{aligned}
\nabla_\theta \sum_i \sum_{z^{(i)}} Q_i(z^{(i)}) \log p(x^{(i)}, z^{(i)}; \theta)\Big|_{\theta=\theta^*} &= \sum_i \sum_{z^{(i)}} Q_i(z^{(i)}) \nabla_\theta \log p(x^{(i)}, z^{(i)}; \theta)\Big|_{\theta=\theta^*} \\
&= \sum_i \sum_{z^{(i)}} Q_i(z^{(i)}) \frac{\nabla_\theta p(x^{(i)}, z^{(i)}; \theta)\Big|_{\theta=\theta^*}}{p(x^{(i)}, z^{(i)}; \theta^*)} \\
&= \sum_i \sum_{z^{(i)}} p(z^{(i)}|x^{(i)}; \theta^*) \frac{\nabla_\theta p(x^{(i)}, z^{(i)}; \theta)\Big|_{\theta=\theta^*}}{p(x^{(i)}, z^{(i)}; \theta^*)} \\
&= \sum_i \sum_{z^{(i)}} \frac{p(x^{(i)}, z^{(i)}; \theta^*)}{p(x^{(i)}; \theta^*)} \frac{\nabla_\theta p(x^{(i)}, z^{(i)}; \theta)\Big|_{\theta=\theta^*}}{p(x^{(i)}, z^{(i)}; \theta^*)} \\
&= \sum_i \frac{1}{p(x^{(i)}; \theta^*)} \sum_{z^{(i)}} \nabla_\theta p(x^{(i)}, z^{(i)}; \theta)\Big|_{\theta=\theta^*} \\
&= \sum_i \frac{\nabla_\theta p(x^{(i)}; \theta)\Big|_{\theta=\theta^*}}{p(x^{(i)}; \theta^*)} \\
&= \sum_i \nabla_\theta \log p(x^{(i)}; \theta^*) \\
&= \nabla_\theta \sum_i \log p(x^{(i)}; \theta^*) \\
&= \nabla_\theta l(\theta)\Big|_{\theta=\theta^*}.
\end{aligned}
$$

Therefore,

$$\nabla_\theta l(\theta)\Big|_{\theta=\theta^*} = 0\,.$$

Problem 3. For a given unit vector u, with $\mathcal{V} = \{\alpha u : \alpha \in \mathbb{R}\}$,

$$\begin{aligned} f_u(x) &= \arg\min_{v \in \mathcal{V}} ||x - v||^2 \\ &= \Big(\arg\min_{\alpha \in \mathbb{R}} ||x - \alpha u||^2\Big) u \,. \end{aligned} \tag{4}$$

Observe that

$$\begin{aligned} ||x - \alpha u||^2 &= (x - \alpha u)^T (x - \alpha u) \\ &= x^T x - 2(x^T u)\alpha + (u^T u)\alpha^2 \\ &= x^T x - 2(x^T u)\alpha + \alpha^2 \,, \end{aligned}$$

which is a quadratic expression in α, and minimizes at $\alpha = x^T u$. Therefore, by (4),

$$f_u(x) = (x^T u) u \,. \tag{5}$$

Hence,

$$\begin{aligned} & \arg\min_{u : u^T u = 1} \sum_{i=1}^m ||x^{(i)} - f_u(x^{(i)})||_2^2 \\ = \;& \arg\min_{u : u^T u = 1} \sum_{i=1}^m ||x^{(i)} - (x^{(i)T} u) u||_2^2 \\ = \;& \arg\min_{u : u^T u = 1} \sum_{i=1}^m \Big(x^{(i)} - (x^{(i)T} u) u\Big)^T \Big(x^{(i)} - (x^{(i)T} u) u\Big) \\ = \;& \arg\min_{u : u^T u = 1} \sum_{i=1}^m \Big(x^{(i)T} x^{(i)} - (x^{(i)T} u) u^T x^{(i)} - (x^{(i)T} u) x^{(i)T} u + (x^{(i)T} u)^2 u^T u\Big) \\ = \;& \arg\min_{u : u^T u = 1} \sum_{i=1}^m \Big(x^{(i)T} x^{(i)} - (x^{(i)T} u)^2 - (x^{(i)T} u)^2 + (x^{(i)T} u)^2\Big) \\ = \;& \arg\min_{u : u^T u = 1} \sum_{i=1}^m \Big(x^{(i)T} x^{(i)} - (x^{(i)T} u)^2\Big) \\ = \;& \arg\min_{u : u^T u = 1} \Big(- \sum_{i=1}^m (x^{(i)T} u)^2 \Big) \\ = \;& \arg\max_{u : u^T u = 1} \sum_{i=1}^m (u^T x^{(i)})(x^{(i)T} u) \\ = \;& \arg\max_{u : u^T u = 1} u^T \Big(\frac{1}{m} \sum_{i=1}^m x^{(i)} x^{(i)T} \Big) u \,, \end{aligned}$$

which is the same as the "variance maximizing" formulation in class. Therefore, solving the minimization problem gives the first principal component.

Problem 4. See the file "Q4.ipynb".

Problem 5(a). In this problem, we make use of a useful property of supremum stated as follows. Let $f, g : A \to \mathbb{R}$ be bounded functions, we have

$$|\sup_A f - \sup_A g| \quad \leq \quad \sup_A |f - g| \; . \tag{6}$$

The proof of this property is not shown here, since it can be easily found in standard advanced calculus or mathematical analysis textbooks.

Since A is finite, $\sup_A$ becomes $\max_A$ so that

$$|\max_A f - \max_A g| \quad \leq \quad \max_A |f - g| \; . \tag{7}$$

Consider

$$\begin{aligned}
& \|B(V_1) - B(V_2)\|_\infty \\
= \; & \max_{s \in S} | R(s) + \gamma \max_{a \in A} \sum_{s' \in S} P_{sa}(s')V_1(s') - R(s) - \gamma \max_{a \in A} \sum_{s' \in S} P_{sa}(s')V_2(s') | \\
= \; & \gamma \max_{s \in S} | \max_{a \in A} \sum_{s' \in S} P_{sa}(s')V_1(s') - \max_{a \in A} \sum_{s' \in S} P_{sa}(s')V_2(s') | \\
\leq \; & \gamma \max_{s \in S} \max_{a \in A} | \sum_{s' \in S} P_{sa}(s')V_1(s') - \sum_{s' \in S} P_{sa}(s')V_2(s') | \\
= \; & \gamma \max_{s \in S} \max_{a \in A} | \sum_{s' \in S} P_{sa}(s')\big(V_1(s') - V_2(s')\big) | \\
= \; & \gamma \max_{s \in S} \max_{a \in A} | \mathbb{E}_{s' \sim P_{sa}} \big(V_1(s') - V_2(s')\big) | \\
\leq \; & \gamma \max_{s \in S} \max_{a \in A} \mathbb{E}_{s' \sim P_{sa}} \, | \big(V_1(s') - V_2(s')\big) | \\
\leq \; & \gamma \max_{s \in S} \max_{a \in A} \max_{s' \in S} | \big(V_1(s') - V_2(s')\big) | \\
= \; & \gamma \max_{s' \in S} | \big(V_1(s') - V_2(s')\big) | \\
= \; & \gamma \|V_1 - V_2\|_\infty \, ,
\end{aligned}$$

where the fourth line follows from (7), the sixth line holds by re-writing the sum as an expectation (expectation is well-defined for bounded random variables), the seventh line is due to the Jensen's inequality (absolute value is a convex function), the eighth line is true because expectation of a random variable is always less than or equal to its maximum value, the ninth line results from dropping unnecessary maximizations (the quantity inside does not depend on s nor a), and the last line makes use of the definition of max-norm.

Problem 5(b). Assume the contrary that B has two distinct fixed points V_1^*, V_2^*. By (a),

$$0 < \|V_1^* - V_2^*\|_\infty = \|B(V_1^*) - B(V_2^*)\|_\infty \leq \gamma \|V_1^* - V_2^*\|_\infty \implies \gamma \geq 1 \, ,$$

which contradicts the fact that $\gamma < 1$. Hence, there is at most one solution to the Bellman equations.

Remarks: Existence and uniqueness of fixed point V^* follow immediately from Banach fixed point theorem.

Problem 6(a)(b). See the files "cart_pole.py", "control.py" and "learning curve.png", which are located in the directory $/ps4/Q6/Inverted\ Pendulum$.

Remarks: It seems to me that we only need to update the state transition probabilities P_{sa}, but not the reward function $R(s)$. From my implementations, it can be easily check that we always get the same estimated reward function, which is a 163-D vector with all elements being 0, except the last one being -1. (Not sure whether there are mistakes in my understandings/implementations, or it is indeed what the problem asks for.)