Introduction to Machine Learning

Fall 2022

University of Science and Technology of China

Lecturer: Jie Wang | Homework 7

Posted: December 23, 2022 | Due: January 3, 2023

Notice, to get the full credits, please present your solutions step by step.

Exercise 1: Singular Value Decomposition

Let $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\operatorname{rank}\mathbf{A} = r$. The SVD of $\mathbf{A}$ is $\mathbf{A} = \mathbf{U\Sigma V}^\top$, where $\mathbf{U} \in \mathbb{R}^{m\times m}, \mathbf{\Sigma} \in \mathbb{R}^{m\times n}, \mathbf{V} \in \mathbb{R}^{n\times n}$, and we sort the diagonal entries of $\mathbf{\Sigma}$ in the descending order $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$. Denote

$$\mathbf{U_1} = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r),\ \mathbf{U_2} = (\mathbf{u}_{r+1}, \ldots, \mathbf{u}_m),$$
$$\mathbf{V_1} = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r),\ \mathbf{V_2} = (\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n).$$

The column space of $\mathbf{A}$ is the set

$$\mathcal{C}(\mathbf{A}) = \{\mathbf{y} \in \mathbb{R}^m : \mathbf{y} = \mathbf{Ax}, \mathbf{x} \in \mathbb{R}^n\}.$$

The null space of $\mathbf{A}$ is the set

$$\mathcal{N}(\mathbf{A}) = \{\mathbf{y} \in \mathbb{R}^n : \mathbf{Ay} = \mathbf{0}\}.$$

1. Show that
   (a) $P_{\mathcal{C}(\mathbf{A})}(\mathbf{x}) = \mathbf{U_1}\mathbf{U_1}^\top\mathbf{x}$;
   (b) $P_{\mathcal{N}(\mathbf{A})}(\mathbf{x}) = \mathbf{V_2}\mathbf{V_2}^\top\mathbf{x}$;
   (c) $P_{\mathcal{C}(\mathbf{A}^\top)}(\mathbf{x}) = \mathbf{V_1}\mathbf{V_1}^\top\mathbf{x}$;
   (d) $P_{\mathcal{N}(\mathbf{A}^\top)}(\mathbf{x}) = \mathbf{U_2}\mathbf{U_2}^\top\mathbf{x}$.
2. The Frobenius norm of $\mathbf{A}$ is

$$\|\mathbf{A}\|_F = \sqrt{\sum_{i=1}^m\sum_{j=1}^n a_{i,j}^2}.$$

   (a) Show that $\|\mathbf{A}\|_F^2 = \operatorname{tr}(\mathbf{A}^\top\mathbf{A})$.
   (b) Let $\mathbf{B} \in \mathbb{R}^{m\times n}$. Suppose that $\mathcal{C}(\mathbf{A})\perp\mathcal{C}(\mathbf{B})$, that is,

$$\langle\mathbf{a}, \mathbf{b}\rangle = 0,\ \forall\, \mathbf{a} \in \mathcal{C}(\mathbf{A}),\ \mathbf{b} \in \mathcal{C}(\mathbf{B}).$$

   Show that

$$\|\mathbf{A} + \mathbf{B}\|_F^2 = \|\mathbf{A}\|_F^2 + \|\mathbf{B}\|_F^2.$$

3. Given $K < r$, $K \in \mathbb{N}$, please solve the problem as follows

$$\min_{\mathbf{X} \in \mathbb{R}^{m \times n}} \{\|\mathbf{A} - \mathbf{X}\|_F : \operatorname{rank} \mathbf{X} \leq K\}.$$

For simplicity, you can assume that all singular values of $\mathbf{A}$ are different.

4. **Programming Exercise.** We provide you with a color image ("Hinton.jpg"). Suppose that $\mathbf{A} = (\mathbf{A}_i)_{i=1}^3$ is the data tensor of the image, where $\mathbf{A}_i$, $i = 1, 2, 3$, represents different channels. We have each $\mathbf{A}_i \in \mathbb{R}^{500 \times 500}$ and $r = \operatorname{rank} \mathbf{A}_i = 500$. In this exercise, you are expected to implement an image compression algorithm following the steps below. You can use your favorite programming language.

   (a) Compute the SVD $\mathbf{A}_i = \mathbf{U}_i \mathbf{\Sigma}_i \mathbf{V}_i^\top = \sum_{j=1}^r \sigma_{i,j} \mathbf{u}_{i,j} \mathbf{v}_{i,j}^\top$, where $i = 1, 2, 3$, $\sigma_{i,1} \geq \sigma_{i,2} \geq \cdots \geq \sigma_{i,r} > 0$ are the diagonal entries of $\mathbf{\Sigma}_i$, $\mathbf{u}_{i,j}$ is the jth column of $\mathbf{U}_i$, and $\mathbf{v}_{i,j}$ is the jth column of $\mathbf{V}_i$.

   (b) Use the first k ($k < r$) terms of SVD to approximate the original image $\mathbf{A}$. Then, we get the compressed images, the data tensors of which are $\mathbf{A}_k = (\mathbf{A}_{i,k})_{i=1}^3 = (\sum_{j=1}^k \sigma_{i,j} \mathbf{u}_{i,j} \mathbf{v}_{i,j}^\top)_{i=1}^3$. Compute $||\mathbf{A} - \mathbf{A}_k||_F$, i.e., $\sum_{i=1}^3 ||\mathbf{A}_i - \mathbf{A}_{i,k}||_F$, for $k = 2, 4, 8, 16, 32, 64, 128, 256$.

   (c) Plot $\mathbf{A}_k$ as images for all the ks in (b).

Figure 1: Hinton

Solution: ■

Exercise 2: Principle Component Analysis

Suppose that we have a set of data instances $\{\mathbf{x}_i\}_{i=1}^n \subset \mathbb{R}^d$. Let $\widetilde{\mathbf{X}} \in \mathbb{R}^{d\times n}$ be the matrix whose i^{th} column is $\mathbf{x}_i - \bar{\mathbf{x}}$, where $\bar{\mathbf{x}}$ is the sample mean, and $\mathbf{S}$ be the sample variance matrix.

1. For $\mathbf{G} \in \mathbb{R}^{d\times K}$, let us define

$$f(\mathbf{G}) = \mathrm{tr}(\mathbf{G}^\top \mathbf{S}\mathbf{G}). \tag{1}$$

   Show that $f(\mathbf{GQ}) = f(\mathbf{G})$ for any orthogonal matrix $\mathbf{Q} \in \mathbb{R}^{K\times K}$.

2. Please find $\mathbf{g}_1$ defined as follows by the Lagrange multiplier method.

$$\mathbf{g}_1 := \underset{\mathbf{g}\in\mathbb{R}^d}{\mathbf{argmax}}\{f(\mathbf{g}) : \|\mathbf{g}\|_2 = 1\}, \tag{2}$$

   where f is defined by (1). Notice that, the vector $\mathbf{g}_1$ is the first principal component vector of the data.

3. Please find $\mathbf{g}_2$ defined as follows by the Lagrange multiplier method.

$$\mathbf{g}_2 := \underset{\mathbf{g}\in\mathbb{R}^d}{\mathbf{argmax}}\{f(\mathbf{g}) : \|\mathbf{g}\|_2 = 1, \langle \mathbf{g}, \mathbf{g}_1\rangle = 0\},$$

   where $\mathbf{g}_1$ is given by (2). Similar to $\mathbf{g}_1$, the vector $\mathbf{g}_2$ is the second principal component vector of the data.

4. Please derive the first K principal component vectors by repeating the above process.

5. What is $f(\mathbf{g}_k)$, $k = 1, \ldots, K$? What about their meaning?

6. When are the first K principal component vectors unique?

7. **Programming Exercise.** We provide you with 130 handwritten 3s, each a digitized 28×28 grayscale image("imgs_3"). Please finish the following steps . You can use your favorite programming language.

   (a) Consider these images as points $\mathbf{x}_i \in \mathbb{R}^{784}, i = 1, \cdots, 130$. Let $\bar{\mathbf{x}}$ be the mean of all $\mathbf{x}_i$. Let $\mathbf{X} \in \mathbb{R}^{130\times 784}$ with $\mathbf{x}_i - \bar{\mathbf{x}}$ as its i^{th} row.

   (b) Calculate the SVD, $\mathbf{X} = \mathbf{U\Sigma V}^\top$. Let $\mathbf{v}_1, \mathbf{v}_2$ be the columns of $\mathbf{V}$ corresponding to the 2 largest singular values, respectively. Please show $\bar{\mathbf{x}}, \mathbf{v}_1$ and $\mathbf{v}_2$, considering them as 28×28 grayscale images.

   (c) What do the three images illustrate?

Solution:

■

Exercise 3: Properties of Transition Matrix

A transition matrix (also called a stochastic matrix, probability matrix) is a square matrix used to describe the transitions of a Markov chain. Each of its entries is a nonnegative real number representing a probability. A right (left) transition matrix is a square matrix with each row (column) summing to one. Without loss of generality, we study the right transition matrix in this exercise. Suppose that $\mathbf{T} \in \mathbb{R}^{n \times n}$ is a right transition matrix.

1. Show that 1 is an eigenvalue of $\mathbf{T}$.
2. Let λ be an eigenvalue of $\mathbf{T}$. Show that $|\lambda| \leq 1$.
3. Show that $\mathbf{I} - \gamma\mathbf{T}$ is invertible, where $\mathbf{I} \in \mathbb{R}^{n \times n}$ is the identity matrix and $\gamma \in (0, 1)$.
4. We will show that $(\mathbf{I} - \gamma\mathbf{T})^{-1} = \sum_{i=0}^{\infty} (\gamma\mathbf{T})^i$.
   (a) For $\mathbf{x} \in \mathbb{R}^n$, the infinity norm is defined by
   $$\|\mathbf{x}\|_\infty = \max_i |x_i|.$$
   The induced norm of a matrix $\mathbf{M} \in \mathbb{R}^{m \times n}$ is
   $$\|\mathbf{M}\|_\infty = \max_{\|\mathbf{x}\|_\infty \leq 1} \|\mathbf{M}\mathbf{x}\|_\infty.$$
      i. Show that $\|\mathbf{M}\|_\infty = \max_i \sum_{j=1}^{n} |m_{i,j}|$.
      ii. Show that $\|\mathbf{A}\mathbf{B}\|_\infty \leq \|\mathbf{A}\|_\infty \|\mathbf{B}\|_\infty$ holds for any $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{B} \in \mathbb{R}^{n \times p}$.
   (b) Show that the sequence $\left\{\sum_{i=0}^{k} (\gamma\mathbf{T})^i\right\}_{k=0}^{\infty}$ converges.
   (c) Let $\left\{\sum_{i=0}^{k} (\gamma\mathbf{T})^i\right\}_{k=0}^{\infty}$ converges to a matrix L. Show that $(\mathbf{I} - \gamma\mathbf{T})^{-1} = L$.

Solution:

■

Exercise 4: Grid World with a Given Policy

Consider the grid world shown in Figure 2. The finite state space is $\mathcal{S} = \{s_i : i = 1, 2, \ldots, 11\}$ and the finite action space is $\mathcal{A} = \{\text{up, down, left, right}\}$.
State transition probabilities: After the agent picks and performs a certain action, there are four possibilities for the next state: the destination state, the current state, the states to the right and left of the current state. If the states are reachable, the corresponding probabilities are 0.7, 0.1, 0.05, and 0.15, respectively; otherwise, the agent stays where it is. The game will terminate if the agent arrives at s_{10} (loss) or s_{11} (win).
Reward: After the agent picks and performs a certain action at its current state, it receives rewards of 100, -100, and 0, if it arrives at states s_{11}, s_{10}, and all the other states, respectively.
Policy: In Figure 2, the arrows show the policy $\pi : \mathcal{S} \to \mathcal{A}$ for the agent. The random variable S_t is the state at time t under the policy π.

1. Find the matrix $\mathbf{M} \in \mathbb{R}^{11\times 11}$ with $m_{i,j} = \mathbf{P}(S_{t+1} = s_j | S_t = s_i)$, i.e., the conditional probability of the agent moving from s_i to s_j.

2. Suppose that the initial state distribution is uniform distribution, that is $\mathbf{P}(S_0 = s_i) = 1/11$, $i = 1, \ldots, 11$.

   (a) Find the distributions $\mathbf{P}(S_1)$ and $\mathbf{P}(S_2)$ by following the policy π.

   (b) Show that the agent would finally arrive at either s_{10} or s_{11}, i.e.,

   $$\lim_{t\to\infty} \mathbf{P}(S_t = s_i) = 0,\ i = 1, \ldots, 9.$$

   (c) Find $\lim_{t\to\infty} \mathbf{P}(S_t = s_{10})$ and $\lim_{t\to\infty} \mathbf{P}(S_t = s_{11})$.

3. Find the value function corresponding to the policy π, where the discount factor $\gamma = 0.9$.

4. Show that the result in (2b) holds for any initial probabilities we choose for $\mathbf{P}(S_0 = s_i)$, $i = 1, \ldots, 11$.

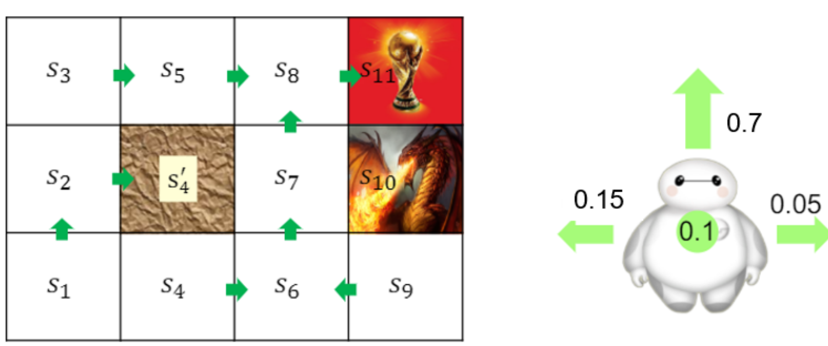

Figure 2: Illustration of a grid world with a policy.

Exercise 5: Optimal Policy

Consider the grid world problem described in Exercise 4. Let π^* be the optimal policy, V^* the corresponding value function, Q^* the corresponding Q function, and $\gamma = 0.9$.

1. Please derive the Bellman Optimality Equation in terms of the value function V^* and the Q function Q^*, respectively.
2. Please choose one of the algorithms we introduced in class to find π^* and V^* respectively and write their pseudocode (hand in your code if you have one).
3. Please design a reward scheme such that following the resulting optimal policy will never lose. Specifically, you need to derive the resulting optimal policy (the proof is not required) and show

$$\lim_{t\to\infty} \mathbf{P}(S_t = s_i) = 0, \; i = 1, \ldots, 10,$$

   whenever $\mathbf{P}(S_0 = s_{10}) = \mathbf{P}(S_0 = s_{11}) = 0$.

Solution:

■

Exercise 6: Value Iteration and Policy Iteration

Consider a Markov Decision Process with bounded rewards and finite state-action pairs. The transition probability is $\mathbf{P}[s'|s,a]$, the discounted factor is $\gamma \in (0,1)$, and the reward function is $r(s,a)$. Let $\pi : \mathcal{S} \to \mathcal{A}$ be a deterministic policy and $Q^\pi(s,a)$ be the accumulated reward by performing the action a first and then following the policy π.

1. Let V^k denote the value function after the k^{th} iteration of the value iteration algorithm. Please show that value iteration achieves linear convergence rate, that is

$$\|V^* - V^{k+1}\|_\infty \leq \gamma \|V^* - V^k\|_\infty.$$

2. (a) Find the Bellman Equation for Q^π.

   (b) Consider a new policy π' given by

$$\pi'(s) = \mathop{\mathbf{argmax}}_{a \in \mathcal{A}} Q^\pi(s,a).$$

   Note that if $\mathbf{argmax}_{a\in\mathcal{A}} Q^\pi(s,a)$ is not unique, we can choose one action arbitrarily. Show that $V^{\pi'}(s) \geq V^\pi(s)$ for all $s \in \mathcal{S}$.

Solution:

■

Exercise 7: Q-learning algorithm (Optional)

1. Consider the Q-learning algorithm for any deterministic MDP with finite state-action pairs and non-negative rewards. Assume that we initialize all $\hat{Q}$ values to zero. Let $\hat{Q}_n(s,a)$ denote the learned $\hat{Q}(s,a)$ value after the n^{th} iteration of the training procedure in Q learning algorithm.

   (a) Please show that $\hat{Q}$ values never decrease during training, that is

   $$\hat{Q}_{n+1}(s,a) = r(s,a) + \gamma \max_{a'} \hat{Q}_n(s',a') \geq \hat{Q}_n(s,a),\ \forall s,a,n,$$

   where s' is the state the agent attains after taking action a at state s.

   (b) Please show that throughout the training process, every $\hat{Q}$ value will remain in the interval between zero and the optimal Q function Q^*, that is

   $$0 \leq \hat{Q}_n(s,a) \leq Q^*(s,a),\ \forall s,a,n.$$

2. Consider the Q-learning algorithm for a stochastic MDP with finite state-action pairs. The transition probability is $\mathbf{P}[s'|s,a]$ and the reward function is deterministic, denoted by $r(s,a)$. Assume that we initialize all $\hat{Q}$ values to zero. Let $\hat{Q}_n(s,a)$ denote the learned $\hat{Q}(s,a)$ value after the n^{th} iteration of the training procedure in Q-learning algorithm. Please show that

$$\mathbb{E}_{s'\sim\mathbf{P}[\cdot|s,a]}\left[r(s,a) + \gamma \max_{a'} \hat{Q}_n(s',a')\right] \geq r(s,a) + \gamma \max_{a'} \mathbb{E}_{s'\sim\mathbf{P}[\cdot|s,a]}\left[\hat{Q}_n(s',a')\right], \forall s,a,n.$$

Solution:

■