Homework 3

肖飞宇　2018210441

1 DIMENSION REDUCTION, PCA

1.1 Minimum Error Formulation

Proof. 对于 R^p 空间存在正交基，

$$\{\mu_i\}, i = 1, \ldots, p \quad \mu_i^\top \mu_j = \delta_{ij} \tag{1.1}$$

那么每一个数据点都可以被表示为

$$\mathbf{x}_n = \sum_{i=1}^{d} \alpha_{ni} \mathbf{u}_i \tag{1.2}$$

那么在R^D空间中近似(D<p)，即

$$\tilde{\mathbf{x}}_n = \sum_{i=1}^{d} z_{ni}\mu_i + \sum_{i=d+1}^{p} b_i \mu_i \tag{1.3}$$

其中的系数z_{ni}和数据x_n有关，问题就转化为选取合适的参数，使得下面的目标函数最小（这就是 Minimum Error Formation）：

$$J = \frac{1}{N}\sum_{n=1}^{N} \|\mathbf{x}_n - \tilde{\mathbf{x}}_n\|^2 \tag{1.4}$$

将$\tilde{\mathbf{x}}_n$代入上式，得到

$$\begin{aligned} J &= \frac{1}{N}\sum_{n=1}^{N} (X_n - \sum_{i=1}^{d} z_{ni}\mu_i - \sum_{i=d+1}^{p} b_i\mu_i)^T (X_n - \sum_{i=1}^{d} z_{ni}\mu_i - \sum_{i=d+1}^{p} b_i\mu_i) & (1.5)\\ &= \frac{1}{N}\sum_{n=1}^{N} \left(x_n^T x_n - 2x_n^T (\sum_{i=1}^{d} z_{ni}\mu_i + \sum_{i=d+1}^{p} b_i\mu_i) + \sum_{i=1}^{d} z_{ni}^2 \right) + \sum_{i=d+1}^{p} b_i^2 & (1.6) \end{aligned}$$

上式对z_{ni}求导有

$$\frac{1}{N}(-2x_n^T\mu_i + 2z_{ni}) = 0 \tag{1.7}$$

得

$$z_{ni} = \mathbf{x}_n^\top \mu_i, i = 1, \ldots, d \tag{1.8}$$

对b_i求导有

$$\frac{1}{N}(-2x_n^T\mu_i) + 2b_i = 0 \tag{1.9}$$

得到

$$b_i = \overline{\mathbf{x}}^\top \mu_i, i = d+1, \ldots, p \tag{1.10}$$

□

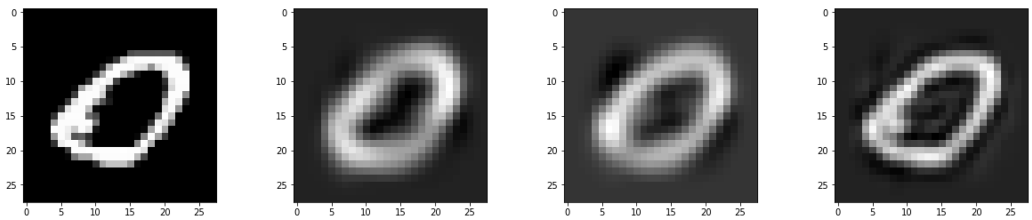
1.2 MNIST

算法

具体的 PCA 算法的实现见 https://github.com/feiyuxiaoThu/SML/blob/master/Assigenments/Assignment3-PCA-Mnist/MnistPCA.py.

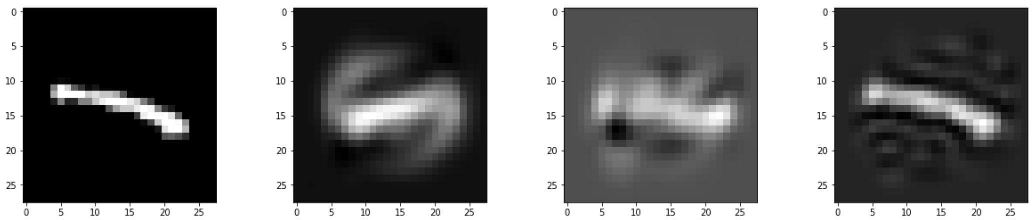
结果

选定数字0和1进行说明，其结果分别如图1.1和图1.2所示



(a) Original Data (b) 30% Information (c) 60% Information (d) 90% Information

Figure 1.1: 数字0



(a) Original Data (b) 30% Information (c) 60% Information (d) 90% Information

Figure 1.2: 数字1

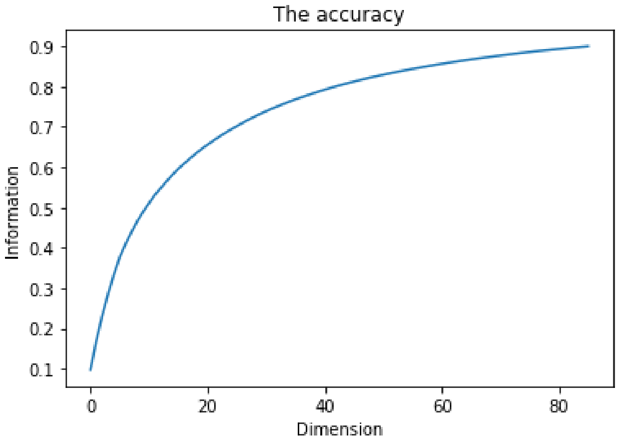
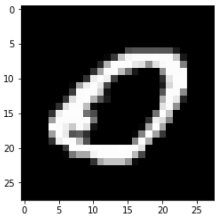
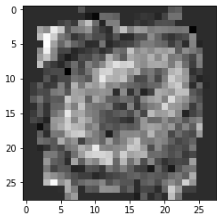


Figure 1.3: Dimension and Information

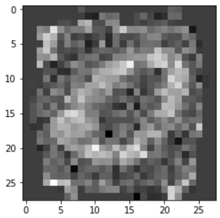
可以将使用的近似维度和准确程度的关系绘制如图1.3： 对于数据不进行中心化的情况，其结果如图1.4和图1.5所示，可以看出不进行中心化结果的噪声非常严重，造成精确度严重下降



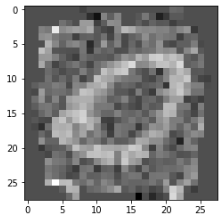
(a) Original Data



(b) 30% Information

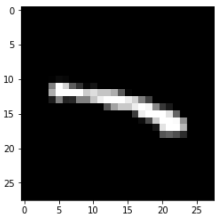


(c) 60% Information

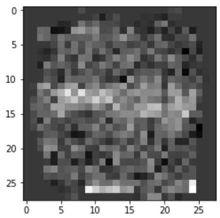


(d) 90% Information

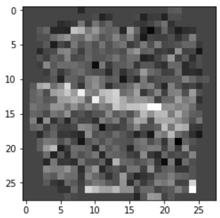
Figure 1.4: 数字0（未中心化）



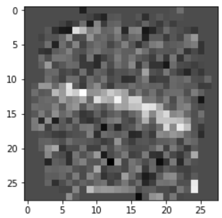
(a) Original Data



(b) 30% Information



(c) 60% Information



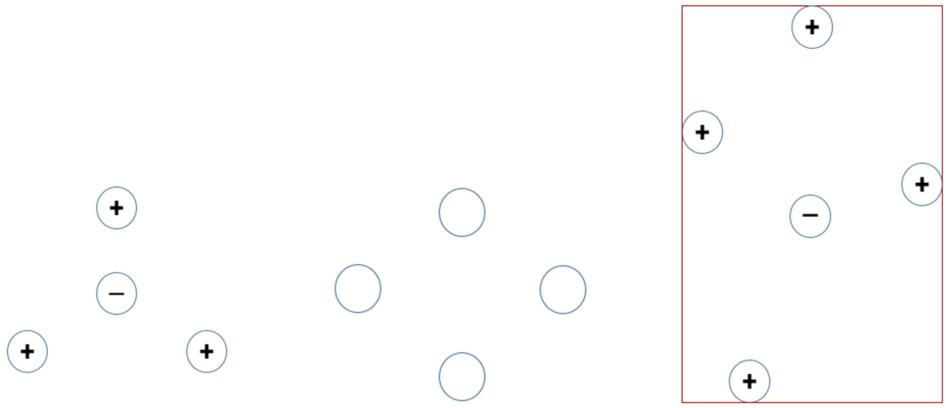
(d) 90% Information

Figure 1.5: 数字1（未中心化）

2 Learning Theory

2.1 VC Dimension

Proof. 首先不是所有四个点的集合都可以被打散，如下图2.1(1)中的组合就不可以被打散，但这并不足以说明VC-dimension至多是3，如图2.1(2)所示，显然可以有一个矩形包含任意给定的点集而不包含其他，所以我们可以得到VC-dimension至少是4。再考察任意五个不同的点的集合 $\{v_1, v_2, v_3, v_4, v_5\} \subseteq \mathbb{R}^2$，如图2.1(3)所示，考虑一个矩形包含了集合中最大的x,最大的y，最小的x，最小的y坐标的点（这些点并不需要不同，但是至多有四个），这些点记为$S \subset \{v_1, v_2, v_3, v_4, v_5\}$。可以看出，任何包含了S的矩形一定会包含全部的点$v_1, v_2, v_3, v_4, v_5$，即至多有一个点不在S中，但是一定在矩形中，记为$v_w$。只需要把S中的点标注为+而把$v_w$标注为−就不可能被打散，可以说明VC-dimension小于5。



(a) 不可以打散的四个点 (b) 可以打散的四个点 (c) 60% Information

Figure 2.1: VC-dimension示意图

□

2.2 Generalization Bound

Proof. 我们首先计算 $H = \{(a < x < b) | a, b \in R\}$的$m_{\mathcal{H}}(N)$，注意到对给定的N个点，数轴被分割成N+1个区域，显然最终的结果取决于a,b所处的区间的位置，这就有$\binom{N+1}{2}$种结果，另外，如果a,b在同一个区间，那么最终的结果就是全是-1。把这两种情况一起考虑可以得到：

$$m_{\mathcal{H}}(N) = \binom{N+1}{2} + 1 = \frac{1}{2}N^2 + \frac{1}{2}N + 1 \tag{2.1}$$

再使用Vapnik-Chervonenkis不等式，有

$$\begin{aligned} P[|E_{in} - E_{out}| > \epsilon] &\le 4 m_{\mathcal{H}}(2N) e^{-\frac{1}{8}\epsilon^2 N} && (2.2) \\ &= (8N^2 + 8N + 4) e^{-\frac{1}{8}\epsilon^2 N} && (2.3) \end{aligned}$$

□

3 Reinforcement Learning

3.1 'Vanilla' Policy Gradient

Proof. 首先，我们的目标函数为：

$$J(\theta) = E_{\tau\sim P(\tau;\theta)}(\sum_t r(s_t,a_t)) = E_{\tau\sim P(\tau;\theta)}(R(\tau)) \tag{3.1}$$

根据期望的定义有

$$J(\theta) = \int p_\theta(\tau)R(\tau)d\tau \tag{3.2}$$

其梯度为

$$\nabla_\theta J(\theta) = \int \nabla_\theta p_\theta(\tau)R(\tau)d\tau \tag{3.3}$$

由链式求导法则有

$$p_\theta(\tau)\nabla_\theta \log p_\theta(\tau) = p_\theta(\tau)\frac{\nabla_\theta p_\theta(\tau)}{p_\theta(\tau)} = \nabla_\theta p_\theta(\tau) \tag{3.4}$$

将上式代入方程3.3，有

$$\nabla_\theta J(\theta) = \int p_\theta(\tau)\nabla_\theta \log p_\theta(\tau)R(\tau)\mathrm{d}\tau = \mathbf{E}_{\tau\sim p_\theta(\tau)}\left[\nabla_\theta \log p_\theta(\tau)R(\tau)\right] \tag{3.5}$$

而又已知

$$p(\tau;\theta) = p(s_1)\prod_t \pi_\theta(a_t|s_t)\,p(s_{t+1}|s_t,a_t) \tag{3.6}$$

两边取对数有

$$\log p(\tau;\theta) = \log p(\mathbf{s}_1) + \sum_{t=1}^{T}\left[\log \pi_\theta(\mathbf{a}_t|\mathbf{s}_t) + \log p(\mathbf{s}_{t+1}|\mathbf{s}_t,\mathbf{a}_t)\right] \tag{3.7}$$

注意到$\nabla_\theta \log p_\theta(\tau)$仅与含有$\theta$的项有关，知其导数为

$$\nabla_\theta \log p(\tau;\theta) = \sum_{t=1}^{T}\nabla_\theta \log \pi_\theta(\mathbf{a}_t|\mathbf{s}_t) \tag{3.8}$$

代入方程3.5得到：

$$\nabla_\theta J(\theta) = E_{\tau\sim P(\tau;\theta)}\left(\sum_t \nabla_\theta \log \pi_\theta(a_t|s_t)\,R(\tau)\right) \tag{3.9}$$

□

3.2 Actor-Cirtic

Proof. 引入一个baseline函数$B(s)$，则有

$$\nabla_\theta J(\theta) = \mathbb{E}_{\pi_\theta}\left[\nabla_\theta \log \pi_\theta(s,a)(Q^{\pi_\theta}(s,a) - B(s)\right] \tag{3.10}$$

计算

$$
\begin{aligned}
\mathbb{E}\left[\nabla_\theta \log P(\tau;\theta) B(s)\right] &= \sum_\tau P(\tau;\theta) \nabla_\theta \log P(\tau;\theta) B(s) && (3.11)\\
&= \sum_\tau P(\tau;\theta) \frac{\nabla_\theta P(\tau;\theta)}{P(\tau;\theta)} B(s) && (3.12)\\
&= \sum_\tau \nabla_\theta P(\tau;\theta) B(s) && (3.13)\\
&= \nabla_\theta \left(\sum_\tau P(\tau) B(s)\right) && (3.14)\\
&= B(s) \nabla_\theta \left(\sum P(\tau)\right) = 0 && (3.15)
\end{aligned}
$$

可以看出不改变原有的期望,而对于 critic 算法

$$
\begin{aligned}
\nabla_\theta J(\theta) =& \mathbb{E}_{\pi_\theta}\left[\nabla_\theta \log \pi_\theta(s,a)\left(R(\tau) - B(s)\right]\right. && (3.16)\\
=& \mathbb{E}_{\pi_\theta}\left(\sum_t \nabla_\theta \log \pi_\theta\left(a_t|s_t\right)\right)\left(\sum_t r\left(s_t,a_t\right) - B(s)\right) && (3.17)\\
=& \mathbb{E}_{\pi_\theta}\left(\sum_t \nabla_\theta \log \pi_\theta\left(a_t|s_t\right)\right)\left(\underbrace{\sum_{k=0}^{t-1} r\left(s_t,a_t\right)}_{\mathbb{A}} + \underbrace{\sum_{k=t} r\left(s_t,a_t\right)}_{\mathbb{B}} - B(s)\right) && (3.18)
\end{aligned}
$$

上式中$\mathbb{A}$和a_t无关，所以可以去掉减小方差，所以最后给出的估计是

$$
\mathbb{E}_{\pi_\theta}\left(\sum_t \nabla_\theta \log \pi_\theta\left(a_t|s_t\right)\right)\left(\sum_{k=t} r\left(s_t,a_t\right) - B(s)\right) \tag{3.19}
$$

□