# 机器学习

# 实验指导书

2019.9

山东大学

# 目录

实验 1 ..... - 2 -
实验 2 最大似然估计 ..... - 3 -
实验 3 非参数估计 ..... - 4 -
实验 4 神经网络学习 ..... - 5 -
实验 5 集成学习 ..... - 6 -

# 实验 1

上机练习 2.5 节第 4 题

# 实验 2 最大似然估计

1、实验目的

（1）掌握用最大似然估计进行参数估计的原理；

（2）当训练样本服从多元正态分布时，计算不同高斯情况下的均值和方差。

2、实验数据

| 样本 | 类 1 | | | 类 2 | | |
|---|---|---|---|---|---|---|
| | $x_1$ | $x_2$ | $x_3$ | $x_1$ | $x_2$ | $x_3$ |
| 1 | 0.011 | 1.03 | -0.21 | 1.36 | 2.17 | 0.14 |
| 2 | 1.27 | 1.28 | 0.08 | 1.41 | 1.45 | -0.38 |
| 3 | 0.13 | 3.12 | 0.16 | 1.22 | 0.99 | 0.69 |
| 4 | -0.21 | 1.23 | -0.11 | 2.46 | 2.19 | 1.31 |
| 5 | -2.18 | 1.39 | -0.19 | 0.68 | 0.79 | 0.87 |
| 6 | 0.34 | 1.96 | -0.16 | 2.51 | 3.22 | 1.35 |
| 7 | -1.38 | 0.94 | 0.45 | 0.60 | 2.44 | 0.92 |
| 8 | -1.02 | 0.82 | 0.17 | 0.64 | 0.13 | 0.97 |
| 9 | -1.44 | 2.31 | 0.14 | 0.85 | 0.58 | 0.99 |
| 10 | 0.26 | 1.94 | 0.08 | 0.66 | 0.51 | 0.88 |

3、实验内容及说明

使用上面给出的三维数据：

1. 编写程序，对类 1 和类 2 中的 3 个特征 $x_i$ 分别求解最大似然估计的均值$\hat{\mu}$和方差$\hat{\sigma}^2$。

2. 编写程序，处理二维数据的情形$p(\mathbf{x})\sim N(\boldsymbol{\mu},\boldsymbol{\Sigma})$。对类 1 和类 2 中任意两个特征的组合分别求解最大似然估计的均值$\hat{\boldsymbol{\mu}}$和方差$\hat{\boldsymbol{\Sigma}}$（每个类有 3 种可能）。

3. 编写程序，处理三维数据的情形$p(\mathbf{x})\sim N(\boldsymbol{\mu},\boldsymbol{\Sigma})$。对类 1 和类 2 中三个特征求解最大似然估计的均值$\hat{\boldsymbol{\mu}}$和方差$\hat{\boldsymbol{\Sigma}}$。

4. 假设三维高斯模型是可分离的，即$\boldsymbol{\Sigma}=diag(\sigma_1^2,\sigma_2^2,\sigma_3^2)$，编写程序估计类 1 和类 2 中的均值和协方差矩阵中的参数。

5. 比较前 4 种方法计算出来的每一个特征的均值 $\mu_i$ 的异同，并加以解释。

6. 比较前 4 种方法计算出来的每一个特征的方差 $\sigma_i$ 的异同，并加以解释。

# 实验 3 非参数估计

1、实验目的

（1）掌握用非参数的方法估计概率密度；

（2）了解 parzen 窗方法的原理；

（3）了解 k 近邻方法的原理

2、实验数据

| 样本 | 类 1 | | | 类 2 | | | 类 3 | | | 类 4 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $x_1$ | $x_2$ | $x_3$ | $x_1$ | $x_2$ | $x_3$ | $x_1$ | $x_2$ | $x_3$ | $x_1$ | $x_2$ | $x_3$ |
| 1 | 0.28 | 1.31 | -6.2 | 0.42 | -0.087 | 0.58 | -0.4 | 0.58 | 0.089 | 0.83 | 1.6 | -0.014 |
| 2 | 0.07 | 0.58 | -0.78 | -0.2 | -3.3 | -3.4 | -0.31 | 0.27 | -0.04 | 1.1 | 1.6 | 0.48 |
| 3 | 1.54 | 2.01 | -1.63 | 1.3 | -0.32 | 1.7 | 0.38 | 0.055 | -0.035 | -0.44 | -0.41 | 0.32 |
| 4 | -0.44 | 1.18 | -4.32 | 0.39 | 0.71 | 0.23 | -0.15 | 0.53 | 0.011 | 0.047 | -0.45 | 1.4 |
| 5 | -0.81 | 0.21 | 5.73 | -1.6 | -5.3 | -0.15 | -0.35 | 0.47 | 0.034 | 0.28 | 0.35 | 3.1 |
| 6 | 1.52 | 3.16 | 2.77 | -0.029 | 0.89 | -4.7 | 0.17 | 0.69 | 0.1 | -0.39 | -0.48 | 0.11 |
| 7 | 2.20 | 2.42 | -0.19 | -0.23 | 1.9 | 2.2 | -0.011 | 0.55 | -0.18 | 0.34 | -0.079 | 0.14 |
| 8 | 0.91 | 1.94 | 6.21 | 0.27 | -0.3 | -0.87 | -0.27 | 0.61 | 0.12 | -0.3 | -0.22 | 2.2 |
| 9 | 0.65 | 1.93 | 4.38 | -1.9 | 0.76 | -2.1 | -0.065 | 0.49 | 0.0012 | 1.1 | 1.2 | -0.46 |
| 10 | -0.26 | 0.82 | -0.96 | 0.87 | -1.0 | -2.6 | -0.12 | 0.054 | -0.063 | 0.18 | -0.11 | -0.49 |

3、实验内容及说明

问题一：

使用上面表格中的数据进行 Parzen 窗估计和设计分类器。窗函数为一个球形的高斯函数如下所示：

$$\varphi\left(\frac{(x-x_i)}{h}\right) \propto exp\left[-(x-x_i)^t(x-x_i)/\left(2h^2\right)\right]$$

编写程序，使用 Parzen 窗估计方法对任意一个的测试样本点 $x$ 进行分类。对分类器的训练则使用表格中的三维数据。令 h = 1，分类样本点为$(0.5,1.0,0.0)^t$，$(0.41,0.82,0.88)^t$，$(0.3,0.44,-0.1)^t$ 。

问题二：

对上面表格中的数据使用 k 近邻方法进行概率密度估计：

1. 编写程序，对于一维的情况，当有 n 个数据样本点时，进行 k-近邻概率密度估计。对表格中的类 1 的特征$x_1$，用程序画出当 k=1，3，5 时的概率密度估计结果。

2. 编写程序，对于二维的情况，当有 n 个数据样本点时，进行 k-近邻概率密度估计。对表格中的类 2 的特征$(x_1,x_2)^t$，用程序画出当 k=1，3，5 时的概率密度估计结果。

3. 编写程序，对表格中的 4 个类别的三维特征，使用 k-近邻概率密度估计方法。并且对下列点处的概率密度进行估计： $(0.14,0.72,4.1)^t$，$(-0.81,0.61,-0.38)^t$，$(0.31,1.51,-0.50)^t$。

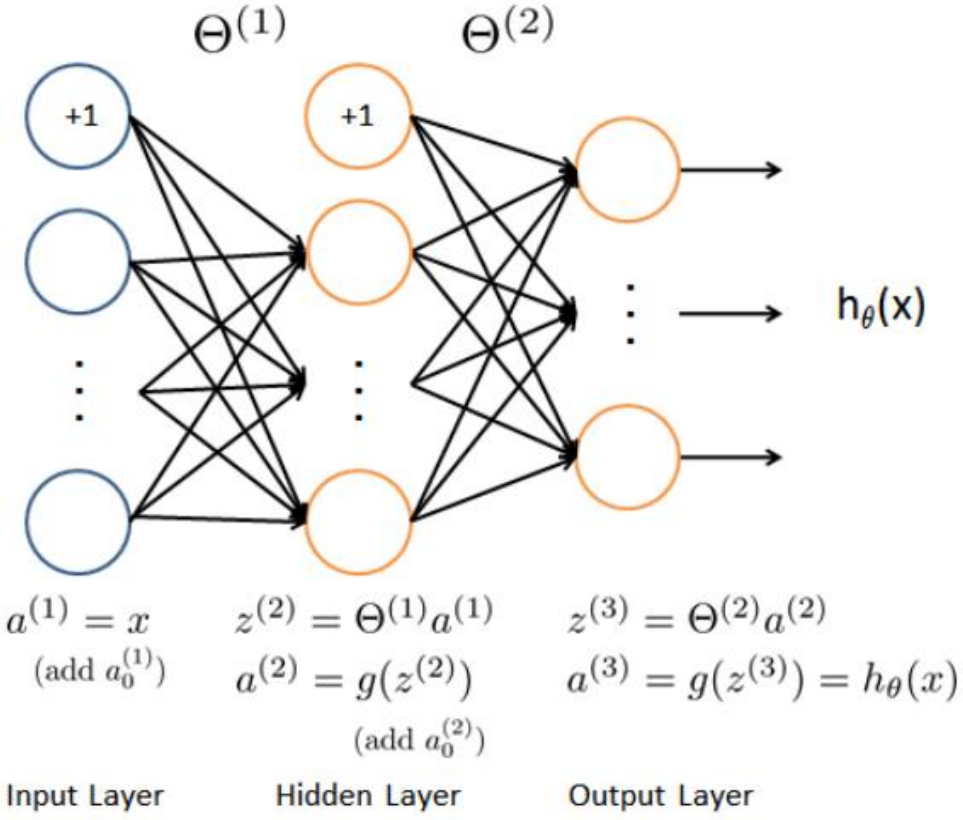
# 实验 4 神经网络学习

1. 实验目的
   （1） 掌握 BP 神经网络的基本原理和基本的设计步骤
   （2） 了解 BP 算法中各参数的作用和意义
2. 实验数据

lms.mat 手写数字数据集，数据集中包含 5000 个训练样本，其中每个训练训练样本都是 20×20 像素的灰度图像的数字。每个像素由一个浮点数表示，表示该位置的灰度强度。

特别规定：数字 0 被标记为 10，而数字 1-9 按照自然顺序被标记为 1-9；

3. 实验内容及说明
   （1） 用神经网络对给定的数据集进行分类；
   （2） 不能使用 TensorFlow 等框架，也不能使用库函数，所有算法都要自己实现；
   （3） 神经网络结构图如下图所示：




整个神经网络包括 3 层——输入层，隐藏层，输出层。输入层有 400 个神经元，隐藏层有 25 个神经元，输出层有 10 个神经元（对应 10 个数字类型）。

（4）附加：可以试着修改神经元数，层数，学习率等参数探究参数对实验结果的影响。

# 实验 5 集成学习

1. 实验目的

   用集成方法对数据集进行分类

2. 实验数据

   实验 4 中的 lms.mat 手写数字数据集

3. 实验内容及说明

   （1） 利用若干算法，针对同一样本数据训练模型，使用投票机制，少数服从多数，用多数算法给出的结果当作最终的决策依据；

   （2） 所选算法包括：

   SVM（核函数为多项式核函数）；

   KNN（k=7）；

   神经网络。

注：实验 4 中的神经网络模型可以使用，也可以使用框架。