13 高斯混合模型

Gaussian Mixture Model

组合若干高斯分布，期望最大化

每当竭力厘清某一数学话题后，我便径直离开，投身另一处昏暗角落；

孜孜以求的人如此奇怪，求解一个问题后，他不会自我陶醉、故步自封，而是踏上新的旅程。

When I have clarified and exhausted a subject, then I turn away from it, in order to go into darkness again; the never satisfied man is so strange if he has completed a structure, then it is not in order to dwell in it peacefully, but in order to begin another.

—— 卡尔·弗里德里希·高斯 (Carl Friedrich Gauss) | 德国数学家、物理学家、天文学家 | 1777 ~ 1855

- `matplotlib.patches.Ellipse()` 绘制椭圆
- `numpy.arctan2()` 输入正切值分子分母两个数，输出为反正切，值域为 `[-pi, pi]`
- `numpy.linalg.eigh ()` 返回实对称矩阵的特征值和特征向量
- `numpy.linalg.norm()` 默认 L2 范数
- `numpy.linalg.svd()` SVD 分解函数
- `plt.quiver()` 绘制箭头图
- `seaborn.barplot()` 绘制直方图
- `sklearn.mixture.GaussianMixture` 高斯混合模型聚类函数

13.1 高斯混合模型

高斯混合模型 (Gaussian Mixture Model, GMM) 是一种常用的无监督机器学习算法，它的核心思维是——用多个高斯密度函数估计样本数据分布。高斯混合模型是一种概率模型，它假定所有数据点都是由有限个参数未知的高斯分布混合产生。

某种意义上讲，高斯混合模型是 K 均值聚类的推广。高斯混合模型和 K 均值聚类都是采用迭代方法求解优化问题。K 均值利用簇质心，最小化簇内残差平方和 SSE；而高斯混合模型利用簇质心和协方差，最大化对数似然函数。

此外，高斯混合模型和本书监督学习部分讲解的贝叶斯分类和高斯判别分析联系紧密。

前文说过，高斯混合模型是若干个高斯分布混合；下面分别一元和二元高斯分布来介绍这以思想。

一元高斯分布混合

大家对一元高斯分布概率密度函数 $f_X(x)$ 再熟悉不过：

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right) \tag{1}$$

其中，μ 为均值/期望值，σ 为标准差。对于一元高斯分布，给定 μ 和 σ 就确定分布形状。

对于单一特征样本数据，高斯混合模型的意义就是利用若干一元高斯分布来描述样本分布。

图 1 给出的是鸢尾花样本数据花萼长度和花萼宽度两个特征。分别观察这两个特征，可以发现用单一高斯分布都不能准确描述数据的边际分布，但是组合三个高斯分布却可以描述数据特征分布。

注意，对于无监督学习，样本数据没有标签，即并不知道样本数据的类别。高斯混合模型算法通过一系列运算估计预测样本数据类别。

通常，称高斯混合模型 GMM 每一高斯分布为一个**分量** (component)。对于一元高斯分布，高斯混合分布算法难点就是确定每个分量各自的参数，μ 和 σ。

本书前文在贝叶斯分类部分介绍过，根据全概率定理，$C_1, C_2, \ldots, C_K$ 为一组不相容分类，对样本空间 Ω 形成分割，下式成立：

$$f_X(x) = \sum_{k=1}^{K} \underbrace{f_{Y,X}(C_k, x)}_{\text{Joint}} \tag{2}$$

根据贝叶斯定理，证据因子、后验概率和联合概率存在如下关系：

$$\underbrace{f_{Y,X}(C_k, X)}_{\text{Joint}} = \underbrace{f_{X|Y}(x|C_k)}_{\text{Likelihood}} \underbrace{p_Y(C_k)}_{\text{Prior}} \tag{3}$$

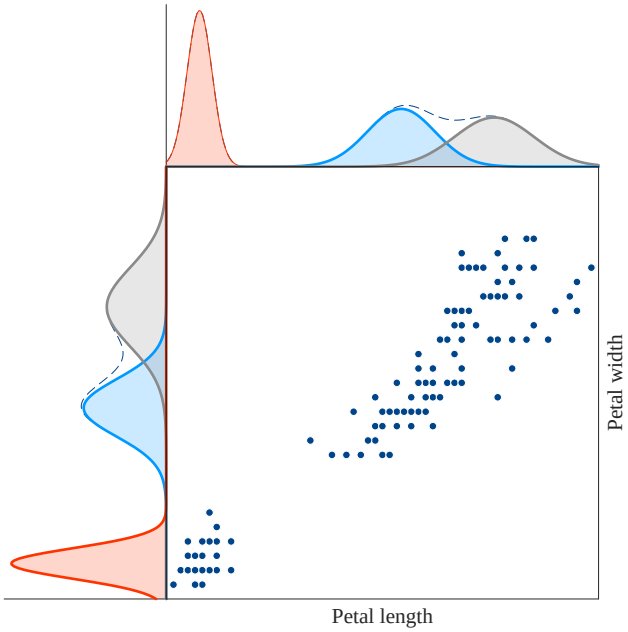


图 1. 用三个高斯一元分布描述样本数据边际分布

将 (3) 代入 (2) 得到：

$$f_X(x) = \sum_{k=1}^{K} \underbrace{f_{X|Y}(x|C_k)}_{\text{Likelihood}} \underbrace{p_Y(C_k)}_{\text{Prior}} \tag{4}$$

对于高斯混合模型 GMM，$f_{X|Y}(x|C_k)$ 为**似然概率** (likelihood) 用高斯分布描述，$p_Y(C_k)$ 为**先验概率** (prior)，表达样本集合中 C_k 类样本占比。

对于无监督学习，样本数据标签未知；因此，高斯混合模型 GMM 迭代过程中，似然概率 $f_{X|Y}(x|C_k)$ 和先验概率 $p_Y(C_k)$ 不断估算更新，直到满足迭代停止条件。

而对于有监督学习，样本标签数据已知，即 C_k 确定；比如，高斯朴素贝叶斯算法，直接就可以估算似然概率 $f_{X|Y}(x|C_k)$ 和先验概率 $p_Y(C_k)$。

对于单一特征样本数据，且 $K = 3$，图 2 对应的边际分布 $p_Y(C_k)$ 可以用三个一元高斯分布叠加获得：

$$\begin{aligned} f_X(x) &= \underbrace{p_Y(C_1)}_{\text{Prior}} \underbrace{f_{X|Y}(x|C_1)}_{\text{Likelihood}} + \underbrace{p_Y(C_2)}_{\text{Prior}} \underbrace{f_{X|Y}(x|C_2)}_{\text{Likelihood}} + \underbrace{p_Y(C_3)}_{\text{Prior}} \underbrace{f_{X|Y}(x|C_3)}_{\text{Likelihood}} \\ &= \alpha_1 N(x, \mu_1, \sigma_1) + \alpha_2 N(x, \mu_2, \sigma_2) + \alpha_3 N(x, \mu_3, \sigma_3) \\ &= \underbrace{\alpha_1 \frac{\exp\left(-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right)}{\sigma_1\sqrt{2\pi}}}_{C_1} + \underbrace{\alpha_2 \frac{\exp\left(-\frac{1}{2}\left(\frac{x-\mu_2}{\sigma_2}\right)^2\right)}{\sigma_2\sqrt{2\pi}}}_{C_2} + \underbrace{\alpha_3 \frac{\exp\left(-\frac{1}{2}\left(\frac{x-\mu_3}{\sigma_3}\right)^2\right)}{\sigma_3\sqrt{2\pi}}}_{C_3} \end{aligned} \tag{5}$$

如图 2 所示，μ_1、μ_2 和 μ_3 为期望值，描述三个正态分布质心位置；σ_1、σ_2 和 σ_3 为标准差，刻画三个正态分布的离散程度；而先验概率 α_1、α_2 和 α_3 给出三个正态分布对 $f_X(x)$ 的贡献。

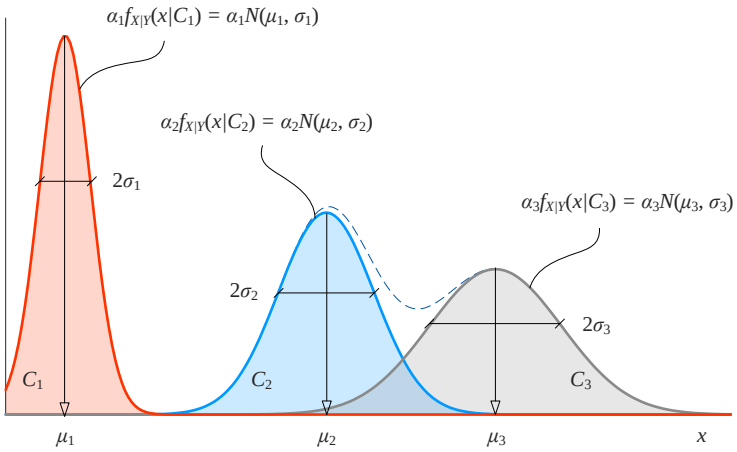


图 2. 三个一元高斯分布重要统计描述量

令：

$$\boldsymbol{\theta} = \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \mu_1 & \mu_2 & \mu_3 & \sigma_1 & \sigma_2 & \sigma_3 \end{bmatrix} \tag{6}$$

三个一元高斯分布叠加产生的高斯混合分布记做 $f_X(x \mid \boldsymbol{\theta})$。

$$f(x \mid \boldsymbol{\theta}) = p_Y(C_1) f_{X|Y}(x|C_1, \boldsymbol{\theta}) + p_Y(C_2) f_{X|Y}(x|C_2, \boldsymbol{\theta}) + p_Y(C_3) f_{X|Y}(x|C_3, \boldsymbol{\theta}) \tag{7}$$

多元高斯分布混合

下面考虑样本数据多特征情况。C_k 类数据条件概率 $f_{X|Y}(x|C_k)$ 分布服从多元高斯分布：

$$f_{X|Y}(x|C_k) = \frac{\exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_k)^{\mathrm{T}} \boldsymbol{\Sigma}_k^{-1} (\boldsymbol{x}-\boldsymbol{\mu}_k)\right)}{\sqrt{(2\pi)^D |\boldsymbol{\Sigma}_k|}} \tag{8}$$

其中，D 为特征数量，即多元高斯分布维数；$\boldsymbol{x}$ 为列向量，$\boldsymbol{\mu}_k$ 为 C_k 类簇质心位置，即期望值/均值；$\boldsymbol{\Sigma}_k$ 为 C_k 类数据协方差矩阵，刻画正态分布离散程度和相关性。

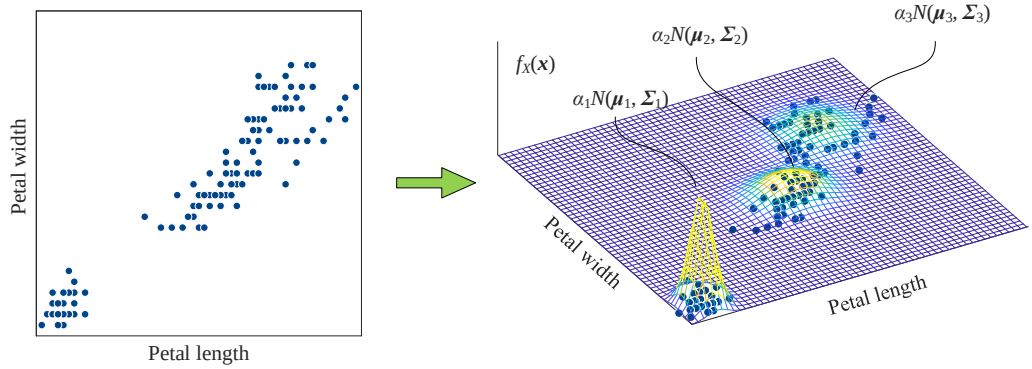


图 3. 三个二元高斯分布叠加描述鸢尾花数据分布

图 3 展示的鸢尾花花萼长度和宽度样本数据分布。显然，样本数据不适合用一个二元高斯分布，也不能用两个二元高斯分布叠加得。但是，每个高斯分布描述一簇数据，采用三个高斯分布叠加就可以比较准确的描述数据分布情况：

$$\begin{aligned} f(\boldsymbol{x}\,|\,\boldsymbol{\theta}) &= p_Y(C_1)\, f_{\boldsymbol{X}|Y}(\boldsymbol{x}|C_1,\boldsymbol{\theta}) + p_Y(C_2)\, f_{\boldsymbol{X}|Y}(\boldsymbol{x}|C_2,\boldsymbol{\theta}) + p_Y(C_3)\, f_{\boldsymbol{X}|Y}(\boldsymbol{x}|C_3,\boldsymbol{\theta}) \\ &= \alpha_1 \underbrace{N(\boldsymbol{x},\boldsymbol{\mu}_1,\boldsymbol{\Sigma}_1)}_{C_1} + \alpha_2 \underbrace{N(\boldsymbol{x},\boldsymbol{\mu}_2,\boldsymbol{\Sigma}_2)}_{C_2} + \alpha_3 \underbrace{N(\boldsymbol{x},\boldsymbol{\mu}_3,\boldsymbol{\Sigma}_3)}_{C_3} \end{aligned} \tag{9}$$

定义参数 $\boldsymbol{\theta}$ 为：

$$\boldsymbol{\theta} = \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \boldsymbol{\mu}_1 & \boldsymbol{\mu}_2 & \boldsymbol{\mu}_3 & \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_3 \end{bmatrix} \tag{10}$$

再次强调，作为无监督学习，样本数据标签未知；高斯混合模型 GMM 通过迭代求解优化问题，迭代过程，参数 $\boldsymbol{\theta}$ 不断更新。当迭代收敛时，参数 $\boldsymbol{\theta}$ 更新变化平缓。因此，(10) 定义的 $\boldsymbol{\theta}$，实际上是某一轮迭代时参数估计的快照。

后验概率

根据贝叶斯定理，计算后验概率：

$$f_{Y|\boldsymbol{X}}(C_k\,|\,\boldsymbol{x},\boldsymbol{\theta}) = \frac{p_Y(C_k)\, f_{\boldsymbol{X}|Y}(\boldsymbol{x}|C_k,\boldsymbol{\theta})}{f_{\boldsymbol{X}}(\boldsymbol{x},\boldsymbol{\theta})} = \frac{p_Y(C_k)\, f_{\boldsymbol{X}|Y}(\boldsymbol{x}|C_k,\boldsymbol{\theta})}{\sum_{k=1}^{K} p_Y(C_k)\, f_{\boldsymbol{X}|Y}(\boldsymbol{x}|C_k,\boldsymbol{\theta})} \tag{11}$$

由 K 个高斯分布构造的混合分布函数如下所示：

$$\begin{aligned} f_{\boldsymbol{X}}(\boldsymbol{x},\boldsymbol{\theta}) &= \sum_{k=1}^{K} p_Y(C_k)\, f_{\boldsymbol{X}|Y}(\boldsymbol{x}|C_k,\boldsymbol{\theta}) \\ &= \sum_{k=1}^{K} \alpha_k N(\boldsymbol{\mu}_k,\boldsymbol{\Sigma}_k) \end{aligned} \tag{12}$$

其中，第 i 个高斯分布参数有两个，分别是均值向量 $\boldsymbol{\mu}_k$ 和协方差矩阵 $\boldsymbol{\Sigma}_k$。α_k 为混合系数，是混合成分的后验概率，$\alpha_i > 0$。

参数 $\boldsymbol{\theta}$ 定义为：

$$\boldsymbol{\theta} = \left\{ \alpha_k, \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k \right\} \quad k = 1, 2, ..., K \tag{13}$$

K 个混合系数之和为 1：

$$\sum_{k=1}^{K} \alpha_k = 1 \tag{14}$$

高斯混合模型，分量数量 K 是一个用户输入值。本章后续会介绍如何选取合适分量数量 K。

三聚类

假设数据聚类为 C_1、C_2 和 C_3 三类，后验概率 $f_{\chi|Y}\left(\boldsymbol{x}|C_1,\boldsymbol{\theta}\right)$、$f_{\chi|Y}\left(\boldsymbol{x}|C_2,\boldsymbol{\theta}\right)$ 和 $f_{\chi|Y}\left(\boldsymbol{x}|C_3,\boldsymbol{\theta}\right)$ 可以通过下式获得：

$$\begin{cases} f_{Y|\chi}\left(C_1|\boldsymbol{x},\boldsymbol{\theta}\right) = \dfrac{p_Y\left(C_1\right) f_{\chi|Y}\left(\boldsymbol{x}|C_1,\boldsymbol{\theta}\right)}{f_{\chi}\left(\boldsymbol{x},\boldsymbol{\theta}\right)} \\ f_{Y|\chi}\left(C_2|\boldsymbol{x},\boldsymbol{\theta}\right) = \dfrac{p_Y\left(C_2\right) f_{\chi|Y}\left(\boldsymbol{x}|C_2,\boldsymbol{\theta}\right)}{f_{\chi}\left(\boldsymbol{x},\boldsymbol{\theta}\right)} \\ f_{Y|\chi}\left(C_3|\boldsymbol{x},\boldsymbol{\theta}\right) = \dfrac{p_Y\left(C_3\right) f_{\chi|Y}\left(\boldsymbol{x}|C_3,\boldsymbol{\theta}\right)}{f_{\chi}\left(\boldsymbol{x},\boldsymbol{\theta}\right)} \end{cases} \tag{15}$$

其中，

$$f_{\chi}\left(\boldsymbol{x},\boldsymbol{\theta}\right) = p_Y\left(C_1\right) f_{\chi|Y}\left(\boldsymbol{x}|C_1,\boldsymbol{\theta}\right) + p_Y\left(C_2\right) f_{\chi|Y}\left(\boldsymbol{x}|C_2,\boldsymbol{\theta}\right) + p_Y\left(C_3\right) f_{\chi|Y}\left(\boldsymbol{x}|C_3,\boldsymbol{\theta}\right) \tag{16}$$

图 4 所示后验概率 $f_{Y|\chi}\left(C_1|\boldsymbol{x},\boldsymbol{\theta}\right)$、$f_{Y|\chi}\left(C_2|\boldsymbol{x},\boldsymbol{\theta}\right)$ 和 $f_{Y|\chi}\left(C_3|\boldsymbol{x},\boldsymbol{\theta}\right)$ 三曲面。比较这三个曲面高度便可以确定预测聚类区域。高斯混合模型迭代过程，$f_{Y|\chi}\left(C_1|\boldsymbol{x},\boldsymbol{\theta}\right)$、$f_{Y|\chi}\left(C_2|\boldsymbol{x},\boldsymbol{\theta}\right)$ 和 $f_{Y|\chi}\left(C_3|\boldsymbol{x},\boldsymbol{\theta}\right)$ 三曲面形状不断变化，决策边界也不断变化。

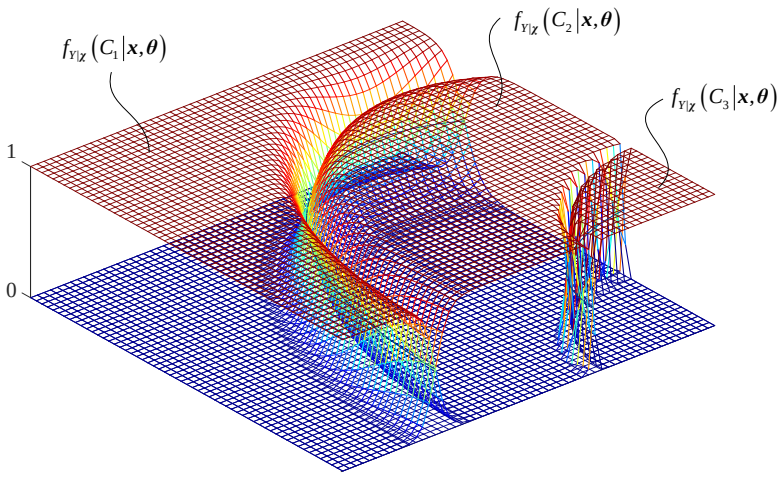


图 4. GMM 模型下后验概率曲面

给定无标记样本数据，可以采用高斯混合模型对数据进行聚类；类似贝叶斯分类，后验概率可以判定聚类决策边界。因此高斯混合模型聚类这个问题的优化目标，便是找到满足条件的参数 $\boldsymbol{\theta}$。下一节，我们将采用**最大期望算法** (expectation maximization, EM)，简称 **EM 算法**，解决这一问题。而下一章将专门讲解最大期望算法。

13.2 四类协方差矩阵

多元高斯分布用来刻画 C_k 类数据条件概率 $f_{X|Y}(\boldsymbol{x} \mid C_k)$；而多元高斯分布中，协方差矩阵 $\boldsymbol{\Sigma}_k$ 决定高斯分布的形状。本书前文在高斯判别分析 GDA 中介绍过六类 GDA，这六类 GDA 中协方差矩阵 $\boldsymbol{\Sigma}_k$ 各有特点。

如表 1 总结，scikit-learn 工具包中 `sklearn.mixture` 高斯混合模型支持四种协方差矩阵——`tied` （平移）、`spherical` （球面）、`diag` （对角）和 `full` (完全)。

`tied` 指的是，所有分量共享一个非对角协方差矩阵 $\boldsymbol{\Sigma}$；`tied` 类似第三类高斯判别分析。每个分量 PDF 等高线为大小相等旋转椭圆。根据本书前文分析，由于不同分量协方差相同，决策边界解析式二次项消去；因此 `tied` 对应的决策边界为直线。

`spherical` 指的是，每个分量协方差矩阵 $\boldsymbol{\Sigma}_k$ 不同，但是每个分量 $\boldsymbol{\Sigma}_k$ 均为对角阵；且 $\boldsymbol{\Sigma}_k$ 对角元素相同，即特征方差相同；`spherical` 类似第四类高斯判别分析。每个分量 PDF 等高线为正圆。`spherical` 对应的决策边界为圆形弧线。

`diag` 指每个分量有各自独立的对角协方差矩阵，也就是 $\boldsymbol{\Sigma}_k$ 为对角阵，特征条件独立；但是对 $\boldsymbol{\Sigma}_k$ 对角线元素大小不做限制。`diag` 对应第五类高斯判别分析。每个分量 PDF 等高线正椭圆，`diag` 对应的决策边界为正圆锥曲线。

`full` 指每个分量有各自独立协方差矩阵，即对 $\boldsymbol{\Sigma}_k$ 不做任何限制。`full` 对应第六类高斯判别分析。`full` 对应的决策边界为任意圆锥曲线。

表 1. 根据方差-协方差矩阵特点将高斯混合模型分为 4 类

参数设置	$\boldsymbol{\Sigma}_i$	$\boldsymbol{\Sigma}_i$ 特点	PDF 等高线	决策边界
tied (第三类)	相同	非对角阵	任意椭圆	直线
spherical (第四类)	不相同	对角阵，对角线元素等值	正圆	正圆
diag (第五类)		对角阵	正椭圆	正圆锥曲线
full (第六类)		非对角阵	任意椭圆	圆锥曲线

和 K 均值聚类算法一样，高斯混合模型 GMM 也需要指定 K 值；高斯混合模型也是利用迭代求解优化问题。不同的是，GMM 利用协方差矩阵，可以估算后验概率/成员值。GMM 的协方差矩阵有四种类型，每种类型对应不同假设，获得不同决策边界类型。

K 均值聚类可以看作是高斯混合模型一个特例。如图 5 所示，K 均值聚类对应的 GMM 特点是，各簇协方差矩阵 $\boldsymbol{\Sigma}_k$ 相同，$\boldsymbol{\Sigma}_k$ 为对角阵，并且 $\boldsymbol{\Sigma}_k$ 主对角线元素相等。

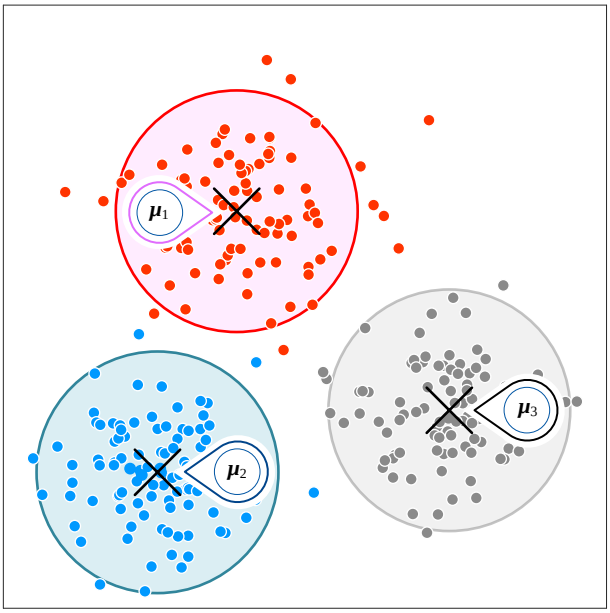


图 5. K 均值聚类可以看作是高斯混合模型一个特例

以鸢尾花数据为例

下面，我们分别利用 `sklearn.mixture` 四种协方差矩阵设置，比较鸢尾花数据的聚类结果。图 6 中，GMM 的协方差矩阵设置为 `tied`；容易发现获得的决策边界为直线，这是因为所有分量公用一个非对角协方差矩阵。图 7 中，GMM 的协方差矩阵设置为 `spherical`；对应的决策边界显然为三段圆弧构造。图 8 中，GMM 的协方差矩阵设置为 `diag`；图 8 中椭圆弧线长度较短，不容易直接判断它们对应的椭圆是否为正圆锥曲线。图 9 中，GMM 的协方差矩阵设置为 `full`；决策边界为任意圆锥曲线。读者可以回顾本书高斯判别分析中有关决策边界形态内容。

另外，图 6~图 9 这四幅图中，还给出高斯分布椭圆等高线的半长轴和半短轴向量指向。

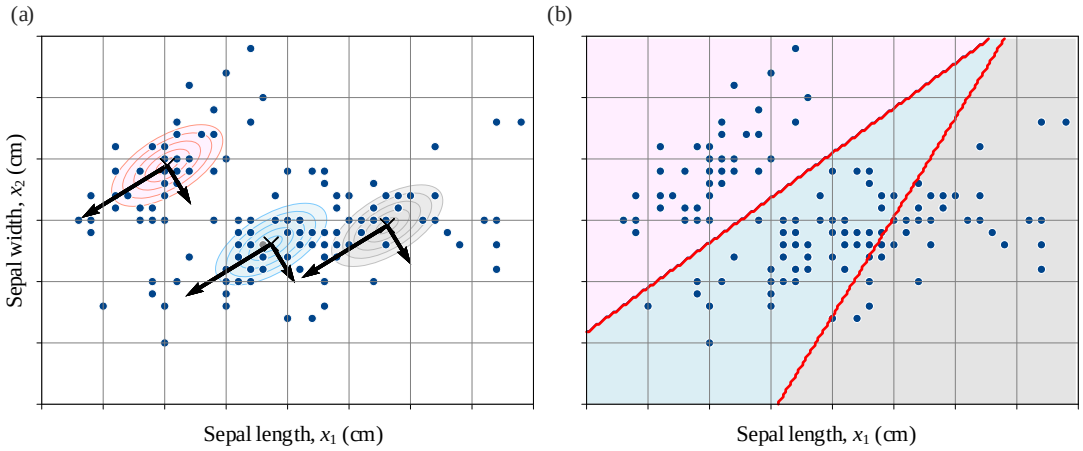


图 6. GMM 的协方差矩阵设置为 `tied`

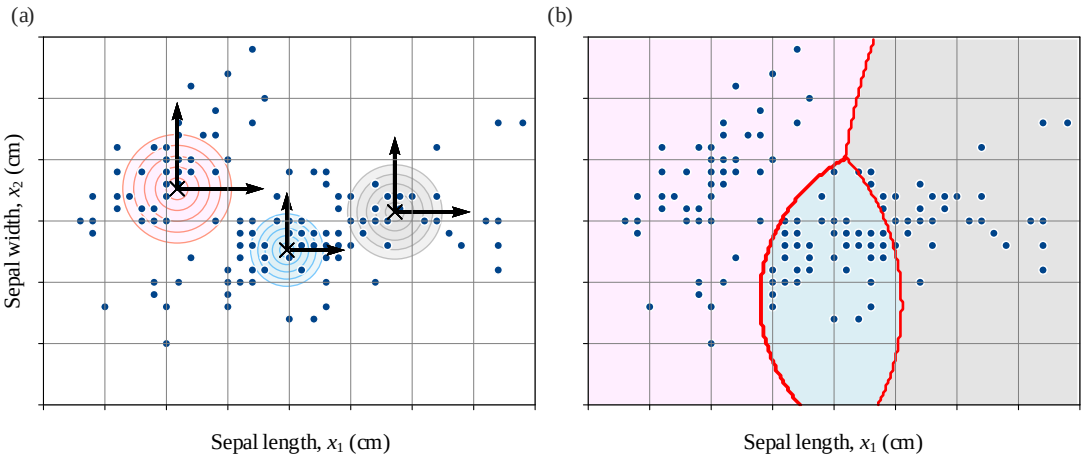


图 7. GMM 的协方差矩阵设置为 `spherical`

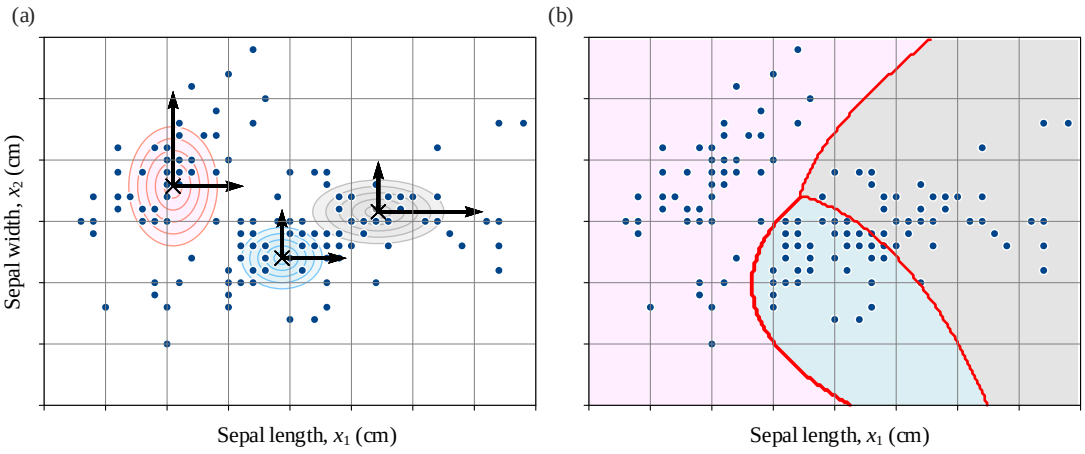


图 8. GMM 的协方差矩阵设置为 `diag`

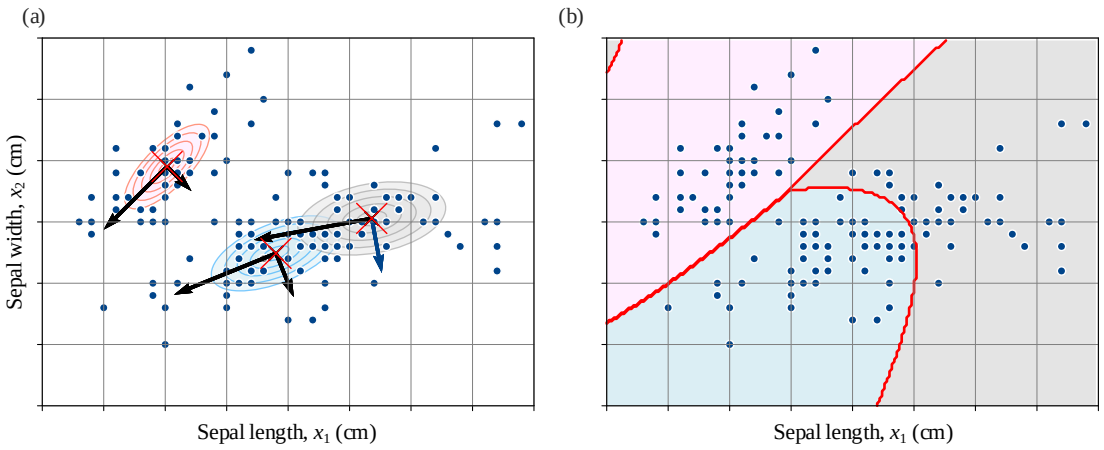
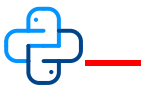


图 9. GMM 的协方差矩阵设置为 `full`



代码 Bk7_Ch12_01.py 完成本节分类问题。

13.3 分量数量

前文介绍，高斯混合模型 GMM 的分量数量 K 是用户输入值。选取合适 K 值，对于 GMM 聚类效果至关重要。本节介绍采用 AIC 和 BIC 选择高斯混合模型分量数量。

赤池信息量准则

丛书读者对 AIC 和 BIC 并不陌生，在《数据科学》一册回归相关内容，我们已经了解过这两个指标。AIC 为**赤池信息量准则** (Akaike information criterion, AIC)，定义如下：

$$\mathrm{AIC} = \underbrace{2K}_{\text{Penalty}} - 2\ln(L) \tag{17}$$

其中，K 是分量数量，即聚类数量；L 是似然函数。

Scikit-learn 工具包中 AIC 计算形式稍有不同。AIC 鼓励数据拟合的优良性；但是，尽量避免出现过度拟合。(17) 中 $2K$ 项为**惩罚项** (penalty)。

贝叶斯信息准则

贝叶斯信息准则 (Bayesian Information Criterion, BIC) 也称**施瓦茨信息准则** (Schwarz information criterion, SIC)，定义如下：

$$\mathrm{BIC} = \underbrace{K\ln(n)}_{\text{Penalty}} - 2\ln(L) \tag{18}$$

其中，n 为样本数据数量。BIC 的惩罚项比 AIC 大。

图 10 所示为三簇数据构成的样本数据。采用高斯混合模型聚类算法，K 取不同值 (K = 1, 2, …, 6)，协方差矩阵分别采用前文介绍的四种设置——`tied` （平移）、`spherical` （球面）、`diag` （对角）和 `full` (完全)。对于这 24 种组合，我们取出对应模型 AIC 和 BIC 结果。

图 11 所示为 AIC 随协方差形状和分量数变化直方图。图 12 所示为 BIC 随协方差形状和分量数变化直方图。可以发现，24 中设置中，`spherical` (球面) 和 K = 3 参数组合对图 10 所示样本数据聚类效果最好。

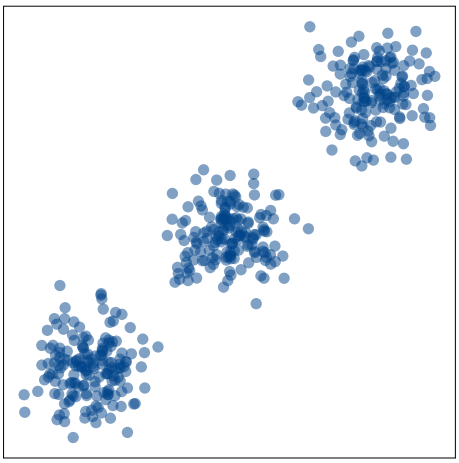


图 10. 三簇数据构成的样本数据

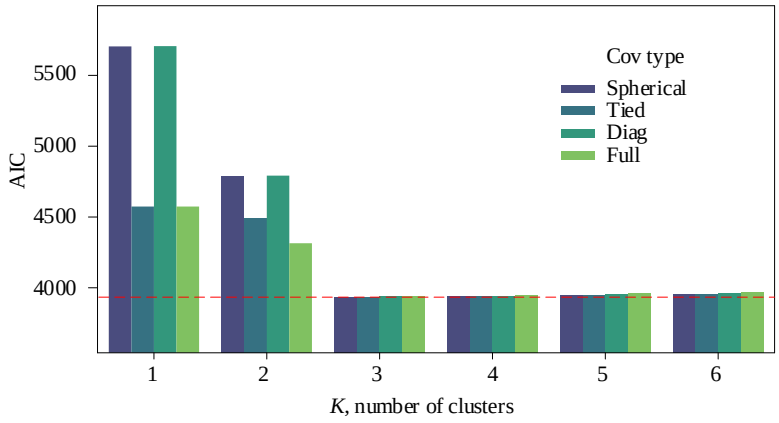


图 11. AIC 随协方差形状和分量数变化

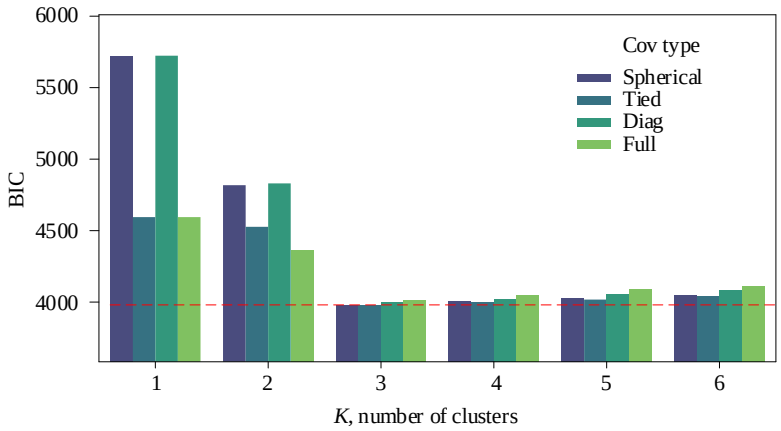


图 12. BIC 随协方差形状和分量数变化

代码 Bk7_Ch12_02.py 绘制本节图像。

13.4 硬聚类和软聚类

本书朴素贝叶斯分类算法中提到，后验概率相当于成员值。**硬聚类** (hard clustering) 指的根据成员值大小，决策边界清楚划定；但是**软聚类** (soft clustering) 则设定缓冲带，当后验概率/成员值在这个缓冲带内，样本数据没有明确的聚类。这样，软聚类的决策边界不再“泾渭分明”，而变成了一条宽带。

给定如图 13 所示 450 个样本数据。利用高斯混合模型算法获得 $f_{Y|X}(C_1 \mid \boldsymbol{x})$ 和 $f_{Y|X}(C_2 \mid \boldsymbol{x})$ 两后验概率曲面，如图 14 所示。

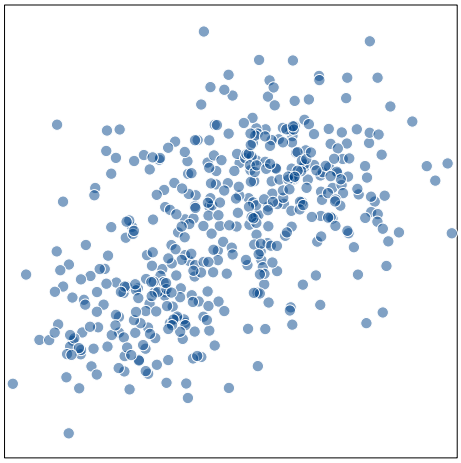


图 13. 样本数据

如图 14 所示，以成员值大小排列这 450 个样本数据；对于二聚类问题，硬聚类以后验概率 0.5 为分界线。当 $f_{Y|X}(C_1|\boldsymbol{x}) = 0.5$ 对应着决策边界；当 $f_{Y|X}(C_1|\boldsymbol{x}) > 0.5$，$\boldsymbol{x}$ 被聚类到 C_1 簇；当 $f_{Y|X}(C_1|\boldsymbol{x}) < 0.5$，$\boldsymbol{x}$ 被聚类到 C_2 簇。

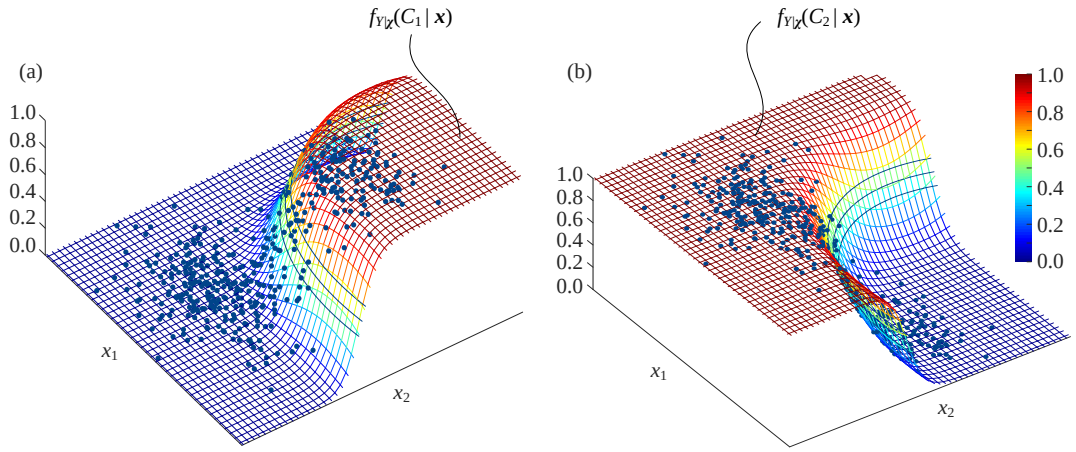


图 14. $f_{Y|X}(C_1|\boldsymbol{x})$ 和 $f_{Y|X}(C_2|\boldsymbol{x})$ 两后验概率曲面

软聚类

而对于软聚类，后验概率在一段阈值内，比如 [0.3, 0.7]，数据没有明确的分类。图 16 所示为聚类结果，加黑圈的样本数据，位于“决策带”之内，没有明确预测分类。

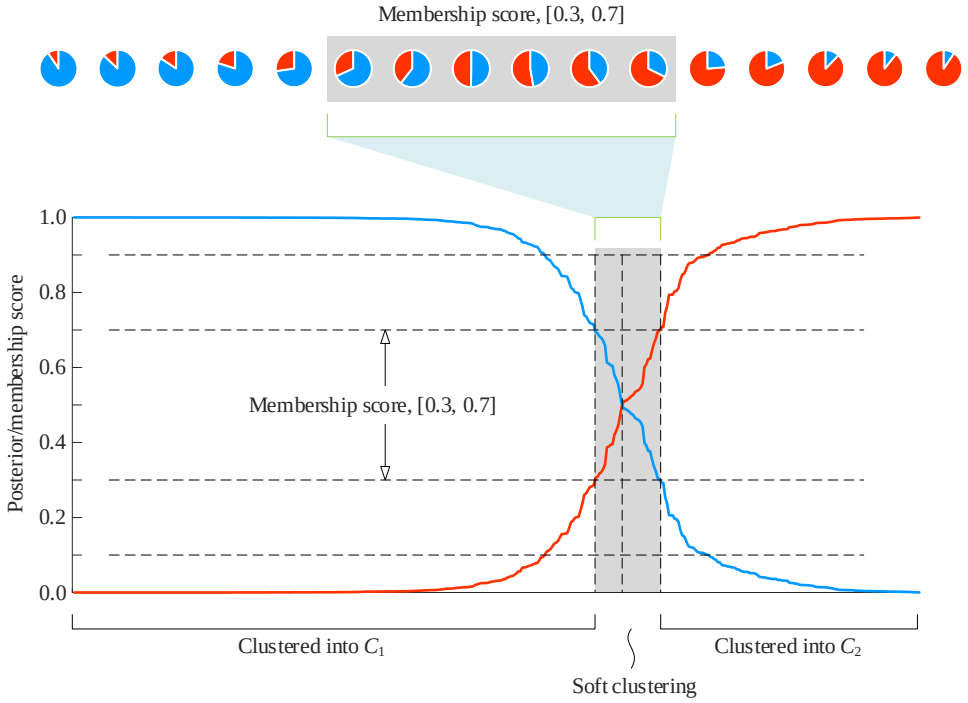


图 15. 成员值与软聚类

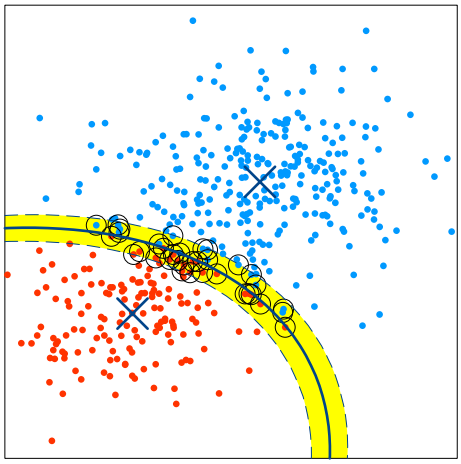


图 16. 软聚类分区和决策带