УДК 519.86

Градиентно-статистический алгоритм.

А.А.Сумкин
Российский университет дружбы народов

Конечные гауссовы смеси широко используются в качестве генерального аппроксиматора при моделирование реальных процессов и явлений в различных областях науки и практики[1-3]. Широкий спектр применения гауссовых смесей вызывает необходимость решения таких задач , как предварительное оценивание числа их мод и компонент , разработка алгоритмов вычисления мод и оптимальных оценок параметров распределения.

Цель работы составить алгоритм с помощью которого можно обнаружить все критические точки .

Пусть плотность вероятности смеси задана выражением :

$$f(x) = (2\pi)^{-\frac{p}{2}} \sum_{i=1}^{k} \pi_i \left|\sum_i\right|^{-\frac{1}{2}} e^{-\frac{1}{2}(X-\mu_i)\sum_i^{-1}(X-\mu_i)'} \quad ,$$

k- число компонент

π_i-априорные вероятности

p- размерность пространства

$\sum_i$ -ковариационная матрица

По заданному закону распределения вероятностей моделируется p-мерная репрезентативная выборка , n независимых случайных величин .

$$A = \{ X_1, X_2, \ldots, X_n\} \quad 1 \le p < \infty$$

$$X_s = (x_{s1}, x_{s2}, \ldots, x_{sp}) \quad s=1,2\ldots n$$

Для всех заданных независимых случайных векторов смеси находим модуль градиента.

$$|\nabla s| = \left(\sum_{i=1}^{p} \left(\frac{\partial f(X_s)}{\partial X_i}\right)^2\right)^{\frac{1}{2}} ,s=1,..,n$$

Для векторов у которых модуль градиента будет меньше заданного положительного числа , считается Гессиан (матрица вторых производных).

$$H = \nabla'\nabla f(X) = \sum_{i=1}^{k} f_i(X) \sum_i^{-1} \left((\mu_i - X)'(\mu_i - X) - \sum_i\right) \sum_i^{-1}$$

Разбив на окрестности и задав условия для этого Гессиана, найдем все критические точки . Результатом работы стало получение алгоритма, находящего все критические точки. Экспериментальная проверка ГСА дала положительный результат.

Литература

1. Titterington D.M., Smith A.F.M., and Makov U.E. Statistical Analysis of Finite Mixture Distributions. JornWiley, 1985. Pp 53-147

2. Волошин Г.Я., Бурлаков И.А., Косенкова С.Т. Статистические методы решения задач распознавания, основанные на аппроксимационном подходе. Владивосток , 1992. Ч.1.271с.

3.Фомин Я.А., Тарловский Г.Р. Статистическая теория распознавания образов . М. Радио и связь , 1986 . 264 с.