

Исследование непараметрических алгоритмов оценивания плотности распределения случайной величины

Цель работы

Исследовать алгоритмы оценивания плотности распределения случайных величин и случайных векторов на основе методов Парзена и k ближайших соседей.

Форма контроля

Письменный отчёт (допускается представление в электронном виде). Опрос в устной форме в соответствии с перечнем контрольных вопросов.

Количество отведённых аудиторных часов

4

Содержание работы

Получить у преподавателя вариант задания и написать код, реализующий соответствующий алгоритм обработки информации. Для ответа на поставленные в задании вопросы провести численный эксперимент или статистическое имитационное моделирование и представить соответствующие графики. Провести анализ полученных результатов и представить его в виде выводов по проделанной работе.

Пример варианта задания

1. Вычислить абсолютную ошибку оценивания плотности распределения случайной величины при использовании оценки Парзена. Построить график зависимости ошибки оценивания от величины параметра оконной функции следующего вида:
 - a. гауссовская функция;
 - b. показательная функция;
 - c. оконная прямоугольная функция;
 - d. оконная треугольная функция.
2. Вычислить среднеквадратичную ошибку оценивания плотности распределения случайной величины по методу Парзена для окна вида:

a.
$$\frac{1}{h} \phi\left(\frac{x - x^{(i)}}{h}\right) = \frac{1}{\pi h} \frac{1}{1 + \left((x - x^{(i)}) / h\right)^2};$$

b.
$$\frac{1}{h} \phi\left(\frac{x - x^{(i)}}{h}\right) = \frac{1}{2\pi h} \left(\frac{\sin\left((x - x^{(i)}) / (2h)\right)}{(x - x^{(i)}) / (2h)} \right)^2.$$

3. Окно какого вида будет обеспечивать оптимальное по критерию среднеквадратичной ошибки оценивание плотности распределения случайной величины в соответствии с методом Парзена? Построить графики зависимостей ошибок от объема обучающей выборки. Сравните следующие виды окон:

- a. гауссовская функция, показательная функция и оконная прямоугольная функция;
 - b. гауссовская функция, оконная прямоугольная функция и оконная треугольная функция.
4. Вычислить абсолютную ошибку оценивания плотности распределения случайного вектора в двумерном пространстве признаков при использовании оценки Парзена. Построить график зависимости ошибки оценивания от величины параметра оконной функции. Используйте одну из следующих оконных функций:
- a. гауссовская функция с использованием диагональной матрицы;
 - b. гауссовская функция с использованием матрицы ковариаций;
 - c. показательная функция;
 - d. оконная прямоугольная функция;
 - e. оконная треугольная функция.
5. Окно какого вида будет обеспечивать оптимальное по критерию среднеквадратичной ошибки оценивание плотности распределения двумерного случайного вектора по методу Парзена? Построить графики зависимостей ошибок от объема обучающей выборки. Сравните следующие виды окон:
- a. гауссовская функция с использованием диагональной матрицы и гауссовская функция с использованием матрицы ковариаций;
 - b. гауссовская функция с использованием диагональной матрицы, показательная функция и оконная треугольная функция;
 - c. гауссовская функция с использованием матрицы ковариаций, оконная прямоугольная функция и оконная треугольная функция.
6. Получить оценки плотности распределения случайной величины на основе метода k ближайших соседей. Определить значение параметра k, при котором достигается минимум среднеквадратичной ошибки оценивания.
7. Реализовать оценку плотности распределения двумерного случайного вектора, плотность которого задается на основе пяти гауссовых функций. Определить оптимальное по критерию среднеквадратичной ошибки оценивания значение параметра k.

Примеры контрольных вопросов

1. При каком значении параметра оконной функции достигается минимум ошибки оценивания по критерию, указанному в варианте исполнителя? Ответ подтвердить графиками, представленными в отчёте.
2. Какой вид оконной функции обеспечивает оптимальную оценку плотности распределения в реализованном Вами исследовании? Ответ подтвердить графиками, представленными в отчёте.