Программа курса «Анализ данных, суррогатное моделирование и оптимизация в прикладных задачах»

2022 год, весенний семестр

1. Данные и методы работы с ними

Числовые характеристики выборки: среднее, медиана, среднеквадратичное отклонение, интерквартильный размах. Коэффициент корреляции Пирсона, ранговый коэффициент корреляции Спирмена.

2. Основы теории вероятностей

Аксиоматика теории вероятностей, вероятностное пространство. Условная вероятность, формула полной вероятности, формула Байеса. Отношение шансов и отношение правдоподобия. Оценка максимального правдоподобия.

3. Случайные величины и их распределения

Понятие случайной величины. Распределения случайной величины, многомерные распределения. Числовые характеристики случайных величин: математическое ожидание, дисперсия, ковариация, коэффициент корреляции.

4. Многомерное нормальное распределение

Гауссовские случайные величины, их свойства. Ковариационная матрица. Линейное преобразование многомерного нормального распределения, генерация выборки гауссовских векторов. Маргинальные и условные распределения.

5. Гауссовские случайные процессы

Случайные процессы: базовые понятие и примеры. Моментные характеристики случайных процессов. Гауссовские случайные процессы. Ковариационные функции. Генерация случайной выборки гауссовских процессов.

6. Регрессия на основе гауссовских процессов

Методы восстановление регрессии, параметрические и непараметрические модели. Ядерные методы. Построение регрессионной модели с помощью гауссовских процессов. Вычисление параметров ковариационной функции апостериорного процесса: апостериорное среднее и апостериорная дисперсия.

7. Оптимизация регрессионной кривой

Обучающая выборка с шумом. Влияние параметров ядра и амплитуды шума на регрессионную кривую. Оптимизация параметров ядра. Многомерный случай.

8. Байесовская оптимизация

Методы планирования экспериментов, метод оптимальных латинских гиперкубов. Подход эксплуатации и эксплорации, функция ожидаемого улучшения. Алгоритм эффективной глобальной оптимизации.