Технологии обработки информации

Лабораторная работа №8

Линейная регрессия

Цель работы

Синтезировать заданный алгоритм линейной регрессии. Выполнить проверку значимости полученной модели регрессии.

Форма контроля

Письменный отчёт (допускается преставление 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 алгоритм линейной регрессии. Получить коэффициенты модели регрессии. Выполнить проверку значимости полученной модели и представить результаты в виде выводов по проделанной работе.

Пример варианта задания

1. Используя стандартные функции Matlab (regress) построить модель линейной регрессии полиномиальной функции с коэффициентами a(1)=2; a(2)=-3; a(3)=17; a(4)=300; a(5)=250; a(6)=-1100. Вычислить коэффициент детерминации для следующих объемов обучающей выборки: 50, 100 и 1000. Дисперсия ошибки измерения выходной переменной = 1, уровень значимости для проверки гипотез по критерию Фишера = 0.01.

2. Используя стандартные функции Matlab (regress) построить модель линейной регрессии полиномиальной функции с коэффициентами a(1)=1; a(2)=2; a(3)=-10; a(4)=100; a(5)=-100; a(6)=1100. Вычислить коэффициент детерминации для следующих объемов обучающей выборки: 50, 100 и 1000. Дисперсия ошибки измерения выходной переменной = 2, уровень значимости для проверки гипотез по критерию Фишера = 0.01.

3. Используя стандартные функции Matlab (regress) построить модель линейной регрессии полиномиальной функции с коэффициентами a(1)=1; a(2)=-2; a(3)=-20; a(4)=-500; a(5)=550; a(6)=600. Вычислить коэффициент детерминации для следующих объемов обучающей выборки: 50, 100 и 1000. Дисперсия ошибки измерения выходной переменной = 3, уровень значимости для проверки гипотез по критерию Фишера = 0.02.

4. Используя стандартные функции Matlab (regress) построить модель линейной регрессии полиномиальной функции с коэффициентами a(1)=5; a(2)=-5; a(3)=5; a(4)=500; a(5)=500; a(6)=1000. Вычислить коэффициент детерминации для следующих объемов обучающей выборки: 50, 100 и 1000. Дисперсия ошибки измерения выходной переменной = 1, уровень значимости для проверки гипотез по критерию Фишера = 0.001.

5. Используя стандартные функции Matlab (regress) построить модель линейной регрессии полиномиальной функции с коэффициентами a(1)=-5; a(2)=5; a(3)=-30; a(4)=200; a(5)=-1000; a(6)=1000. Вычислить коэффициент детерминации для следующих объемов обучающей выборки: 50, 100 и 1000. Дисперсия ошибки измерения выходной переменной = 10, уровень значимости для проверки гипотез по критерию Фишера = 0.05.

6. Используя стандартные функции Matlab (regress) построить модель линейной регрессии гармонического ряда с коэффициентами a(1)=2; a(2)=-3; a(3)=17; a(4)=5; a(5)=2; a(6)=-1. Вычислить коэффициент детерминации для следующих объемов обучающей выборки: 50, 100 и 1000. Дисперсия ошибки измерения выходной переменной = 2, уровень значимости для проверки гипотез по критерию Фишера = 0.03.
7. Используя стандартные функции Matlab (regress) построить модель линейной регрессии гармонического ряда с коэффициентами a(1)=1; a(2)=-1; a(3)=2; a(4)=-2; a(5)=3; a(6)=-3. Вычислить коэффициент детерминации для следующих объемов обучающей выборки: 50, 100 и 1000. Дисперсия ошибки измерения выходной переменной = 10, уровень значимости для проверки гипотез по критерию Фишера = 0.001.
8. Используя стандартные функции Matlab (regress) построить модель линейной регрессии гармонического ряда с коэффициентами a(1)=1; a(2)=1; a(3)=2; a(4)=2; a(5)=3; a(6)=3. Вычислить коэффициент детерминации для следующих объемов обучающей выборки: 50, 100 и 1000. Дисперсия ошибки измерения выходной переменной = 5, уровень значимости для проверки гипотез по критерию Фишера = 0.05.
9. Используя стандартные функции Matlab (regress) построить модель линейной регрессии гармонического ряда с коэффициентами a(1)=4; a(2)=-5; a(3)=1; a(4)=5; a(5)=2; a(6)=1. Вычислить коэффициент детерминации для следующих объемов обучающей выборки: 50, 100 и 1000. Дисперсия ошибки измерения выходной переменной = 5, уровень значимости для проверки гипотез по критерию Фишера = 0.01.
10. Используя стандартные функции Matlab (regress) построить модель линейной регрессии гармонического ряда с коэффициентами a(1)=3; a(2)=6; a(3)=7; a(4)=-5; a(5)=4; a(6)=-1. Вычислить коэффициент детерминации для следующих объемов обучающей выборки: 50, 100 и 1000. Дисперсия ошибки измерения выходной переменной = 0.1, уровень значимости для проверки гипотез по критерию Фишера = 0.05.
11. Используя метод наименьших квадратов реализовать алгоритм гребневой регрессии для двумерных векторов входной переменной и параметра степени обусловленности матрицы ковариации = 1e-7. Вычислить значение параметра регуляризации, при котором коэффициент обусловленности матрицы **XtX** был равен 1000. Вычислить значения невязки (на обучающей и тестовой выборке) и СКО оценивания коэффициентов модели при следующих объемах выборок: 50, 200, 1000.
12. Используя метод наименьших квадратов реализовать алгоритм гребневой регрессии для двумерных векторов входной переменной и параметра степени обусловленности матрицы ковариации = 1e-8. Вычислить значение параметра регуляризации, при котором коэффициент обусловленности матрицы **XtX** был равен 1000. Вычислить значения невязки (на обучающей и тестовой выборке) и СКО оценивания коэффициентов модели при следующих объемах выборок: 50, 200, 1000.
13. Используя метод наименьших квадратов реализовать алгоритм гребневой регрессии для двумерных векторов входной переменной и параметра степени обусловленности матрицы ковариации = 1e-9. Вычислить значение параметра регуляризации, при котором коэффициент обусловленности матрицы **XtX** был равен 1000. Вычислить значения невязки (на обучающей и

тестовой выборке) и СКО оценивания коэффициентов модели при следующих объемах выборок: 50, 200, 1000.

14. Используя метод наименьших квадратов реализовать алгоритм гребневой регрессии для двумерных векторов входной переменной и параметра степени обусловленности матрицы ковариации = 1e-7. Вычислить значение параметра регуляризации, при котором коэффициент обусловленности матрицы **XtX** был равен 100. Вычислить значения невязки (на обучающей и тестовой выборке) и СКО оценивания коэффициентов модели при следующих объемах выборок: 50, 200, 1000.
15. Используя метод наименьших квадратов реализовать алгоритм гребневой регрессии для трехмерных векторов входной переменной и параметра степени обусловленности матрицы ковариации = 1e-8. Вычислить значение параметра регуляризации, при котором коэффициент обусловленности матрицы **XtX** был равен 100. Вычислить значения невязки (на обучающей и тестовой выборке) и СКО оценивания коэффициентов модели при следующих объемах выборок: 50, 200, 1000.
16. Используя метод наименьших квадратов реализовать алгоритм гребневой регрессии для трехмерных векторов входной переменной и параметра степени обусловленности матрицы ковариации = 1e-9. Вычислить значение параметра регуляризации, при котором коэффициент обусловленности матрицы **XtX** был равен 100. Вычислить значения невязки (на обучающей и тестовой выборке) и СКО оценивания коэффициентов модели при следующих объемах выборок: 50, 200, 1000.
17. Используя метод наименьших квадратов реализовать алгоритм гребневой регрессии для трехмерных векторов входной переменной и параметра степени обусловленности матрицы ковариации = 1e-6. Вычислить значение параметра регуляризации, при котором коэффициент обусловленности матрицы **XtX** был равен 2000. Вычислить значения невязки (на обучающей и тестовой выборке) и СКО оценивания коэффициентов модели при следующих объемах выборок: 50, 200, 1000.
18. Используя метод наименьших квадратов реализовать алгоритм гребневой регрессии для трехмерных векторов входной переменной и параметра степени обусловленности матрицы ковариации = 1e-7. Вычислить значение параметра регуляризации, при котором коэффициент обусловленности матрицы **XtX** был равен 2000. Вычислить значения невязки (на обучающей и тестовой выборке) и СКО оценивания коэффициентов модели при следующих объемах выборок: 50, 200, 1000.
19. Используя метод наименьших квадратов реализовать алгоритм гребневой регрессии для четырехмерных векторов входной переменной и параметра степени обусловленности матрицы ковариации = 1e-8. Вычислить значение параметра регуляризации, при котором коэффициент обусловленности матрицы **XtX** был равен 500. Вычислить значения невязки (на обучающей и тестовой выборке) и СКО оценивания коэффициентов модели при следующих объемах выборок: 50, 200, 1000.
20. Используя метод наименьших квадратов реализовать алгоритм гребневой регрессии для четырехмерных векторов входной переменной и параметра степени обусловленности матрицы ковариации = 1e-9. Вычислить значение параметра регуляризации, при котором коэффициент обусловленности матрицы **XtX** был равен 500. Вычислить значения невязки (на обучающей и

тестовой выборке) и СКО оценивания коэффициентов модели при следующих объемах выборок: 50, 200, 1000.

Примеры контрольных вопросов

1. Как объем обучающей выборки влияет на значимость модели регрессии?
2. Каким образом используемое значение параметра регуляризации влияет на качество получаемого решения?
3. В каких случаях требуется использование регуляризации?