Язык R в примерах и задачах

Юрченков Иван Александрович

14 09 2020

Table of Contents

# Введение в R

Значение переменных могут печататься как функцией **print()**, так и функцией **cat()**. Функция print отличается тем, что на вход принимает лишь одну структуру данных, и относится к перегруженным функциям относительно внутренних классов. Функция cat выводит переменное количество объектов “как есть”, то есть для данного вывода необходимо реально указывать пробелы и переносы строки.

# 1. Переменные и типы данных в R

## 1.1 Объявление переменных

Переменная - именованная область памяти. Такое определение переменной в языке R имеет своё определённое воплощение. Каждая объявленная в коде скрипта переменная сохраняется в своём окружении (environment) со своей областью видимости.

В языке R, переменная может быть объявлена несколькими способами, а именно:

# Объявление пермеменных --------------------------------------------  
x1 <- 5 # Правое присвоение  
x2 = 6 # Правое присвоение  
7 -> x3 # Левое присвоение

Стрелки влево и вправо являются позиционными, то есть имеет место позиционное присвоеие значений переменных в программе. Названия хранящихся в памяти проекта переменных, созданных ранее могут быть отображены с помощью функции **ls()**. Получить значение переменной в коде по строке, содержащей только имя переменной можно осуществить с помощью функции **get()**.

# Отображение имён переменных в памяти и их значений ---------------  
print(ls()) # print() выводит значения ls()

## [1] "x1" "x2" "x3"

print(get("x1")) # get() получает значение по строке имени

## [1] 5

Таким образом можно поэкспериментировать с объявлением переменных в коде проекта. Важно понимать, что разработчиками R поддерживается нативный вариант с правым присваиванием с помощью стрелки влево **" <- "**. Исходя из этого, не стоит принимать за серьёзное применение присваивания следующий пример:

# Эксперимент по присвоению целых значений переменным ---------------   
x1 <- x2 <- 8  
cat(x1, " ", x2, " ", x3, "\n")

## 8 8 7

x1 -> x3  
cat(x1, " ", x2, " ", x3, "\n")

## 8 8 8

x4 = 3 -> x5  
cat(x4, " ", x5)

## 3 3

Для имён переменных существуют некоторые правила, которые схожи с правилами именования переменных в других языках программирования.

## 1.2 Типы данных

Данные и переменные в языке R представлены следующими своими представители:

# 2. Векторы

# 3. Функции, ветвления и циклы

# 4. Матрицы

# 5. Фреймы данных

# 6. Статистика

# 7. Визуализация данных

# 8. Списки

# 9. Работа со строками в R

# 10. Факторные переменные

# 11. Больше о функциях в R

# 12. Работа с фреймами данных в R

# 13. Обработка данных

# 14. Линейная регрессия

# 15. Временные ряды в R

# 16. Уменьшение размерности данных. Метод главных компонент

# 17. Логистическая регрессия. Классификация

# 18. Байесовские методы

# 19. Факторный анализ в R

# 20. Кластеризация данных

# 21. Деревья решений

# 22. Случайный лес

# 23. Нейронные сети в R

# 24. Приложения в математических методах

# Темы курсовых работ

## Программирование на языке R

1. Реализация алгоритма многопараметрической линейной регрессии. Решение задач прогнозирования
2. Реализация алгоритма многоклассового классификационного метода на основе логистической регрессии
3. Реализация алгоритма Байесовской классификации
4. Реализация алгоритма кластеризации методом к-средних
5. Реализация EM-алгоритма кластеризации
6. Реализация алгоритмов уменьшения размерности данных. Метод главных компонент
7. Реализация алгоритмов уменьшения размерности данных. Метод независимых компонент
8. Реализация алгоритмов уменьшения размерности данных. Нейронная сеть автокодировщик
9. Факторный анализ. Реализация методов Фишера для линейного дискриминантного анализа
10. Реализация нейронной сети прямого распространения в задачах регрессионного анализа
11. Реализация нейронной сети прямого распространения в задачах прогнозирования временных рядов
12. Реализация методов авторегрессии
13. Реализация алгоритма дерева решений. Дерево решений в задачах классификации
14. Реализация алгоритма дерева решений. Дерево решений в задачах регрессии
15. Реализация алгоритма случайного леса. Случайный лес в задачах регрессионного анализа
16. Реализация алгоритмов решения операторных уравнений для матриц высокой размерности.

## Анализ данных в R

1. Прогнозирование временных рядов. Прогнозирование средних доходов населения России по отношению к доходам федерального бюджета. Заработная плата: <http://sophist.hse.ru/hse/1/tables/WAG_M.htm>, доходы федерального бюджета: <http://sophist.hse.ru/hse/1/tables/GOV_M.htm>
2. Прогнозирование временных рядов. Прогнозирование цен акций мировых компаний: <https://finance.yahoo.com>