

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
Факультет прикладной математики и информатики
Кафедра математического моделирования и анализа данных

Аннотация к дипломной работе

**Статистическая классификация многомерных наблюдений
в пространстве кривых Эндрюса**

Ротько Валерий Петрович

Научный руководитель: доктор физ.-мат. наук,
профессор Жук Евгений Евгеньевич

Минск, 2022

РЕФЕРАТ

Дипломная работа, 35 с., 9 рис., 6 источников, 6 приложений.

Ключевые слова: КРИВЫЕ ЭНДРЮСА, ВИЗУАЛИЗАЦИЯ, ЧИСЛЕННЫЕ ПРИЗНАКИ, ПРОЕКЦИЯ, ТРИГОНОМЕТРИЧЕСКАЯ СИСТЕМА.

Объекты исследования – многомерные данные с численными признаками.

Цель работы – исследовать возможности кривых Эндрюса и их модификаций по визуализации многомерных данных.

Методы исследования – анализ, изучение теоретических материалов.

Область применения – биология, социология, анализ данных, неврология, контроль качества продукции, выявление выбросов во временных рядах.

РЭФЕРАТ

Дыпломная праца, 35 с., 9 мал., 6 крыніцы, 6 прыкладанняў.

Ключавыя словы: КРЫВЫЯ ЭНДРЮСА, ВІЗУАЛІЗАЦЫЯ, КОЛЬКАСНЫЯ ПРЫКМЕТЫ, ПРАЕКЦЫЯ, ТРЫГАНАМЕТРЫЧНАЯ СІСТЭМА.

Аб'екты даследавання – шматмерныя данныя з колькаснымі прыкметамі.

Мэта працы – даследаваць магчымасці крывых Эндрюса і іх мадыфікацый па візуалізацыі шматмерных данных.

Метады даследавання – аналіз, вывучэнне тэарэтычных матэрыялаў.

Вобласць прымянення – біялогія, сацыялогія, аналіз данных, неўралогія, кантроль якасці прадукцыі, выяўленне выкідаў у часовых шэрагах.

ABSTRACT

Graduate work, 35 pages., 9 figures, 6 sources, 6 applications.

Keywords: ANDREWS PLOT, VISUALIZATION, NUMERIC FEATURES, PROJECTION, TRIGONOMETRIC SYSTEM.

Object of research – numerical multidimensional data.

Purpose of work – explore the possibilities of Andrews plots and their modifications to visualize multidimensional data.

Methods of research – analysis, study of theoretical materials.

Application area – biology, sociology, data analysis, neurology, product quality control, identification of outliers in time series.

ВВЕДЕНИЕ

Вне зависимости от конечной цели анализа тех или иных данных, возникает потребность в их качественной визуализации. Этот этап, конечно, не представляет сложности при визуализации данных с низкими размерностями, легко представляемыми человеческим мозгом, однако при увеличении их числа необходимо применять понижающие размерность методы визуализации. При этом хотелось бы, чтобы используемый метод сохранял важные взаимосвязи и математические свойства исходных данных.

Одним из способов представления многомерных данных являются кривые Эндрюса. При своей простоте, они позволяют сохранить ключевые свойства и взаимосвязи между объектами выборки, сопоставляя каждому наблюдению ровно одну кривую. При этом важно отметить существование множества вариаций кривых Эндрюса, как двумерных, так и трехмерных, предоставляющих различные способы восприятия и трактовки полученных результатов.

В данной работе рассмотрены свойства и визуализации как самого базового варианта кривых Эндрюса, так и их более сложных вариантов, а именно, кривых Эндрюса с масштабированием данных, с пропусками признаков, двумерных параметрических и трехмерных параметрических кривых Эндрюса.