

Динамическое выравнивание многомерных временных рядов*

Гончаров А. В., Моргачев Г. И., Смирнов В., Липницкая Т.

morgachev.gi@phystech.edu, smirnov.vs@phystech.edu, tanya.lipnizky@yandex.ru

МФТИ

В работе рассматривается задача получения расстояния между многомерными временными рядами. На текущий момент для этого используется алгоритм DTW и его эффективные модификации. В классическом понимании DTW реализует способ получения расстояния между двумя одномерными рядами. Исследуется возможность его обобщения на многомерный случай. Рассматриваются существующий подходы к решению данной задачи и исследуется зависимость качества выравнивания от использования различных функций расстояния. В качестве прикладной задачи рассматривается кластеризация данных мозговой активности обезьян. Качество кластеризации на основе полученной функции расстояния сравнивается с алгоритмом на основе авторегрессионной модели.

Ключевые слова: *Временные ряды, многомерные временные ряды, DTW.*

1 Введение

В ряде задач анализа данных исследователям приходится иметь дело с измерениями, зависящими от времени. Последовательности такие измерений часто называют временными рядами. Для применения многих стандартных методов к подобным данным на них необходимо задать функцию расстояния. Стандартный подход, базирующийся на сумме поточечным расстояний, плохо отражают схожесть временных рядов [1]. Одним из способов решения этой проблемы является выравнивание рядов друг относительно друга. [2] и его модификаций [3]. В работе рассматривается обобщение DTW на случай многомерных рядов.

В работах [4], [5] предлагаются подходы к решению этой задачи. В работе [4] предлагается способ выравнивания многомерных рядов, основанный на нормализации исходных данных и нахождении векторной нормы. В работе [5] рассматривается алгоритм, позволяющий выполнить выравнивание временных рядов между координатами.

Предлагается рассмотреть влияние различных функций расстояния на точность выравнивания: расстояние Минковского при $p = 1, 2$, косинусная мера.

Зачастую, перед работой с временными рядами полезно провести их предварительную кластеризацию. Качество полученной функции расстояния будет оцениваться как качество кластеризации временных рядов метрическими методами.

В качестве данных для задачи выступают данные мозговой активности обезьян и ее зависимости от положения глаз. Данные представляют собой электрокортикограмму - зависимости потенциалов в 128 точках мозга и положения глаз от времени. Рассматривается кластеризация данных мозговой активности, что соответствует кластеризация 128-мерных временных рядов.

Исследуется возможность обобщения DTW для выравнивания многомерных временных рядов относительно общей временной шкалы. Также рассматривается возможность выравнивания значений координат друг относительно друга.

* Работа выполнена при финансовой поддержке РФФИ, проект № 00-00-00000. Научный руководитель: Гончаров А. В. Задачу поставил: Гончаров А. В. Консультант: Гончаров А. В.

Расстояния между выравненными рядами получается как сумма поточечных расстояний и используется для кластеризация методом k ближайших соседей. Для оценки качества полученных расстояний предлагается сравнивать качества кластеризация с кластеризация на основании коэффициентов авторегрессионной модели для различных функций расстояния между выравненными рядами.

2 Название раздела

2.1 Название параграфа.

2.2 Теоретическую часть работы

3 Заключение

Литература

- [1] Hui Ding, Goce Trajcevski, Peter Scheuermann, Xiaoyue Wang, and Eamonn Keogh. *Querying and mining of time series data: Experimental comparison of representations and distance measures*, volume 1, pages 1542–1552. 2 edition, 8 2008.
- [2] Eamonn J. Keogh and Michael J. Pazzani. Derivative dynamic time warping. In *In SIAM International Conference on Data Mining*, 2001.
- [3] Stan Salvador and Philip Chan. Toward accurate dynamic time warping in linear time and space. *Intell. Data Anal*, 11(5):561–580, 2007.
- [4] G.A. ten Holt, M.J.T. Reinders, and E.A. Hendriks. Multi-dimensional dynamic time warping for gesture recognition. In *Thirteenth annual conference of the Advanced School for Computing and Imaging*, 2007.
- [5] Parinya Sanguansat. Multiple multidimensional sequence alignment using generalized dynamic time warping. 2012.