Springer Series in the Data Sciences

Nickolay Trendafilov
Michele Gallo

Multivariate Data Analysis on Matrix Manifolds

(with Manopt)

Springer

Springer Series in the Data Sciences

Springer Series in the Data Sciences focuses primarily on monographs and graduate level textbooks. The target audience includes students and researchers working in and across the fields of mathematics, theoretical computer science, and statistics. Data Analysis and Interpretation is a broad field encompassing some of the fastest-growing subjects in interdisciplinary statistics, mathematics and computer science. It encompasses a process of inspecting, cleaning, transforming, and modeling data with the goal of discovering useful information, suggesting conclusions, and supporting decision making. Data analysis has multiple facets and approaches, including diverse techniques under a variety of names, in different business, science, and social science domains. Springer Series in the Data Sciences addresses the needs of a broad spectrum of scientists and students who are utilizing quantitative methods in their daily research. The series is broad but structured, including topics within all core areas of the data sciences. The breadth of the series reflects the variation of scholarly projects currently underway in the field of machine learning.

More information about this series at http://www.springer.com/series/13852

Nickolay Trendafilov · Michele Gallo

Multivariate Data Analysis on Matrix Manifolds

(with Manopt)

Springer

Nickolay Trendafilov
School of Mathematics and Statistics
Open University
Milton Keynes, Buckinghamshire, UK

Michele Gallo
Department of Human and Social Sciences
University of Naples "L'Orientale"
Naples, Italy

ISSN 2365-5674 ISSN 2365-5682 (electronic)
Springer Series in the Data Sciences
ISBN 978-3-030-76973-4 ISBN 978-3-030-76974-1 (eBook)
https://doi.org/10.1007/978-3-030-76974-1

Mathematics Subject Classification: 58A05, 58C05, 62–07, 62H12, 62H25, 62H30, 62H99, 65C60, 65Fxx, 65K99, 90C26, 90C51

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To the memory of

my mother, Zdravka, and my father, Trendafil

and to my family

wife, Irina, son, Iassen,
and grandchildren, Veronica and Christian

Preface

We want to start with few remarks predating considerably the emerging of the idea for writing this book and our collaboration in general. They are related to the first author's own experience which made him explore matrix manifolds in data analysis problems.

NTT was in an early stage of his research career, after his Ph.D. was completed on a completely different topic. He was assigned to study factor analysis (FA) and do some programming for a particular software product. While working on FA, NTT realized that the most interesting part for him is the FA interpretation and the so-called rotation methods (see Section 4.5). NTT recognized that the main goal of FA is to produce simple for interpretation results, which is achieved by orthogonal rotation of some initial, usually difficult to interpret, FA solution Λ (known as factor loadings matrix). However, how we can define what is "simple for interpretation results"? The problem is really tough, especially if you try to capture its meaning in a single mathematical expression/formula. Because of that, a huge number of different formulas were proposed each of them claiming to approximate in some sense the idea for "simple for interpretation results". In FA, these formulas are called simple structure criteria. They are supposed to measure/quantify the simplicity of a certain FA solution. Let f be such a criterion. Then, if Λ_2 is simpler than Λ_1, and if Λ_3 is even simpler than Λ_2, then we should have $f(\Lambda_1) > f(\Lambda_2) > f(\Lambda_3)$, or $f(\Lambda_1) < f(\Lambda_2) < f(\Lambda_3)$ depending on the sign of f. Clearly, one would be interested to find the solution Λ, for which f reaches its smallest/largest value, and the problem smells of optimization.

For reasons to become clear in Section 4.5, the rotation methods transform the initial (difficult to interpret) solution Λ_0 into another, simpler, one $\Lambda_0 Q$, where Q is an orthogonal matrix, which gives the name "rotation methods". Thus, in formal terms, the rotation methods look for an orthogonal matrix Q, such that $f(\Lambda_0) > f(\Lambda_0 Q)$. The problem stays the same no matter what rotation criterion f is considered. Thus, one needs to solve

$$\min_{Q} f(\Lambda_0 Q), \tag{1}$$

where Q is orthogonal and f is an arbitrary criterion.

NTT was surprised to see in the psychometric literature, e.g. Harman (1976); Mulaik (1972), that the introduction of any new simplicity criterion, f, was inevitably connected to a new method for solving (1). At that moment his knowledge in optimization was quite limited, but intuitively NTT felt that this cannot be right: a new optimization method for every new function f. It sounded ridiculous!

And then, NTT started to look for a general approach that can work for every f, as far as the variable Q is an orthogonal matrix. In this way, NTT realized that the rotational problem can be naturally defined as optimization on the matrix manifold of all orthogonal matrices. The understanding that, in fact, all MDA problems can be seen defined on specific matrix manifold(s) followed then straight away.

Milton Keynes, UK — Nickolay Trendafilov
Naples, Italy — Michele Gallo

Contents