Differential geometric properties of textile plot

University of Tokyo Tomonari Sei
Osaka Prefecture University Ushio Tanaka

Abstract

The textile plot is a method for data visualization proposed by Ku-
masaka and Shibata (2008), which transforms a data matrix to draw a
parallel coordinate plot. From a differential geometric point of view, we
investigate a set of matrices induced by the textile plot, which we call the
textile set.

Preliminaries and results

We define the textile set by a set of matrices induced by the textile plot. We will
use bold uppercase letters for matrices and bold lowercase letters for column
vectors.

Let \(X = (x_1, \ldots, x_p) \in \mathbb{R}^{n \times p}\) be a data matrix of \(n\) individuals and \(p\)
variates, where each \(x_i\) is a column vector. For example, imagine a data of
characteristics of \(n\) students in a school. Then the first column of \(X\) may
represent age, the second is height, the third is weight, and so forth. Note that,
in each column, every element has the common unit ([years], [cm], [kg], etc. ).
For simplicity, we assume that the data matrix has no missing value and that
each variate is numeric. Each column of \(X\) is assumed to be non-degenerate.
In the following, without loss of generality, we assume that the data is scaled:
\(x'_j 1_n = 0\) and \(\|x_j\| = 1\), where \(x'\) is the transpose of \(x\), \(\|x\| = (x'x)^{1/2}\)
is the Euclidean norm, and \(1_n = (1, \ldots, 1)' \in \mathbb{R}^n\).

The textile plot generates another matrix \(Y \in \mathbb{R}^{n \times p}\) from \(X\) by location
and scale transformations as follows: The matrix \(Y = (y_1, \ldots, y_p)\) is defined by
\(y_j = b_j x_j, j = 1, \ldots, p\), where \((b_1, \ldots, b_p)'\) is the unit eigenvector corresponding
to the maximum eigenvalue of the sample correlation matrix \((x'_i x_j)_{i,j=1}^p\).

**Definition** (Textile set). Let \(n\) and \(p\) be positive integers. The textile set \(T_{n,p}\)
is defined as a matrix \(Y = (y_1, \ldots, y_p) \subset \mathbb{R}^{n \times p}\) satisfying
\[
\exists \lambda \in \mathbb{R} \quad \forall i \in \{1, \ldots, p\} \quad \sum_{j=1}^p y_i'y_j = \lambda \|y_i\|^2, \sum_{j=1}^p \|y_j\|^2 = 1.
\]

Here let us state our results with small \(p\) only. In fact, it is possible to gen-
erate the case of low-dimensional \(p\) to that of high-dimension. Our observation
is the following:

**Theorem.** \(T_{n,1} = S^{n-1}\), namely \(T_{n,1}\) is \((n-1)\)-dimensional unit sphere of
\(\mathbb{R}^n\), and \(T_{n,2}\) is the cup of certain manifolds, each of which is diffeomorphic to
\(S^{n-1} \times S^{n-1}\).