Geometric Properties of textile plot

Tomonari SEI and Ushio TANAKA

University of Tokyo and Osaka Prefecture University

at École Polytechnique, Oct 28, 2015

Introduction

- The textile plot proposed by Kumasaka and Shibata (2008) is a method for data visualization.
- The method transforms a data matrix into another matrix,

$$\mathbb{R}^{n\times p} \ni \mathbf{X} \;\mapsto\; \mathbf{Y} \in \mathbb{R}^{n\times p},$$

  in order to draw a parallel coordinate plot.
- The parallel coordinate plot is a standard 2-dimensional graphical tool for visualizing multivariate data at a glance.
- In this talk, we investigate a set of matrices induced by the textile plot, which we call the textile set, from a differential geometrical point of view.
- It is shown that the textile set is written as the union of two differentiable manifolds if data matrices are "generic".

Introduction

- The textile plot proposed by Kumasaka and Shibata (2008) is a method for data visualization.
- The method transforms a data matrix into another matrix,

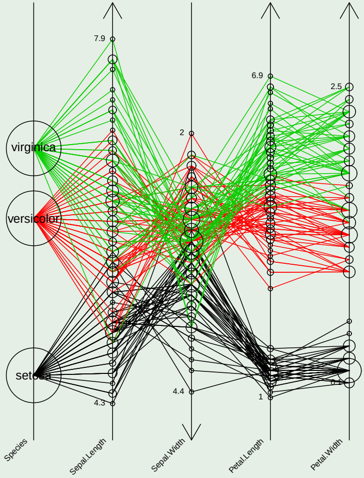
$$\mathbb{R}^{n\times p} \ni \mathbf{X} \;\mapsto\; \mathbf{Y} \in \mathbb{R}^{n\times p},$$

  in order to draw a parallel coordinate plot.
- The parallel coordinate plot is a standard 2-dimensional graphical tool for visualizing multivariate data at a glance.
- In this talk, we investigate a set of matrices induced by the textile plot, which we call the textile set, from a differential geometrical point of view.
- It is shown that the textile set is written as the union of two differentiable manifolds if data matrices are "generic".

Textile plot

Example (Kumasaka and Shibata, 2008)

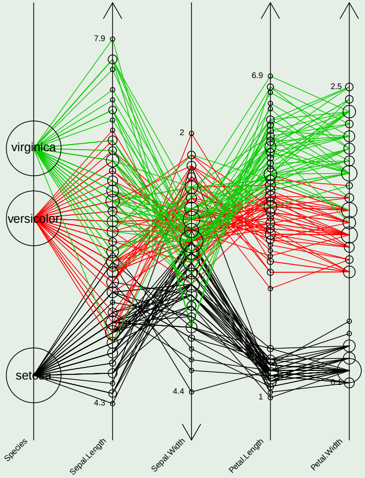
- Textile plot for the iris data. (150 cases, 5 attributes)
- Each variate is transformed by a location-scale transformation.
- Categorical data is quantified.
- Missing data is admitted.
- Order of axes can be maintained.



Textile plot

Example (Kumasaka and Shibata, 2008)

- Textile plot for the iris data. (150 cases, 5 attributes)
- Each variate is transformed by a location-scale transformation.
- Categorical data is quantified.
- Missing data is admitted.
- Order of axes can be maintained.



Textile plot

Let us recall the method of the textile plot.

- For simplicity, we assume no categorical variate and no missing value.
- Let $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_p) \in \mathbb{R}^{n \times p}$ be the data matrix.
- Without loss of generality, assume the sample mean and sample variance of each $\mathbf{x}_j$ are 0 and 1, respectively.
- The data is transformed into $\mathbf{Y} = (\mathbf{y}_1, \ldots, \mathbf{y}_p)$, where

$$\mathbf{y}_j = a_j + b_j \mathbf{x}_j, \quad a_j, b_j \in \mathbb{R}, \ j = 1, \ldots, p.$$

- The coefficients a_j and b_j are determined by the following procedure.

Textile plot

Let us recall the method of the textile plot.

- For simplicity, we assume no categorical variate and no missing value.
- Let $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_p) \in \mathbb{R}^{n \times p}$ be the data matrix.
- Without loss of generality, assume the sample mean and sample variance of each $\mathbf{x}_j$ are 0 and 1, respectively.
- The data is transformed into $\mathbf{Y} = (\mathbf{y}_1, \ldots, \mathbf{y}_p)$, where

$$\mathbf{y}_j = a_j + b_j \mathbf{x}_j, \quad a_j, b_j \in \mathbb{R}, \ j = 1, \ldots, p.$$

- The coefficients a_j and b_j are determined by the following procedure.

Textile plot

Let us recall the method of the textile plot.

- For simplicity, we assume no categorical variate and no missing value.
- Let $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_p) \in \mathbb{R}^{n \times p}$ be the data matrix.
- Without loss of generality, assume the sample mean and sample variance of each $\mathbf{x}_j$ are 0 and 1, respectively.
- The data is transformed into $\mathbf{Y} = (\mathbf{y}_1, \ldots, \mathbf{y}_p)$, where

$$\mathbf{y}_j = a_j + b_j \mathbf{x}_j, \quad a_j, b_j \in \mathbb{R}, \; j = 1, \ldots, p.$$

- The coefficients a_j and b_j are determined by the following procedure.

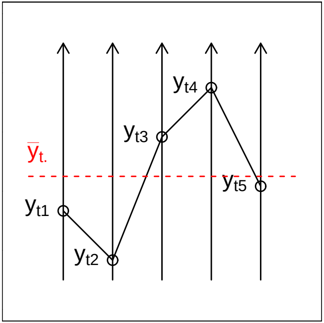
Textile plot

- Coefficients $\mathbf{a} = (a_j)$ and $\mathbf{b} = (b_j)$ are the solution of the following minimization problem:

$$\underset{\mathbf{a},\mathbf{b}}{\text{Minimize}} \quad \sum_{t=1}^{n} \sum_{j=1}^{p} (y_{tj} - \bar{y}_{t\cdot})^2$$

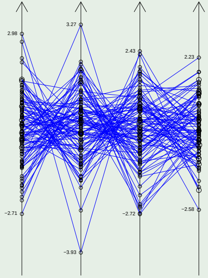
$$\text{subject to} \quad \mathbf{y}_j = a_j + b_j \mathbf{x}_j, \quad \sum_{j=1}^{p} \|\mathbf{y}_j\|^2 = 1.$$

- Intuition: as horizontal as possible.
- Solution: $\mathbf{a} = \mathbf{0}$ and $\mathbf{b}$ is the eigenvector corresponding to the maximum eigenvalue of the covariance matrix of $\mathbf{X}$.

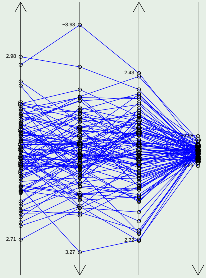


Example ($n = 100,\ p = 4$)

$\mathbf{X} \in \mathbb{R}^{100\times 4}$. Each row $\sim N(\mathbf{0}, \mathbf{\Sigma})$, $\mathbf{\Sigma} = \begin{pmatrix} 1 & -0.6 & 0.5 & 0.1 \\ -0.6 & 1 & -0.6 & -0.2 \\ 0.5 & -0.6 & 1 & 0.0 \\ 0.1 & -0.2 & 0.0 & 1 \end{pmatrix}$.



(a) raw data **X**



(b) textile plot **Y**

Our motivation

- The textile plot transforms the data matrix $\mathbf{X}$ into $\mathbf{Y}$.
- Denote the map by $\mathbf{Y} = \tau(\mathbf{X})$.
- What is the image $\tau(\mathbb{R}^{n \times p})$?
- We can show that $\mathbf{Y} \in \tau(\mathbb{R}^{n \times p})$ satisfies two conditions:

$$\exists \lambda \geq 0, \quad \forall i = 1, \ldots, p, \quad \sum_{j=1}^{p} \mathbf{y}_i' \mathbf{y}_j = \lambda \|\mathbf{y}_i\|^2$$

and

$$\sum_{j=1}^{p} \|\mathbf{y}_j\|^2 = 1.$$

This motivates the following definition of the textile set.

Our motivation

- The textile plot transforms the data matrix $\mathbf{X}$ into $\mathbf{Y}$.
- Denote the map by $\mathbf{Y} = \tau(\mathbf{X})$.
- What is the image $\tau(\mathbb{R}^{n \times p})$?
- We can show that $\mathbf{Y} \in \tau(\mathbb{R}^{n \times p})$ satisfies two conditions:

$$\exists \lambda \geq 0, \quad \forall i = 1, \ldots, p, \quad \sum_{j=1}^{p} \mathbf{y}_i' \mathbf{y}_j = \lambda \|\mathbf{y}_i\|^2$$

and

$$\sum_{j=1}^{p} \|\mathbf{y}_j\|^2 = 1.$$

This motivates the following definition of the textile set.

Our motivation

- The textile plot transforms the data matrix $\mathbf{X}$ into $\mathbf{Y}$.
- Denote the map by $\mathbf{Y} = \tau(\mathbf{X})$.
- What is the image $\tau(\mathbb{R}^{n\times p})$?
- We can show that $\mathbf{Y} \in \tau(\mathbb{R}^{n\times p})$ satisfies two conditions:

$$\exists \lambda \geq 0, \quad \forall i = 1, \ldots, p, \quad \sum_{j=1}^{p} \mathbf{y}_i' \mathbf{y}_j = \lambda \|\mathbf{y}_i\|^2$$

and

$$\sum_{j=1}^{p} \|\mathbf{y}_j\|^2 = 1.$$

This motivates the following definition of the textile set.

Textile set

Definition

The textile set is defined by

$$T_{n,p} = \{ \mathbf{Y} \in \mathbb{R}^{n\times p} \mid \exists\lambda \geq 0,\ \forall i,\ \sum_j \mathbf{y}_i'\mathbf{y}_j = \lambda\|\mathbf{y}_i\|^2,\ \sum_j \|\mathbf{y}_j\|^2 = 1 \},$$

The unnormalized textile set is defined by

$$U_{n,p} = \{ \mathbf{Y} \in \mathbb{R}^{n\times p} \mid \exists\lambda \geq 0,\ \forall i,\ \sum_j \mathbf{y}_i'\mathbf{y}_j = \lambda\|\mathbf{y}_i\|^2 \}.$$

- We are interested in mathematical properties of $T_{n,p}$ and $U_{n,p}$.
- Bad news: statistical implication such is a future work.
- Let us begin with small p case.

Textile set

Definition

The textile set is defined by

$$T_{n,p} = \{\, \mathbf{Y} \in \mathbb{R}^{n\times p} \mid \exists\lambda \geq 0,\ \forall i,\ \sum_j \mathbf{y}_i'\mathbf{y}_j = \lambda\|\mathbf{y}_i\|^2,\ \sum_j \|\mathbf{y}_j\|^2 = 1 \,\},$$

The unnormalized textile set is defined by

$$U_{n,p} = \{\, \mathbf{Y} \in \mathbb{R}^{n\times p} \mid \exists\lambda \geq 0,\ \forall i,\ \sum_j \mathbf{y}_i'\mathbf{y}_j = \lambda\|\mathbf{y}_i\|^2 \}.$$

- We are interested in mathematical properties of $T_{n,p}$ and $U_{n,p}$.
- Bad news: statistical implication such is a future work.
- Let us begin with small p case.

Textile set

Definition

The textile set is defined by

$$T_{n,p} = \{\, \mathbf{Y} \in \mathbb{R}^{n \times p} \mid \exists \lambda \geq 0,\ \forall i,\ \sum_j \mathbf{y}_i' \mathbf{y}_j = \lambda \|\mathbf{y}_i\|^2,\ \sum_j \|\mathbf{y}_j\|^2 = 1 \,\},$$

The unnormalized textile set is defined by

$$U_{n,p} = \{\, \mathbf{Y} \in \mathbb{R}^{n \times p} \mid \exists \lambda \geq 0,\ \forall i,\ \sum_j \mathbf{y}_i' \mathbf{y}_j = \lambda \|\mathbf{y}_i\|^2 \}.$$

- We are interested in mathematical properties of $T_{n,p}$ and $U_{n,p}$.
- Bad news: statistical implication such is a future work.
- Let us begin with small p case.

Textile set

Definition

The textile set is defined by

$$T_{n,p} = \{ \mathbf{Y} \in \mathbb{R}^{n \times p} \mid \exists \lambda \geq 0,\ \forall i,\ \sum_j \mathbf{y}_i' \mathbf{y}_j = \lambda \|\mathbf{y}_i\|^2,\ \sum_j \|\mathbf{y}_j\|^2 = 1 \},$$

The unnormalized textile set is defined by

$$U_{n,p} = \{ \mathbf{Y} \in \mathbb{R}^{n \times p} \mid \exists \lambda \geq 0,\ \forall i,\ \sum_j \mathbf{y}_i' \mathbf{y}_j = \lambda \|\mathbf{y}_i\|^2 \}.$$

- We are interested in mathematical properties of $T_{n,p}$ and $U_{n,p}$.
- Bad news: statistical implication such is a future work.
- Let us begin with small p case.

Textile set

Definition

The textile set is defined by

$$T_{n,p} = \{\, \mathbf{Y} \in \mathbb{R}^{n \times p} \mid \exists \lambda \geq 0,\ \forall i,\ \sum_j \mathbf{y}_i' \mathbf{y}_j = \lambda \|\mathbf{y}_i\|^2,\ \sum_j \|\mathbf{y}_j\|^2 = 1 \,\},$$

The unnormalized textile set is defined by

$$U_{n,p} = \{\, \mathbf{Y} \in \mathbb{R}^{n \times p} \mid \exists \lambda \geq 0,\ \forall i,\ \sum_j \mathbf{y}_i' \mathbf{y}_j = \lambda \|\mathbf{y}_i\|^2 \}.$$

- We are interested in mathematical properties of $T_{n,p}$ and $U_{n,p}$.
- Bad news: statistical implication such is a future work.
- Let us begin with small p case.

$T_{n,p}$ with small p

Lemma ($p = 1$)

$T_{n,1} = \mathbb{S}^{n-1}$, *the unit sphere.*

Lemma ($p = 2$)

$T_{n,2} = A \cup B$, *where*

$$A = \{(\mathbf{y}_1, \mathbf{y}_2) \mid \|\mathbf{y}_1\| = \|\mathbf{y}_2\| = 1/\sqrt{2}\},$$
$$B = \{(\mathbf{y}_1, \mathbf{y}_2) \mid \|\mathbf{y}_1 - \mathbf{y}_2\| = \|\mathbf{y}_1 + \mathbf{y}_2\| = 1\},$$

each of which is diffeomorphic to $\mathbb{S}^{n-1} \times \mathbb{S}^{n-1}$. *Their intersection* $A \cap B$ *is diffeomorphic to the Stiefel manifold* $V_{n,2}$.

→ See next slide for $n = p = 2$ case.

$T_{n,p}$ with small p

Lemma ($p = 1$)

$T_{n,1} = \mathbb{S}^{n-1}$, the unit sphere.

Lemma ($p = 2$)

$T_{n,2} = A \cup B$, where

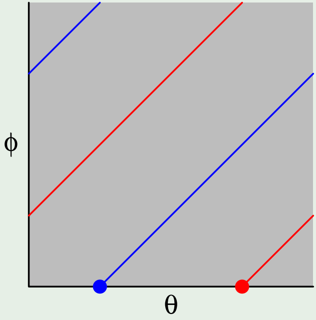
$$A = \{(\mathbf{y}_1, \mathbf{y}_2) \mid \|\mathbf{y}_1\| = \|\mathbf{y}_2\| = 1/\sqrt{2}\},$$
$$B = \{(\mathbf{y}_1, \mathbf{y}_2) \mid \|\mathbf{y}_1 - \mathbf{y}_2\| = \|\mathbf{y}_1 + \mathbf{y}_2\| = 1\},$$

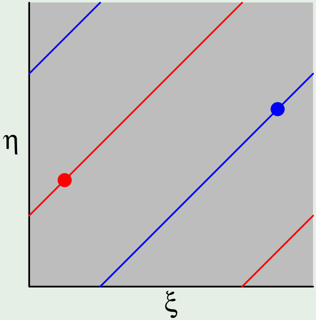
each of which is diffeomorphic to $\mathbb{S}^{n-1} \times \mathbb{S}^{n-1}$. Their intersection $A \cap B$ is diffeomorphic to the Stiefel manifold $V_{n,2}$.

$\rightarrow$ See next slide for $n = p = 2$ case.

Example ($n = p = 2$)

$T_{2,2} \subset \mathbb{R}^4$ is the union of two tori, glued along $O(2)$.





$$T_{2,2} = \left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} \cos\theta & \cos\phi \\ \sin\theta & \sin\phi \end{pmatrix} \right\} \cup \left\{ \frac{1}{2} \begin{pmatrix} \cos\xi + \cos\eta & \cos\xi - \cos\eta \\ \sin\xi + \sin\eta & \sin\xi - \sin\eta \end{pmatrix} \right\}$$

For general dimension p

To state our main result, we define two concepts: noncompact Stiefel manifold and canonical form.

Definition (e.g. Absil et al. (2008))

Let $n \geq p$. Denote by V^* the set of all column full-rank matrices:

$$V^* := \{ \mathbf{Y} \in \mathbb{R}^{n \times p} \mid \mathrm{rank}(\mathbf{Y}) = p \}.$$

V^* is called the noncompact Stiefel manifold.

- Note that $\dim(V^*) = np$ and $\overline{V^*} = \mathbb{R}^{n \times p}$.
- The orthogonal group $O(n)$ acts on V^*.
- By the Gram-Schmidt orthonormalization, the quotient space $V^*/O(n)$ is identified with upper-triangular matrices with positive diagonals. → see next slide.

For general dimension p

To state our main result, we define two concepts: noncompact Stiefel manifold and canonical form.

Definition (e.g. Absil et al. (2008))

Let $n \geq p$. Denote by V^* the set of all column full-rank matrices:

$$V^* := \{\, \mathbf{Y} \in \mathbb{R}^{n \times p} \mid \mathrm{rank}(\mathbf{Y}) = p \,\}.$$

V^* is called the noncompact Stiefel manifold.

- Note that $\dim(V^*) = np$ and $\overline{V^*} = \mathbb{R}^{n \times p}$.
- The orthogonal group $O(n)$ acts on V^*.
- By the Gram-Schmidt orthonormalization, the quotient space $V^*/O(n)$ is identified with upper-triangular matrices with positive diagonals. → see next slide.

For general dimension p

To state our main result, we define two concepts: noncompact Stiefel manifold and canonical form.

Definition (e.g. Absil et al. (2008))

Let $n \geq p$. Denote by V^* the set of all column full-rank matrices:

$$V^* := \{ \mathbf{Y} \in \mathbb{R}^{n \times p} \mid \mathrm{rank}(\mathbf{Y}) = p \}.$$

V^* is called the noncompact Stiefel manifold.

- Note that $\dim(V^*) = np$ and $\overline{V^*} = \mathbb{R}^{n \times p}$.
- The orthogonal group $O(n)$ acts on V^*.
- By the Gram-Schmidt orthonormalization, the quotient space $V^*/O(n)$ is identified with upper-triangular matrices with positive diagonals. → see next slide.

Noncompact Stiefel manifold and canonical form

Definition (Canonical form)

Let us denote by V^{**} the set of all matrices written as

$$\begin{pmatrix} y_{11} & \cdots & y_{1p} \\ 0 & \ddots & \vdots \\ \vdots & \ddots & y_{pp} \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}, \quad y_{ii} > 0,\ 1 \le i \le p.$$

We call it a canonical form.

Note that $V^{**} \subset V^{*}$ and $V^{*}/O(n) \simeq V^{**}$.

Noncompact Stiefel manifold and canonical form

Definition (Canonical form)

Let us denote by V^{**} the set of all matrices written as

$$\begin{pmatrix} y_{11} & \cdots & y_{1p} \\ 0 & \ddots & \vdots \\ \vdots & \ddots & y_{pp} \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}, \quad y_{ii} > 0,\ 1 \le i \le p.$$

We call it a canonical form.

Note that $V^{**} \subset V^{*}$ and $V^{*}/O(n) \simeq V^{**}$.

Restriction of unnormalized textile set

V^*: non-compact Stiefel manifold, V^{**}: set of canonical forms.

Definition

Denote the restriction of $U_{n,p}$ to V^* and V^{**} by

$$U^*_{n,p} = U_{n,p} \cap V^*, \qquad U^{**}_{n,p} = U_{n,p} \cap V^{**},$$

respectively.

- The group $O(n)$ acts on $U^*_{n,p}$.
- The quotient space $U^*_{n,p}/O(n)$ is identified with $U^{**}_{n,p}$.
- So it is essential to study $U^{**}_{n,p}$.

Restriction of unnormalized textile set

V^*: non-compact Stiefel manifold, V^{**}: set of canonical forms.

Definition

Denote the restriction of $U_{n,p}$ to V^* and V^{**} by

$$U^*_{n,p} = U_{n,p} \cap V^*, \qquad U^{**}_{n,p} = U_{n,p} \cap V^{**},$$

respectively.

- The group $O(n)$ acts on $U^*_{n,p}$.
- The quotient space $U^*_{n,p}/O(n)$ is identified with $U^{**}_{n,p}$.
- So it is essential to study $U^{**}_{n,p}$.

Restriction of unnormalized textile set

V^*: non-compact Stiefel manifold, V^{**}: set of canonical forms.

Definition

Denote the restriction of $U_{n,p}$ to V^* and V^{**} by

$$U^*_{n,p} = U_{n,p} \cap V^*, \qquad U^{**}_{n,p} = U_{n,p} \cap V^{**},$$

respectively.

- The group $O(n)$ acts on $U^*_{n,p}$.
- The quotient space $U^*_{n,p}/O(n)$ is identified with $U^{**}_{n,p}$.
- So it is essential to study $U^{**}_{n,p}$.

$U^{**}_{n,p}$ for small p

Let us check examples.

Example ($n = p = 1$)

$U^{**}_{1,1} = \{(1)\}$.

Example ($n = p = 2$)

Let $\mathbf{Y} = \begin{pmatrix} y_{11} & y_{12} \\ 0 & y_{22} \end{pmatrix}$ with $y_{11}, y_{22} > 0$. Then

$$U^{**}_{2,2} = \{y_{12} = 0\} \cup \{y_{11}^2 = y_{12}^2 + y_{22}^2\},$$

union of a plane and a cone.

$U^{**}_{n,p}$ for small p

Let us check examples.

Example ($n = p = 1$)

$U^{**}_{1,1} = \{(1)\}$.

Example ($n = p = 2$)

Let $\mathbf{Y} = \begin{pmatrix} y_{11} & y_{12} \\ 0 & y_{22} \end{pmatrix}$ with $y_{11}, y_{22} > 0$. Then

$$U^{**}_{2,2} \;=\; \{y_{12} = 0\} \;\cup\; \{y_{11}^2 = y_{12}^2 + y_{22}^2\},$$

union of a plane and a cone.

Main theorem

The differential geometrical property of $U^{**}_{n,p}$ is given as follows:

Theorem

Let $n \geq p \geq 3$. Then we have the following decomposition

$$U^{**}_{n,p} = M_1 \cup M_2,$$

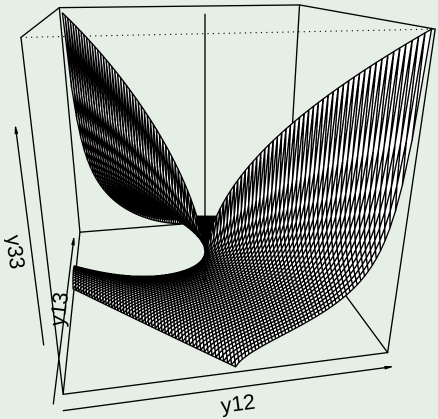
where each M_i is a differentiable manifold, the dimensions of which are given by

$$\dim M_1 = \frac{p(p+1)}{2} - (p-1),$$
$$\dim M_2 = \frac{p(p+1)}{2} - p,$$

respectively. M_2 is connected while M_1 may not.

Example

- $U_{3,3}^{**}$ is the union of 4-dim and 3-dim manifolds.
- We look at a cross section with $y_{11} = y_{22} = 1$:



Union of a surface and a vertical line.

Corollary

Let $n \geq p \geq 3$. Then we have

$$U^*_{n,p} = \pi^{-1}(M_1) \cup \pi^{-1}(M_2),$$

where π denotes the map of Gram-Schmidt orthonormalization. The dimensions are

$$\dim \pi^{-1}(M_1) = np - (p-1),$$
$$\dim \pi^{-1}(M_2) = np - p.$$

Other results

We state other results. First we have $n = 1$ case.

Lemma

If $n = 1$, then the textile set $T_{1,p}$ is the union of a $(p-2)$-dimensional manifold and $2(2^p - 1)$ isolated points.

Example

$U^{**}_{1,3}$ consists of a circle and 14 points:

$$U^{**}_{1,3} = (S^2 \cap \{y_1 + y_2 + y_3 = 1\}) \cup \{\pm(\tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}), \pm(\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}, 0), \pm(\tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}}), \pm(0, \tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}), \pm(1,0,0), \pm(0,1,0), \pm(0,0,1)\}$$

.

Other results

We state other results. First we have $n = 1$ case.

Lemma

If $n = 1$, then the textile set $T_{1,p}$ is the union of a $(p-2)$-dimensional manifold and $2(2^p - 1)$ isolated points.

Example

$U^{**}_{1,3}$ consists of a circle and 14 points:

$$U^{**}_{1,3} = (S^2 \cap \{y_1 + y_2 + y_3 = 1\}) \\ \cup \{\pm(\tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}), \pm(\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}, 0), \pm(\tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}}), \pm(0, \tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}), \\ \pm(1,0,0), \pm(0,1,0), \pm(0,0,1)\}$$

.

Differential geometrical characterization of $f_\lambda{}^{-1}(\mathbf{O})$

Fix $\lambda \geq 0$ arbitrarily. We define the map $f_\lambda : \mathbb{R}^{n\times p} \to \mathbb{R}^{p+1}$ by

$$f_\lambda(\mathbf{y}_1, \ldots, \mathbf{y}_p) := \begin{pmatrix} \sum_j \mathbf{y}_1{}'\mathbf{y}_j - \lambda\|\mathbf{y}_1\|^2 \\ \vdots \\ \sum_j \mathbf{y}_p{}'\mathbf{y}_j - \lambda\|\mathbf{y}_p\|^2 \\ \sum_j \|\mathbf{y}_j\|^2 - 1 \end{pmatrix}.$$

Lemma

We have a classification of $T_{n,p}$, namely

$$T_{n,p} = \bigsqcup_{\lambda\geq 0} f_\lambda{}^{-1}(\mathbf{O}) = \bigsqcup_{0\leq\lambda\leq n} f_\lambda{}^{-1}(\mathbf{O}).$$

Differential geometrical characterization of ${f_\lambda}^{-1}(\mathbf{O})$

Fix $\lambda \geq 0$ arbitrarily. We define the map $f_\lambda : \mathbb{R}^{n\times p} \to \mathbb{R}^{p+1}$ by

$$f_\lambda(\mathbf{y}_1, \ldots, \mathbf{y}_p) := \begin{pmatrix} \sum_j {\mathbf{y}_1}'\mathbf{y}_j - \lambda\|\mathbf{y}_1\|^2 \\ \vdots \\ \sum_j {\mathbf{y}_p}'\mathbf{y}_j - \lambda\|\mathbf{y}_p\|^2 \\ \sum_j \|\mathbf{y}_j\|^2 - 1 \end{pmatrix}.$$

Lemma

We have a classification of $T_{n,p}$, namely

$$T_{n,p} = \bigsqcup_{\lambda\geq 0} {f_\lambda}^{-1}(\mathbf{O}) = \bigsqcup_{0\leq\lambda\leq n} {f_\lambda}^{-1}(\mathbf{O}).$$

Differential geometrical characterization of ${f_\lambda}^{-1}(\mathbf{O})$

Lastly, we state a characterization of ${f_\lambda}^{-1}(\mathbf{O})$ from the viewpoint of differential geometry.

Theorem

Let $\lambda \geq 0$. ${f_\lambda}^{-1}(\mathbf{O})$ is a regular sub-manifold of $\mathbb{R}^{n \times p}$ with codimension $p+1$ whenever

$$\lambda > 0,$$

$$y_{11}y_{jj} - y_{1j}y_{j1} \neq 0, \quad j = 2, \ldots, p,$$

$$\exists \ell \in \{2, \ldots, p\}; \sum_{j=2}^{p} y_{ij} + y_{i\ell}(1 - 2\lambda) \neq 0, \quad i = 1, \ldots, n.$$

Present and future study

Summary:

- We defined the textile set $T_{n,p}$ and find its geometric properties.

Present and future study:

1. Characterize the classification $f_\lambda{}^{-1}(\mathbf{O})$ with induced Riemannian metric from $\mathbb{R}^{np}$ by (global) Riemannian geometry: geodesic, curvature etc.
2. Investigate differential geometrical and topological properties of $T_{n,p}$ and $f_\lambda{}^{-1}(\mathbf{O})$, including its group action.
3. Can one find statistical implication such as sample distribution theory?

Merci beaucoup!

Present and future study

Summary:

- We defined the textile set $T_{n,p}$ and find its geometric properties.

Present and future study:

1. Characterize the classification ${f_\lambda}^{-1}(\mathbf{O})$ with induced Riemannian metric from $\mathbb{R}^{np}$ by (global) Riemannian geometry: geodesic, curvature etc.
2. Investigate differential geometrical and topological properties of $T_{n,p}$ and ${f_\lambda}^{-1}(\mathbf{O})$, including its group action.
3. Can one find statistical implication such as sample distribution theory?

Merci beaucoup!

Present and future study

Summary:

- We defined the textile set $T_{n,p}$ and find its geometric properties.

Present and future study:

1. Characterize the classification ${f_\lambda}^{-1}(\mathbf{O})$ with induced Riemannian metric from $\mathbb{R}^{np}$ by (global) Riemannian geometry: geodesic, curvature etc.
2. Investigate differential geometrical and topological properties of $T_{n,p}$ and ${f_\lambda}^{-1}(\mathbf{O})$, including its group action.
3. Can one find statistical implication such as sample distribution theory?

Merci beaucoup!

References

1. Absil, P.-A., Mahony, R., and Sepulchre, R. (2008), *Optimization Algorithms on Matrix Manifolds*, Princeton University Press.
2. Honda, K. and Nakano, J. (2007), 3 dimensional parallel coordinate plot, *Proceedings of the Institute of Statistical Mathematics*, **55**, 69–83.
3. Inselberg, A. (2009), *Parallel Coordinates: VISUAL Multidimensional Geometry and its Applications*, Springer.
4. Kumasaka, N. and Shibata, R. (2008), High-dimensional data visualisation: The textile plot, *Computational Statistics and Data Analysis*, **52**, 3616–3644.