(STA6557) Final presentation

Principal Component Analysis for Riemannian Manifold

Takuo Ichihara

Outline

1. Principal Component Analysis (PCA) review
2. PCA based on geodesics
3. Distance to geodesics on spheres
4. The PCG omit the intrinsic mean
5. Algorithms for geodesic PCA means on spheres
6. Examples

1. Principal Component Analysis (PCA) review

For simplicity, set the number of parameters is 2.

Let $x_1, \cdots\cdots, x_n$ ： random sample, $x_i = \begin{bmatrix} a_i \\ b_i \end{bmatrix}$, $\bar{x} = \begin{bmatrix} \bar{a} \\ \bar{b} \end{bmatrix}$ (1≦j≦n), a unit vector u=$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$.

Then,

$(x_i' - \bar{x}'\,)$u = $(a_i - \bar{a})u_1 + (b_i - \bar{b})u_2$ ← the length of the projection of $(x_i - \bar{x})$ onto u

Consider the variance of x_i'u

$$\sum_{i=1}^{n}(x_i'u - \bar{x}'u)^2 = \frac{1}{n}\sum_{i=1}^{n}((a_i - \bar{a})u_1 + (b_i - \bar{b})u_2)^2$$

$$= \frac{1}{n}[u_1 \quad u_2]\begin{bmatrix} \sum(a_i - \bar{a})(a_i - \bar{a}) & \sum(b_i - \bar{b})(a_i - \bar{a}) \\ \sum(b_i - \bar{b})(a_i - \bar{a}) & \sum(b_i - \bar{b})(b_i - \bar{b}) \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$= [u_1 \quad u_2]\begin{bmatrix} \frac{1}{n}\sum(a_i - \bar{a})(a_i - \bar{a}) & \frac{1}{n}\sum(b_i - \bar{b})(a_i - \bar{a}) \\ \frac{1}{n}\sum(b_i - \bar{b})(a_i - \bar{a}) & \frac{1}{n}\sum(b_i - \bar{b})(b_i - \bar{b}) \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$= u'\Sigma u$ (where Σ is the covariance matrix)

Next, consider the maximization of u’Σu with the constraint of $u'u = {u_1}^2 + {u_2}^2 = 1$ (or u:unit vector).

By Lagrange multiplier,

$$\frac{\partial}{\partial u}\left(u'\Sigma u + \lambda(1 - u'u)\right) = 0$$

$\Rightarrow 2\Sigma u - 2\lambda u = 0$

$\Rightarrow \Sigma u = \lambda u$

Thus, λ:eigenvalue of Σ , u:eigenvector of Σ

In addition, $u'\Sigma u = u'\lambda u = \lambda u'u = \lambda$

Hence, max(u' Σu) is the largest eigenvalue of Σ, and we call the corresponding eigenvector a 1-st Principal Component (PC).

2-nd PC can be found in the similar way.

Please note that 2-nd PC is orthogonal to the 1-st PC.

The above is the case with 2 parameters, but the case with general p parameters can be reached to eigenvalue problems as well.

The contribution ratio of each PC is

$$\frac{\lambda_1}{\sum_{i=1}^{p}\lambda_i}, \frac{\lambda_2}{\sum_{i=1}^{p}\lambda_i}, \cdots\cdots, \frac{\lambda_p}{\sum_{i=1}^{p}\lambda_i}$$

And the cumulative contribution ratio is

$$\frac{\lambda_1}{\sum_{i=1}^{p}\lambda_i}, \frac{\lambda_1+\lambda_2}{\sum_{i=1}^{p}\lambda_i}, \cdots\cdots, \frac{\sum_{i=1}^{p}\lambda_i}{\sum_{i=1}^{p}\lambda_i} (=1)$$

Next, suppose that we have

$\Sigma u_1 = \lambda_1 u_1$, $\Sigma u_2 = \lambda_2 u_2$,……, $\Sigma u_p = \lambda_p u_p$ (where λ_i:eigenvalue, u_i: eigenvector, $\lambda_1 \geqq \lambda_2 \geqq \cdots \ldots \geqq \lambda_p$)

By spectral decomposition,

$$\Sigma = \lambda_1 u_1 u_1' + \lambda_2 u_2 u_2' + \cdots + \lambda_p u_p u_p{}'$$

Suppose that the cumulative contribution rate until j-th PC is almost 1, then $\lambda_{j+1} \fallingdotseq 0, \cdots\cdots, \lambda_p \fallingdotseq 0$

Thus,

$$\Sigma \fallingdotseq \lambda_1 u_1 u_1' + \lambda_2 u_2 u_2' + \cdots + \lambda_j u_j u_j{}'$$

The above approximation holds when there are some number of large eigenvalues, and other eigenvalues are close to zero.

2. PCA based on geodesics

2-1. Means and PCG

M: m-dimensional Riemannian manifold with induced metric $d(\cdot,\cdot)$

G(M): $\{\gamma : \gamma$ is a geodesic on M of maximal length$\}$

$D(p,\gamma) := \inf_{q\in\gamma} d(p,q)$ where $p\in M$, $\gamma\in G(M)$

X: M-valued random variable

$\bar{p}\in M$ is called an intrinsic mean of X if it minimize

$p\to E[d(X,p)^2]$ ・・・・・・(*)

A geodesic $\gamma_1\in G(M)$ is called a first Principal component geodesic to X if it minimizes

$\gamma\to E[d(X,\gamma)^2]$ ・・・・・・(**)

A geodesic $\gamma_2\in G(M)$ minimizes (**) over all geodesics $\gamma\in G(M)$ that have at least one point in common with γ_1, and orthogonal to γ_1 at all common points a second PCG to X.

Every point $\hat{p}$ that minimizes (*) over all common points of γ_1 and γ_2 is called Principal Component Geodesic Mean (PCGM).

PCGs of higher order are defined analogously.

$\overline{p^{(j)}}\in\gamma_j$ $(1\leqq j\leqq m)$ of the function $p\to E[d(X^{(j)},p)^2]$ on the geodesic γ_j is an intrinsic mean of X on γ_j .

2-2. geodesic variance

Let $\gamma_1, \cdots\cdots, \gamma_m$: PCG

$\bar{p}$: intrinsic mean

$\hat{p}$: PCGM

$\overline{p^{(1)}}$: intrinsic mean of an M-valued random variable X

In Euclidean space, $\hat{p} = \bar{p} = \overline{p^{(1)}}$, and for the total variance

$V_{Eucl}(X) := E[d(X, \bar{p})^2] = \sum_{s=1}^{m} V_{Eucl}^{(s)}(X)$

with the variances explained by the s-th PC ($1 \leqq s \leqq m$) given by

$V_{Eucl}^{(s)}(X) = E[d(X^{(s)}, \bar{p})^2] = E[\frac{1}{m-1}\sum_{j=1}^{m} d(X,\gamma_j)^2 - d(X,\gamma_s)^2]$ ……(*)

We can generalize <u>**the second term**</u> of (*) with the geodesic variance explained by the s-th PCG as obtained by projection.

$V_{gp}^{(s)}(X) = E[d(X^{(s)}, \hat{p})^2] = \sum_{s=1}^{m} V_{gp}^{(s)}(X)$

$V_{gp}(X) = \sum_{s=1}^{m} V_{gp}^{(s)}(X)$

The generalization of <u>**the third term**</u> of (*) with the geodesic variance explained by the s-th PCG as obtained by residuals.

$V_{gr}^{(s)}(X) = E[\frac{1}{m-1}\sum_{j=1}^{m} d(X,\gamma_j)^2 - d(X,\gamma_s)^2]$ ($1 \leqq s \leqq m$)

$V_{gr}(X) = \sum_{s=1}^{m} V_{gr}^{(s)}(X)$

By mixing the two generalizations above,

$V_{gm}(X) := E[d(X^{(1)}, \overline{p^{(1)}})^2] + E[d(X, \gamma_1)^2] = E[d(X^{(1)}, \overline{p^{(1)}})^2] + E[d(X, X^{(1)})^2]$

2-3-1. 1-st PCG

Let Φ_1: $\mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^{2n-2m+1}$, $(x, v) \to \begin{pmatrix} \varphi(x) \\ d\varphi(x)v \\ \langle v, v \rangle - 1 \end{pmatrix}$, $F(x,v) = \sum_{i=1}^{N} d(p_i, \gamma_{x,v})^2$

Finding 1-st PCG is equivalent to solving the following

* finding $(x^*, v^*) \in \mathbb{R}^n \times \mathbb{R}^n$ s.t. $F(x^*, v^*) = \inf\{F(x,v): x, v \in \mathbb{R}^n, \Phi_1(x,v) = 0\}$

(the standard method involves employing Lagrange multipliers, $dF + \lambda' d\Phi_1 = 0$……(a))

2-3-2. 2-nd PCG and PCGM

Given a 1-st PCG $t \to \gamma_{x,v}(t)$, a 2-nd PCG must pass the point $y = \gamma_{x,v}(\tau)$ with an initial direction $w \in T_y M$, orthogonal to $\dot{\gamma}_{x,v}(\tau)$.

with

Φ_2 : $\mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^{n-m+2}$, $(\tau, w) \to \begin{pmatrix} d\varphi(\gamma_{x,v}(\tau))w \\ \langle \dot{\gamma}_{x,v}(\tau) w \rangle \\ \langle w, w \rangle - 1 \end{pmatrix}$ and $F_2(\tau, w) := F(\gamma_{x,v}(\tau), w)$.

Finding 2-nd PCG is equivalent to solving the following

* finding $(\hat{\tau}, \hat{w}) \in \mathbb{R} \times \mathbb{R}^n$ s.t. $F_2(\hat{\tau}, \hat{w}) = \inf\{F_2(\tau, w): \tau \in \mathbb{R}, w \in \mathbb{R}^n, \Phi_2(\tau, w) = 0\}$

(the standard method involves employing Lagrange multipliers, $dF_2 + \lambda' d\Phi_2 = 0$)

Suppose that we have $\hat{\tau}, \hat{w}$.

And let $v_2 := \hat{w}$, $\gamma_{x,v} = \hat{\gamma}_{\hat{x},v}$ where $\hat{x} := \gamma_{x,v}(\hat{\tau})$

$v_1 := \frac{\dot{\gamma}_{x,v}(\hat{\tau})}{|\dot{\gamma}_{x,v}(\hat{\tau})|}$

Note that $\hat{x}$ is a Principal Component Geodesic Mean (PCGM).

2-3-3. Higher order PCG

All PCG of order j (3≦j≦m) pass through the PCGM $\hat{x}$∈M, and v_j is perpendicular to all lower order PCG at $\hat{x}$.

Suppose that we have j-1 PCG ($\gamma_{\hat{x},v_1}$, ……, $\gamma_{\hat{x},v_{j-1}}$)

And defining

$$\Phi_j : \mathbb{R}^n \times \mathbb{R}^{n-m+j} \to \mathbb{R}^{n-m+2}, \quad v \to \begin{pmatrix} d\varphi(\gamma_{x,v}(\tau))w \\ \langle v, v_1 \rangle \\ \vdots \\ \langle v, v_{j-1} \rangle \\ \langle v, v \rangle - 1 \end{pmatrix} \text{ and } F_3(v) := F(\hat{x}, v) .$$

Finding j-th PCG is equivalent to solving the following

* finding v∈$\mathbb{R}^n$ s.t. $F_3(v_j)$ = $\inf\{F_3(v): v \in \mathbb{R}^n, \Phi_j(v) = 0\}$

(the standard method involves employing Lagrange multipliers, $dF_3+\lambda' d\Phi_j=0$)

2-3-4. Intrinsic mean

$G(x) := \sum_{i=1}^{N} d(x, p_i)^2$

Find intrinsic mean is equivalent to solving

* finding $\bar{x}$∈$\mathbb{R}^n$ s.t. $G(\bar{x}) = \inf\{G(x): x \in \mathbb{R}^n,\ \varphi(x) = 0\}$

(the standard method involves employing Lagrange multipliers, $dG+\lambda' d\phi=0$)

2-3-5. Intrinsic mean on a geodesic

Given a geodesic t → $\gamma(t) := \gamma_{x,v}(t)$, best approximating the orthogonal projection onto $\gamma_{x,v}$, q_i of the data points p_i (i=1, …, N) is equivalent to

$G_1(t) := \sum_{i=1}^{N} d(q_i, \gamma_{x,v}(t))^2$

3. Distance to geodesics on spheres

Geodesics on spheres are the great circles given by

$\gamma : t \to \gamma(t) = a \cdot \cos(t) + b \cdot \sin(t)$ for $a, b \in S$, $\langle a, b \rangle = 0$, $t \in \mathbb{R}$

And the Riemannian distance between $z \in S$ and γ is

$d(z, \gamma) = \arccos\sqrt{\langle a, z \rangle^2 + \langle b, z \rangle^2}$

4. The PCG omit the intrinsic mean

(Main theorem)

Let $p_1, p_2, p_3 \in S^2(1)$ are the following.

$p_1 := (\cos\alpha, \sin\alpha, 0)$

$p_2 := (\cos\alpha, -\sin\alpha, 0)$

$p_3 \equiv p_{3\delta} := (\cos\delta, 0, \sin\delta)$

Then, for sufficiently small $\delta > 0$ and α $(0 < \alpha < \frac{\pi}{2})$

(1) There is a unique 1-st PCG to the points

(2) The 1-st PCG does not pass through the intrinsic Frechet mean of the points

(3) $\frac{\delta}{3} < \mu(\delta) = \frac{\delta}{1+2\alpha\cot(\alpha)} + O(\delta^3) < v(\delta) = \frac{\delta}{1+2cos^2\alpha} + O(\delta^3) < \delta$

5. Algorithms for Geodesic PCA and means on sphere

We apply the method to a unit sphere S := $S^m \in \mathbb{R}^{m+1}$ and data points $p_1, \cdots, p_N \in$S.

The unit sphere is defined by ϕ(x) = $\langle x,x \rangle$−1 = 0

and every tangent space is given by T_xS = $\{v \in \mathbb{R}^n : d\varphi(x)v = 2\langle x,v \rangle = 0\}$

Thus, $\Phi_1(x,v)$:= $(\langle x,x \rangle - 1, 2\langle x,v \rangle, \langle v,v \rangle - 1)'$ for x, v$\in \mathbb{R}^{m+1}$

Then, $\gamma_{x,v}$(t) := x・cos(t)+v・sin(t) is a geodesic on S if and only if $\Phi_1(x,v)$ = (0, 0, 0)'

And we have

F(x, v) = $\sum_{i=1}^{N} d(p_i, \gamma(x,v))^2$ = $\sum_{i=1}^{N} arccos^2 \sqrt{\langle x,p_i \rangle^2 + \langle v,p_i \rangle^2}$

G(x) = $\sum_{i=1}^{N} d(p_i, x)^2$ = $\sum_{i=1}^{N} arccos^2 \langle x,p_i \rangle$

5-1. 1-st Principal Component great circle

For convenience, let x^{j_0} = 0 and x^{j_1} ≧ 0 (or $x^{j_1} < 0$) for suitable component indices j_0 and j_1 (1≦j_0≠j_1≦m+1).

By Lagrange multiplier,

ζ_i := $\sqrt{\langle x,p_i \rangle^2 + \langle v,p_i \rangle^2}$, ξ_i := $-\frac{1}{2\zeta_i} \cdot \frac{d}{d\zeta_i} arccos^2 \zeta_i$ = $\frac{arccos\zeta_i}{\zeta_i \sqrt{1-{\zeta_i}^2}}$

By function (a) and λ=(λ_1, λ_2, λ_3),

$\sum_{i=1}^{N} \xi_i \langle x,p_i \rangle p_i$ = $\lambda_1 x + \lambda_2 v$, $\sum_{i=1}^{N} \xi_i \langle v,p_i \rangle p_i$ = $\lambda_2 x + \lambda_3 v$

By solving for x, v

$\sum_{i=1}^{N} \xi_i (\lambda_3 \langle x,p_i \rangle - \lambda_2 \langle v,p_i \rangle) p_i$ = $(\lambda_1 \lambda_3 - {\lambda_2}^2) x$

$\sum_{i=1}^{N} \xi_i (\lambda_2 \langle x,p_i \rangle - \lambda_1 \langle v,p_i \rangle) p_i$ = $({\lambda_2}^2 - \lambda_1 \lambda_3) v$

with

$\sum_{i=1}^{N} \xi_i \langle x,p_i \rangle^2$ = λ_1 , $\sum_{i=1}^{N} \xi_i \langle x,p_i \rangle \langle v,p_i \rangle$ = λ_2 , $\sum_{i=1}^{N} \xi_i \langle v,p_i \rangle^2$ = λ_3

Denoting by $\Psi_1(x,v)$ and $\Psi_2(x,v)$ the two lefthand sides of the fixed-point, we define the algorithm

$(x_n, v_n) \to (x_{n+1}, v_{n+1})$

Where $x_{n+1} = \frac{\Psi_1(x_n,v_n)}{\|\Psi_1(x_n,v_n)\|}$, $v_{n+1} = \frac{\Psi_1(x_n,v_n)-\langle\Psi_2(x_n,v_n),x_{n+1}\rangle x_{n+1}}{\|\Psi_1(x_n,v_n)-\langle\Psi_2(x_n,v_n),x_{n+1}\rangle x_{n+1}\|}$

For the starting direction v_0, either take the normalized part of any vector $x_0 - p_i$ orthogonal to x_0,

Or choose among the vectors $p_i - \bar{p}$ of maximal spherical length and again normalize that part orthogonal to $\bar{p}$.

5-2. The 2-nd Principal Component great circle

Having found a 1-st PCG, $\gamma_1 = \gamma_{x,v}$, determined by x, v∈S, $\langle x,v\rangle = 0$,

Suppose that $\gamma_2(t) = \gamma_{y,w}(t) = ycos(t) + wsint(t)$, with $y = y(\tau) = xcos\tau + vsin\tau$ for a suitable $\tau \in \mathbb{R}$, is a 2-nd PCG.

We consider

$$d\varphi\left(\gamma_{x,v}(\tau)\right)w = \langle 2\left(xcos(\tau) + vsin(\tau)\right), w\rangle = 0$$

$$\langle \dot{\gamma}_{x,v}(\tau), w\rangle = \langle -xsin(\tau) + vcos(\tau), w\rangle = 0$$

The equations are equivalent to

$$\langle x,w\rangle = 0, \langle v,w\rangle = 0$$

Thus, we consider the constraint

$$\Phi_2(w) = (\langle x,w\rangle, \langle v,w\rangle, \langle w,w\rangle - 1)' = 0$$

With the constraint, we minimize

$$F_2(\tau,w) := \sum_{i=1}^{N} d(p_i, \gamma_{y(\tau),w})^2 = \sum_{i=1}^{N} arccos^2\sqrt{(\langle x,p_i\rangle\cos(\tau) + \langle v,p_i\rangle\sin(\tau))^2 + \langle w,p_i\rangle^2}$$

Let

$a_i := a_i(\tau) := \langle v,p_i\rangle\cos(\tau) + \langle v,p_i\rangle\sin(\tau)$

$b_i := b_i(\tau) := \langle v,p_i\rangle\cos(\tau) - \langle x,p_i\rangle\sin(\tau)$

$\zeta_i = \zeta_i(\tau, w) := \sqrt{a_i{}^2 + \langle w, p_i \rangle^2}$

$\xi_i = \xi_i(\tau, w) := \dfrac{\arccos(\zeta_i)}{\zeta\sqrt{1 - \zeta_i{}^2}}$

$\lambda = (2\lambda_1, 2\lambda_2, \lambda_3)'$

By the Lagrange equation

$\sum_{i=1}^{N} \xi_i\, a_i b_i = 0$, $\sum_{i=1}^{N} \xi_i \langle w, p_i \rangle p_i = \lambda_1 x + \lambda_2 v + \lambda_3 w$

Thus,

$\sum_{i=1}^{N} \xi_i \langle w, p_i \rangle \langle x, p_i \rangle = \lambda_1$, $\sum_{i=1}^{N} \xi_i \langle w, p_i \rangle \langle v, p_i \rangle = \lambda_2$, $\sum_{i=1}^{N} \xi_i \langle w, p_i \rangle^2 = \lambda_3$

And let

$\Psi_1(\tau, w) := \frac{\sum_{i=1}^{N} \xi_i \langle w, p_i \rangle p_i}{\sum_{i=1}^{N} \xi_i \langle w, p_i \rangle^2}$, $\Psi_2(\tau, w) := \sum_{i=1}^{N} \xi_i\, a_i b_i$

Then, we obtain the algorithm, start with suitable initial values $\tau^{(0)}$, $w^{(0)}$,

$(\tau^{(0)}, w^{(0)}) = \left(0, \frac{p_1 - p_0 - \langle p_1 - p_0, x \rangle x - \langle p_1 - p_0, v \rangle v}{\| p_1 - p_0 - \langle p_1 - p_0, x \rangle x - \langle p_1 - p_0, v \rangle v \|}\right)$

And compute $(\tau^{(n+1)}, w^{(n+1)})$ from $(\tau^{(n)}, w^{(n)})$ by

$z^{(n+1)} := \Psi_1(\tau^{(n)}, w^{(n)})$

$w^{(n+1)} := \frac{z^{(n+1)} - \langle z^{(n+1)}, x \rangle x - \langle z^{(n+1)}, v \rangle v}{\| z^{(n+1)} - \langle z^{(n+1)}, x \rangle x - \langle z^{(n+1)}, v \rangle v \|}$

$\tau^{(n+1)}$:= (solution to $\Psi_2(\tau, w^{(n+1)}) = 0\ in\ [-\frac{\pi}{2}, \frac{\pi}{2})$

5-3. Higher order principal component great circle

For simplicity, set $x := \hat{x}$

Suppose that we have PCG γ_{x,v_1}, ……, $\gamma_{x,v_{j-1}}$ (3≦j≦m)

And let

$$\zeta_i = \sqrt{\langle x, p_i\rangle^2 + \langle v, p_i\rangle^2} \ , \ \xi_i = \frac{arccos\zeta_i}{\zeta_i\sqrt{1-\zeta_i{}^2}}$$

Then by Lagrange equation,

$$\sum_{i=1}^{N} \xi_i \langle v, p_i\rangle p_i = \lambda_0 x + \sum_{s=1}^{j-1} \lambda_s v_s + \lambda_j v$$

Starting with a suitable v^0, we compute $v^{(n+1)}$ from $v^{(n)}$ using the following algorithm,

$$z^{(n+1)} := \sum_{i=1}^{N} \xi_i{}^{(n)} \langle v^{(n)}, p_i\rangle p_i$$

$$\lambda_0^{(n+1)} := \langle z^{(n+1)}, x\rangle$$

$$\lambda_s^{(n+1)} := \langle z^{(n+1)}, v_s\rangle$$

$$\lambda_j^{(n+1)} := \sum_{i=1}^{N} \xi_i{}^{(n)} \langle v^{(n)}, p_i\rangle^2$$

$$v^{(n+1)} := sgn(\lambda_j) \frac{z^{(n+1)} - \lambda_0^{(n+1)} x - \sum_{s=1}^{j-1} \lambda_s^{(n+1)} v_s}{\left\| z^{(n+1)} - \lambda_0^{(n+1)} x - \sum_{s=1}^{j-1} \lambda_s^{(n+1)} v_s \right\|}$$

5-4. The spherical mean

Set

ζ_i := $\langle x, p_i \rangle$, ξ_i := $\arccos\frac{\zeta_i}{\sqrt{1-\zeta_i{}^2}}$

And by single Lagrange multiplier $\lambda \in \mathbb{R}$,

$\sum_{i=1}^{N} \xi_i \, p_i = \lambda x$ with $\sum_{i=1}^{N} \xi_i \, \langle p_i, x \rangle = \lambda$

By the Lagrange equation with Ψ(x) := $\frac{1}{\lambda}\sum_{i=1}^{N} \xi_i \, p_i$, we gain the following algorithm for intrinsic mean

$x_n \rightarrow x_{n+1} = \frac{\Psi(x_n)}{\|\Psi(x_n)\|}$

5-5. intrinsic mean a great circle

Suppose that we have a spherical geodesic t → $\dot{\gamma}_{x,v}(t)$, determined by $\gamma_{x,v}(0)$ =x and $\dot{\gamma}_{x,v}(0)$ = v.

with

α_i := $\arctan\frac{\langle v, p_i \rangle}{\langle x, p_i \rangle} \in [-\frac{\pi}{2}, \frac{\pi}{2})$

and

[−π, π) ∋ t_i := $\begin{cases} \alpha_i \mod 2\pi & \text{if } \langle v, p_i \rangle \langle x, p_i \rangle > 0 \text{ or } \langle v, p_i \rangle < 0 < \langle x, p_i \rangle \\ \alpha_i + \pi \mod 2\pi & \text{otherwise} \end{cases}$

Then, the geodesic projections of the data points $p_1, \cdots\cdots, p_N$ onto $\gamma_{x,v}$ are given by

q_i = $x cos(t_i) + v sin(t_i)$ (i=1, ……, N)

The function G_1 is given by

$G_1(\tau)$ = $\sum_{i=1}^{N} arccos^2 (\cos(t)\cos(t_i) + \sin(t)\sin(t_i))$ = $\sum_{i=1}^{N} (E_i \, \delta_i (t - t_i) + \frac{1-E_i}{2} 2\pi)^2$

where δ_i = $\mathrm{sgn}(t - t_i)$, E_i = $\mathrm{sgn}(2\pi - |t - t_i|)$

This quantity is uniquely minimized by

$$t = \frac{1}{N}\sum_{i=1}^{N} t_i - \frac{2\pi}{N}\sum_{i=1}^{N} E_i\,\delta_i \frac{1-E_i}{2}$$

The second sum can be any integer between -N to N.

We will determine $t^* := \frac{1}{N}\sum_{i=1}^{N} t_i$ and check which $t^* + \frac{2\pi k}{N}$ (k=0,……,N-1) minimize G_1 .

This value yields an intrinsic mean on the geodesic.

6. Examples

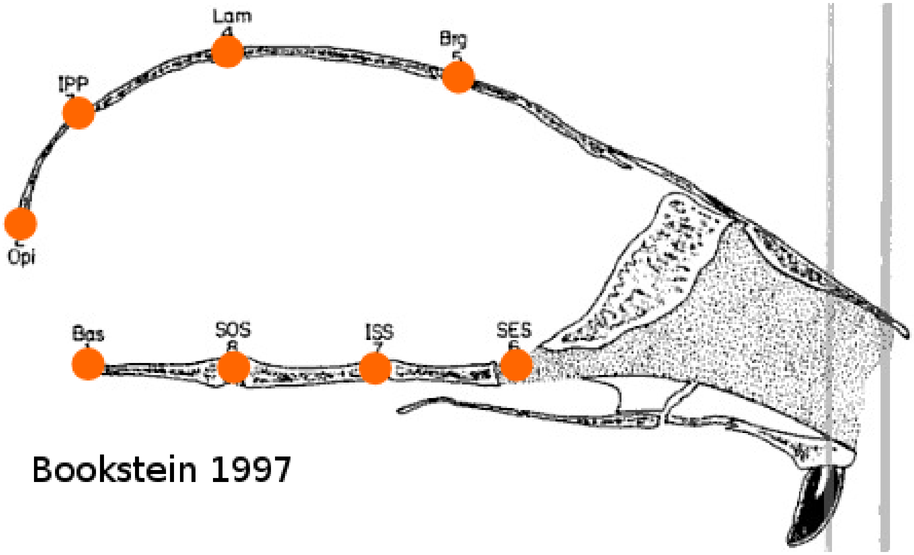
6-1. Rat calvarial landmarks data by Bookstein

ID	Identification number (ID=1,……,21)
Date	Days passed after birth (Date=7,14,21,30,40,60,90,150)

The following landmarks (due to the limitation of the main theorem, we use only Landmark 1,5,6 in this example)

Basion	Landmark 1	Ventral extreme of foremen magnum (tip of basioccipital bone in midplane)
Opisthion	Landmark 2	Dorsal extreme of foremen magnum (tip of supraoccipital bone in midplane)
Interparietal suture	Landmark 3	Interparietal-supraoccipital suture where it crosses the midplane
Lambda	Landmark 4	Parietal-interparietal suture where it crosses the midplane
Bregma	Landmark 5	Frontoparietal (coronal) suture where it crosses the midpoint
Spheno-ethmoid synchondrosis along	Landmark 6	“Middle” of gap between cribriform plate of ethmoid bone and presphenoid bone axis of prephenoid
Interphenoidal suture	Landmark 7	“Middle” of the presphenoid-basisphenoid synchondrosis in the midplane

Spheno-occipital synchondrosis Landmark 8   “Middle” of the basisphenoid-basisphenoid synchondrosis in the midplane



● and solid geodesic represent the 1-st rat.

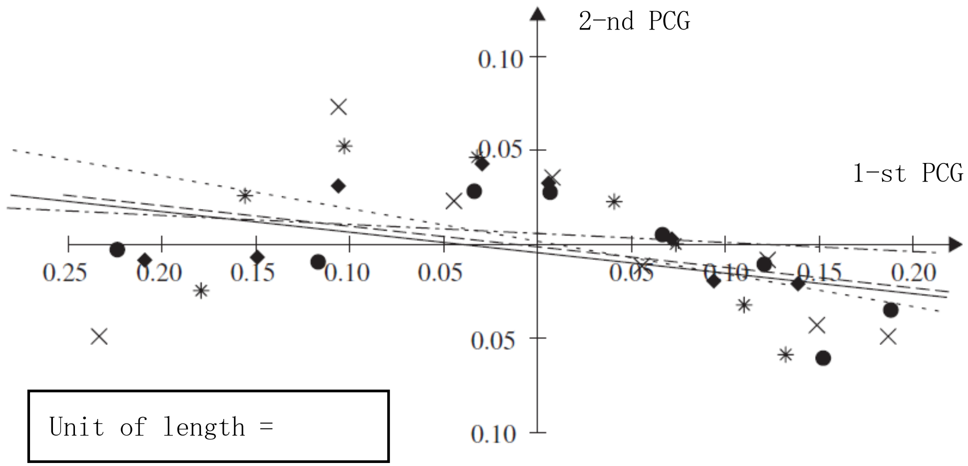
×and long dashed geodesic represent the 2-nd rat.

◆ and dash-dot geodesic represent the 3-rd rat.

＊ and short-dashed geodesic represent the 4-th rat.

The amount of variance explained by 1-st PC is

*95.48% in Euclidean PC



Reference

1. PRINCIPAL COMPONENT ANALYSIS FOR RIEMANNIAN MANIFOLDS, WITH AN APPLICATION TO TRIANGULAR SHAPE SPACES - STEPHAN HUCKEMANN and HERBERT ZIEZOLD - Adv. Appl. Prob. (SGSA) 38, 299-319 (2006)

2. Bookstein, F.L. (1991). Morphometric Tools for Landmark Data: Geometry and Biology. Cambridge University Press.

3. Pattern Recognition and Machine Learning by Christopher M. Bishop