Štatistická analýza tvaru a obraz

Mnohorozmerné štatistické metódy

Stanislav Katina

[1]Ústav matematiky a statistiky
Přírodovědecká fakulta
Masarykova Univerzita v Brne

Tento učební text vznikl za přispění Evropského sociálního fondu a státního rozpočtu ČR prostřednictvím Operačního programu Vzdělávání pro konkurenceschopnost v rámci projektu Univerzitní výuka matematiky v měnícím se světě (CZ.1.07/2.2.00/15.0203).

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Distance-based PCA

Classical PCA

Definition (Distance-based PCA)

The ***Principal Component Analysis*** (PCA) finds a set of standardized linear combinations, called ***principal components*** (PCs), which are orthogonal and taken together explain all the variance of the random vector

$$\mathbf{X}_{k\times 1} = (X_1, ..., X_k)^T \text{ with } E(\mathbf{X}) = \boldsymbol{\mu}_{\mathbf{X}} \text{ and } Var(\mathbf{X}) = \boldsymbol{\Sigma}_{\mathbf{X}},$$

where $\boldsymbol{\mu}_{\mathbf{X}}$ is k-vector and $\boldsymbol{\Sigma}_{\mathbf{X}}$ is $k \times k$ matrix. Let $\mathbf{X}_i^T, i = 1, 2, ...n$ be a random sample of k-vectors (the rows of $\mathbf{X}_{n\times k}$), where $k \leq n - 1$. Then the ***principal component transformation*** is defined as

$$\mathbf{X}_{n\times k} \to \mathbf{Y}_{n\times k} = \left(\mathbf{X}_{n\times k} - \mathbf{1}_n \boldsymbol{\mu}_{\mathbf{X}}^T\right) \boldsymbol{\Gamma}_{k\times k},$$

where $\boldsymbol{\Gamma}$ is orthogonal, $\boldsymbol{\Gamma}^T \boldsymbol{\Sigma}_{\mathbf{X}} \boldsymbol{\Gamma} = \boldsymbol{\Lambda} = diag(\lambda_1, ...\lambda_k)$, $\lambda_1 \geq ... \geq \lambda_k \geq 0$, $\lambda_j, j = 1, 2, ...k$ are ***eigenvalues*** of $\boldsymbol{\Gamma}$ and γ_j (jth column of $\boldsymbol{\Gamma}$) are ***eigenvectors*** of $\boldsymbol{\Gamma}$. The jth PC of $\mathbf{X}_{n\times k}$ is defined as jth column of $\mathbf{Y}_{n\times k}$ by equation $\mathbf{Y}_j = \left(\mathbf{X}_{n\times k} - \mathbf{1}_n \boldsymbol{\mu}_{\mathbf{X}}^T\right) \gamma_j$, where γ_j is the jth column of $\boldsymbol{\Gamma}$ and is called jth vector of ***PC loadings***, and $R_{ij} = Y_{ij}, i = 1, 2, ...n$ are ***PC scores*** of ith individual ($R_{ij} = Y_{ij}$ is ith element of n-vector $\mathbf{Y}_j$).

Distance-based PCA

Classical PCA

Definition (Distance-based PCA; cont.)

SVD of ***covariance matrix*** $\boldsymbol{\Sigma}_{\mathbf{X}}$ is defined as follows

$$\boldsymbol{\Sigma}_{\mathbf{X}} = \boldsymbol{\Gamma}\boldsymbol{\Lambda}\boldsymbol{\Gamma}^T = \sum_{j=1}^{k} \lambda_j \gamma_j \gamma_j^T.$$

Let $\mathbf{H} = \mathbf{I} - \left(\frac{1}{n}\mathbf{1}\mathbf{1}^T\right)$ be ***centring matrix***, then SVD of $\boldsymbol{\Sigma}_{\mathbf{Y}}$ can be written as

$$\boldsymbol{\Sigma}_{\mathbf{Y}} = \frac{1}{n}\mathbf{Y}^T\mathbf{H}\mathbf{Y} = \frac{1}{n}\boldsymbol{\Gamma}^T \left(\mathbf{X} - \mathbf{1}\boldsymbol{\mu}_{\mathbf{X}}^T\right) \mathbf{H} \left(\mathbf{X} - \mathbf{1}\boldsymbol{\mu}_{\mathbf{X}}^T\right) \boldsymbol{\Gamma} = \frac{1}{n}\boldsymbol{\Gamma}^T\mathbf{X}^T\mathbf{H}\mathbf{X}\boldsymbol{\Gamma} = \boldsymbol{\Gamma}^T\boldsymbol{\Sigma}_{\mathbf{X}}\boldsymbol{\Gamma}.$$

If $\mathbf{X} = (X_1, ..., X_k)^T \sim N_k(\boldsymbol{\mu}_{\mathbf{X}}, \boldsymbol{\Sigma}_{\mathbf{X}})$, then

1. $E(Y_j) = 0$ and $Var(Y_j) = \gamma_j^T \boldsymbol{\Sigma}_{\mathbf{X}} \gamma_j = \lambda_j$
2. ***covariance of transformed variables*** is equal to $Cov(Y_i, Y_j) = \gamma_i^T \boldsymbol{\Sigma}_{\mathbf{X}} \gamma_j = \lambda_j \gamma_i^T \gamma_j = 0, i \neq j$, $Var(Y_1) \geq Var(Y_2) \geq ... \geq Var(Y_k)$, $\boldsymbol{\Sigma}_{\mathbf{X}} \gamma_j = \lambda_j \gamma_j$
3. ***covariance of original and transformed variables*** $Cov(X_i, Y_j) = \gamma_{ij} \lambda_j$
4. ***correlation coefficient*** $\rho(X_i, Y_j) = \left(\gamma_{ij} \sqrt{\lambda_j / (\boldsymbol{\Sigma}_{\mathbf{X}})_{ii}}\right); i, j = 1, 2, ...k$

Distance-based PCA

Classical PCA

Definition (Distance-based PCA; cont.)

Total variance is equal to

$$tr(\boldsymbol{\Sigma}_{\mathbf{X}}) = tr\left(\boldsymbol{\Gamma}\boldsymbol{\Lambda}\boldsymbol{\Gamma}^T\right) = tr(\boldsymbol{\Lambda}) = \sum_{j=1}^{k} \lambda_j,$$

and ***generalized variance***

$$\det(\boldsymbol{\Sigma}_{\mathbf{X}}) = \prod_{j=1}^{k} \lambda_j.$$

Definition (Distance-based PCA; cont.)

If $\widehat{\mu}_{\mathbf{X}} = \overline{\mathbf{x}}$, $\widehat{\mathbf{\Sigma}}_{\mathbf{X}} = \mathbf{S}_{\mathbf{X}} = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{x}_i - \overline{\mathbf{x}})(\mathbf{x}_i - \overline{\mathbf{x}})^T = \frac{1}{n}\mathbf{X}^T\mathbf{H}\mathbf{X}$, then ***sample PCs*** are defined as follows

$$\widehat{\mathbf{Y}}_{n\times p} = \left(\mathbf{X}_{n\times p} - \mathbf{1}_n\overline{\mathbf{x}}^T\right)\widehat{\mathbf{\Gamma}},$$

$\mathbf{S}_{\mathbf{X}} = \widehat{\mathbf{\Gamma}}\widehat{\mathbf{\Lambda}}\widehat{\mathbf{\Gamma}}^T = \sum_{j=1}^{p}\widehat{\lambda}_j\widehat{\gamma}_j\widehat{\gamma}_j^T$, $\widehat{\mathbf{\Sigma}}_{\mathbf{Y}} = \widehat{\mathbf{\Gamma}}^T\mathbf{S}_{\mathbf{X}}\widehat{\mathbf{\Gamma}}$. If $\mathbf{X} = (X_1, ..., X_k)^T \sim N_k(\mu_{\mathbf{X}}, \mathbf{\Sigma}_{\mathbf{X}})$, then

1. $\overline{\mathbf{y}}_j = 0$ and $\widehat{Var}(\widehat{\mathbf{y}}_j) = \widehat{\gamma}_j^T\mathbf{S}_{\mathbf{X}}\widehat{\gamma}_j = \widehat{\lambda}_j$
2. ***sample covariance of transformed variables*** is equal to
   $\widehat{Cov}(\widehat{\mathbf{y}}_i, \widehat{\mathbf{y}}_j) = \widehat{\gamma}_i^T\mathbf{S}_{\mathbf{X}}\widehat{\gamma}_j = \lambda_j\widehat{\gamma}_i^T\widehat{\gamma}_j = 0, i \neq j$,
   $\widehat{Var}(\widehat{\mathbf{y}}_1) \geq \widehat{Var}(\widehat{\mathbf{y}}_2) \geq ... \geq \widehat{Var}(\widehat{\mathbf{y}}_k)$, $\mathbf{S}_{\mathbf{X}}\widehat{\gamma}_j = \widehat{\lambda}_j\widehat{\gamma}_j$
3. ***sample covariance of original and transformed variables***
   $\widehat{Cov}(\mathbf{x}_i, \widehat{\mathbf{y}}_j) = \widehat{\gamma}_{ij}\widehat{\lambda}_j$
4. ***sample correlation coefficient***
   $\widehat{\rho(\mathbf{x}_i, \mathbf{y}_j)} = r(\mathbf{x}_i, \widehat{\mathbf{y}}_j) = \left(\widehat{\gamma}_{ij}\sqrt{\widehat{\lambda}_j / (\mathbf{S}_{\mathbf{X}})_{ii}}\right); i, j = 1, 2, ...k$

Definition (Spatial PCA)

- let $\mathbf{y}_i$ represent a k-vector of ***EEG responses*** for individual $i, i = 1, 2, ...n$, measured in k ***sensor locations*** on the human head in $\mathbb{R}^3$ (projected to $\mathbb{R}^2$, in our case)
- in general, these sensor locations might be different for each individual—but here, we consider their $x^{(1)}$- and $x^{(2)}$-coordinates be the same and form a $k \times 2$ matrix $\mathbf{X}$
- with respect to $\mathbf{X}$, $\mathbf{y}_i$ are y-coordinates of the surface $(x_{ij}^{(1)}, x_{ij}^{(2)}, y_{ij})$, $j = 1, 2, ...k$. Let $\overline{\mathbf{y}}$ be ***mean response***
- ***spatial PCA*** is generalized PCA, where PCs are calculated with respect to the bending energy matrix $\mathbf{B}_e$ or its inverse
- consider a random sample of n surface values (here EEG/ERP values) $\mathbf{y}_i = (y_{i1}, y_{i2}, \ldots y_{ik})^T$, $i = 1, 2, ...n$
- the ***bending energy matrix*** $\mathbf{B}_e$ is calculated for the mean position of the electrodes $\overline{\mathbf{X}}$ (here fixed position $\mathbf{X}$ of the electrodes on the head
- let $\widehat{\mathbf{\Sigma}} = \frac{1}{n}\mathbf{Y}_c^T\mathbf{Y}_c$ be $k \times k$ ***sample covariance matrix***, where ith row of $\mathbf{Y}_c$ is equal to $\mathbf{y}_{ic} = \mathbf{y}_i - \overline{\mathbf{y}}$

Definition (Spatial PCA, cont.)

- let
  $$\widehat{\mathbf{\Sigma}}_B = \left(\mathbf{B}_e^-\right)^{\alpha/2}\widehat{\mathbf{\Sigma}}\left(\mathbf{B}_e^-\right)^{\alpha/2}$$
  be the sample covariance matrix of $\left(\mathbf{B}_e^-\right)^{\alpha/2}\mathbf{y}_{ic}$, i.e. ***generalized sample covariance matrix*** of $\mathbf{y}_{ic}$
- the non-zero ***eigenvalues*** of $\widehat{\mathbf{\Sigma}}_B$ are $\widehat{l}_j$ with corresponding ***eigenvectors*** $\widehat{\mathbf{g}}_j$ (***PC loadings***)
- ***Moore-Penrose generalized inverse*** of $\mathbf{B}_e^{\alpha/2}$, $\left(\mathbf{B}_e^-\right)^{\alpha/2} = \sum_j \widehat{\lambda}_j^{-\alpha/2}\widehat{\gamma}_j^T\widehat{\gamma}_j$
- the ***PC scores*** are
  $$r_{ij} = \widehat{\mathbf{g}}_j^T\left(\mathbf{B}_e^-\right)^{\alpha/2}\mathbf{y}_{ic}; i = 1, 2, ...n; j = 1, 2, ...k$$

Definition (Spatial PCA, cont.)

- PCs and PC scores are useful tools for describing the non-affine surface variation in particular, ***the effect of the jth PC*** can be viewed by plotting
  $$\mathbf{y}(c_j, j, \alpha) = \overline{\mathbf{y}} \pm c_j\mathbf{B}_e^{\alpha/2}\widehat{\mathbf{g}}_j\widehat{l}_j^{1/2}, r_j = c_j\widehat{l}_j^{1/2}$$
  for various values of $r_j \in \langle 0, \max(|r_{ij}|)\rangle$ (or reasonable magnification of $\max(|r_{ij}|)$; alternatively, fixing $c_j = 1$, magnification of $\widehat{l}_j^{1/2}$, standard deviation of PC$_j$ scores), where $\mathbf{B}_e^{\alpha/2} = \sum_j \widehat{\lambda}_j^{\alpha/2}\widehat{\gamma}_j\widehat{\gamma}_j^T$
- to emphasize ***large scale variability*** (***global bending***), $\alpha = 1$
- for ***small scale variability*** (***local bending***), $\alpha = -1$, and
- if $\alpha = 0$, then we take $\mathbf{B}_e^0 = \mathbf{I}$ as the $k \times k$ identity matrix and the procedure is exactly the same as classical PCA
- ***visualization the effect of each PC***—grid of gray-scale rectangles with colors corresponding to the surface values with superimposed contours built up based on TPS, where the fixed positions of the electrodes were re-sampled in the convex hull data-space

Definition (Spatial PCA, cont.)

- a ***PC summary of the surface data***

$$\mathbf{y}_i = \overline{\mathbf{y}} + \mathbf{B}_e^{\alpha/2} \sum_{j=1}^{k} r_{ij} \widehat{\mathbf{g}}_j = \overline{\mathbf{y}} + \sum_{j=1}^{k} \widehat{\mathbf{g}}_j \widehat{\mathbf{g}}_j^T \mathbf{y}_{ic}, i = 1, ...n$$

- ***PC summary for any q-subset of PCs*** $\{PC_{j_1}, ... PC_{j_q}\}$, $q \geq 1$, can be written as

$$\mathbf{y}_i(PC_{(j_1,...j_q)}) = \overline{\mathbf{y}} + \mathbf{B}_e^{\alpha/2} \sum_{j_1,...j_q} r_{ij} \widehat{\mathbf{g}}_j, i = 1, ...n$$

and then $\mathbf{Y}_{PC(j_1,...j_q)}$ is the matrix of $\mathbf{y}_i(PC_{(j_1,...j_q)})$

Definition (Spatial PCA, cont.)

- ***affine contributions to the variability***—an ***affine subspace PCA*** on the $n \times k$ matrix $\mathbf{Y}_A$ with the rows $\mathbf{y}_{A,i}$, $i = 1, 2, ...n$
- ***non-affine contribution to the variability*** —a ***non-affine subspace PCA*** on the $n \times k$ matrix $\mathbf{Y}_{NA}$ with the rows $\mathbf{y}_{NA,i}$, $i = 1, 2, ...n$
- ***in affine subspace***, $\widehat{\mathbf{\Sigma}}_A$ stands for ***sample covariance matrix*** of $\widehat{\mathbf{y}}_i$ and spatial PCA is calculated with respect to bending energy matrix $\mathbf{B}_e^0 = \mathbf{I}$;
- ***in non-affine subspace***, we have $\left(\mathbf{B}_e^-\right)^{\alpha/2} \widehat{\mathbf{\Sigma}}_{NA} \left(\mathbf{B}_e^-\right)^{\alpha/2}$, because pure bending is independent of affine component
- to find the affine component we use ***linear regression model (LRM)*** $\mathbf{y}_i = \overline{\mathbf{y}}\beta_i + \epsilon_i$, where $\widehat{\beta}_i = (\overline{\mathbf{y}}^T\overline{\mathbf{y}})^{-1}\overline{\mathbf{y}}^T\mathbf{y}_i$, $\mathbf{y}_i, i = 1, 2, ...n$, are the rows of $n \times k$ matrix $\mathbf{Y}$; then $\mathbf{y}_{A,i} = \overline{\mathbf{y}}\widehat{\beta}_i$ is the ***affine component***; finally, we get ***non-affine component (residuals of LRM)***, $\mathbf{y}_{NA,i} = \mathbf{y}_i - \mathbf{y}_{A,i}$

Definition (Spatial PCA, cont.)

Consider the ***null model***

$$\mathbf{y} = \boldsymbol{\mu} + \boldsymbol{\epsilon},$$

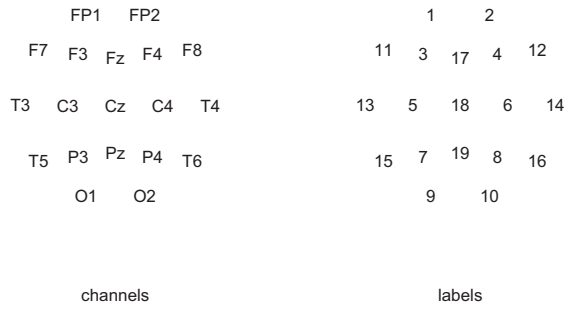
where $\mathbf{y} = (y_1, y_i, \ldots y_k)^T$, $\boldsymbol{\epsilon} \sim N_k(\mathbf{0}, \mathbf{\Sigma}_B)$ Special case of this model can be written as

$$\mathbf{y} = \boldsymbol{\mu} + \mathbf{B}_e^{\alpha/2} \sum_{j=1}^{q} c_j \mathbf{g}_j l_j^{1/2} + \boldsymbol{\epsilon},$$

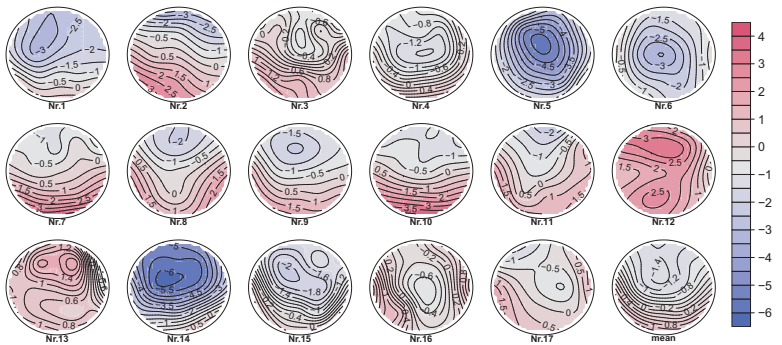
where $c_j \sim N(0, 1)$, $\boldsymbol{\epsilon} \sim N_k(\mathbf{0}, \sigma^2 \mathbf{I}_{k\times k})$ independently, $\boldsymbol{\mu}\mathbf{g}_j = 0, \mathbf{g}_j^T\mathbf{g}_j = 1$ and $\mathbf{g}_i^T\mathbf{g}_j = 0$ $(i \neq j)$. Then

$$\widehat{\mathbf{\Sigma}}_B = \sum_{i=1}^{q} l_j \mathbf{g}_j \mathbf{g}_j^T + \sigma^2 \mathbf{I}_{k\times k}.$$

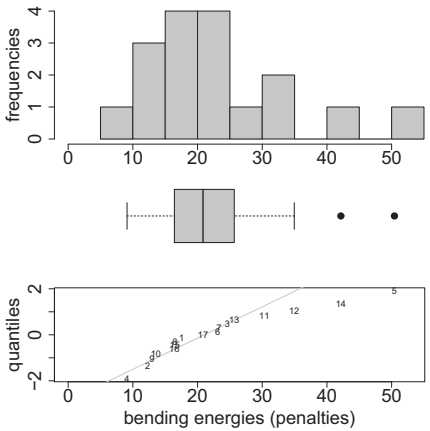
Finally, $\mathbf{\Sigma}_B$ can be estimated by $\widehat{\mathbf{\Sigma}}_B$ and $\widehat{\sigma}^2 = \frac{1}{k-q} \sum_{j=q+1}^{k} l_j$



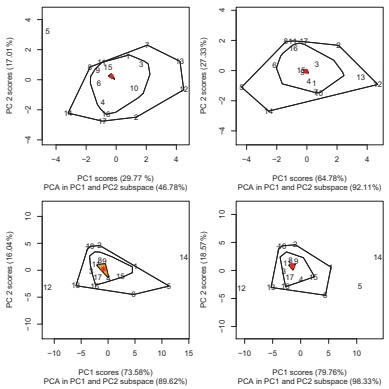
Obrázok: UI 10-20 systém pozícií elektród EEG s $k = 19$ elektródami



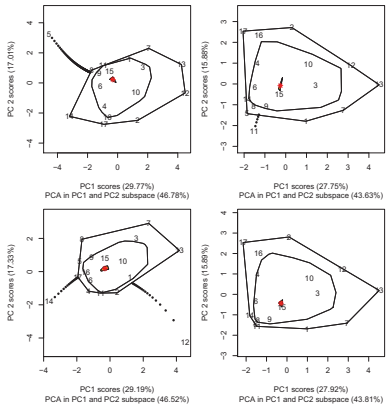
Obrázok: TPS sieť farebných štvoruholníkov s farbami korešpondujúcimi vyhladeným hodnotám plochy superponovanými kontúrami (použitím optimálnej λ vypočítanej pomocou GCV)



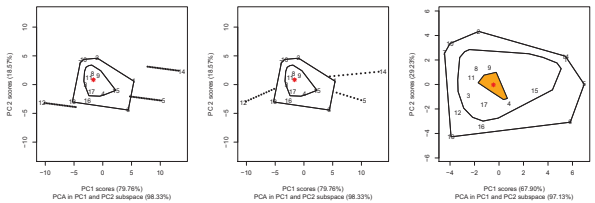
Obrázok: Histogram, boxplot, and quantile plot of penalties (bending energies; outliers—Nr.14 and 5)



Obrázok: Spatial PCA—***PCA of local bending patterns*** (outlier Nr.5; upper left), ***classical PCA*** (outliers Nr.12 and 14; bottom left), ***global bending patterns*** (upper right), and ***PCA in the affine subspace*** (outlier Nr.5, 12, and 14; bottom right)



Obrázok: Iterative process of outlier detection and relaxation in the subspace of first two PCs of local bending patterns with '*curves décolletage*'—first PCA (outlier Nr.5; upper left), second PCA (outlier Nr.12 and 14; upper right), third PCA (outlier Nr.11; bottom left), final PCA (without outliers; bottom right)

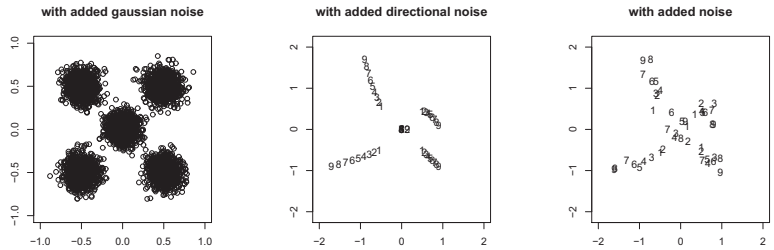


Obrázok: Iterative process of outlier detection and relaxation in the subspace of first two affine PCs with '*curves décolletage*'—initial PCA with incorrect relaxation direction (outlier Nr.5, 12, and 14; upper left), initial PCA with correct relaxation direction (outlier Nr.5, 12, and 14; upper right), final PCA (without outliers; bottom)

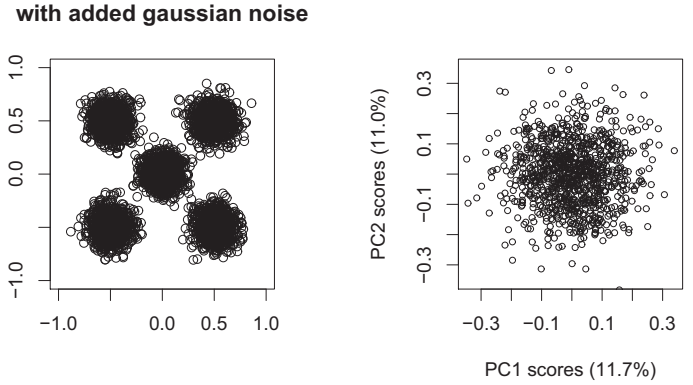
Euler angles ψ, θ, and ϕ in degrees (clockwise around $x^{(1)}$-, $x^{(2)}$- and y-axis) of original and ***affine-relaxed surfaces*** (***OLS planes***) were calculated from a 3D rotation matrix. Additionally, ***translation in absolute and relative scale*** (in the range of **y** values including whole sample) was calculated as a difference of original and affine-relaxed surface centres.

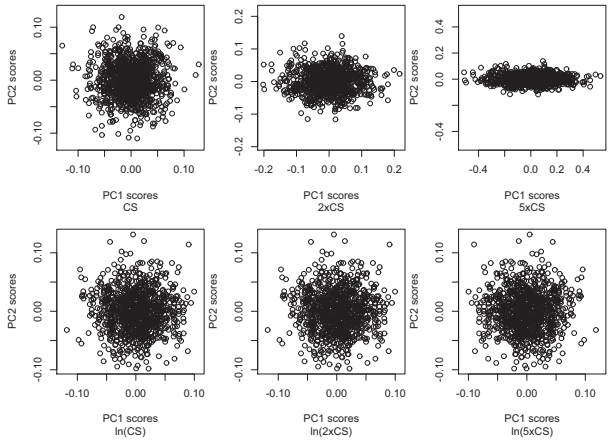
Tabuľka: ***Affine outliers*—angles of rotation about particular axes** (clockwise, in degrees)—ψ about $x^{(1)}$-axis, θ about $x^{(2)}$-axis, ϕ about y-axis; ***translation of surface centers*** in absolute (t.abs) and relative (t.relat; in % of the range of **y** of the whole sample) scale

outliers	ψ	θ	ϕ	t.abs	t.relat
Nr. 5	−0.65	0.15	−0.07	1.00	13%
Nr. 12	−12.22	−3.24	4.72	−1.37	−17%
Nr. 14	2.00	1.26	−1.38	1.81	23%

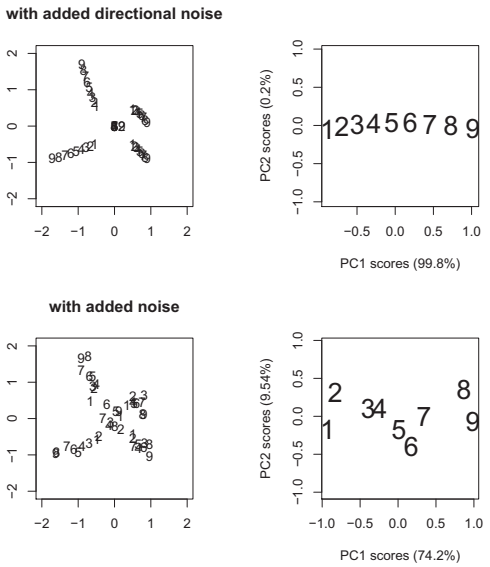


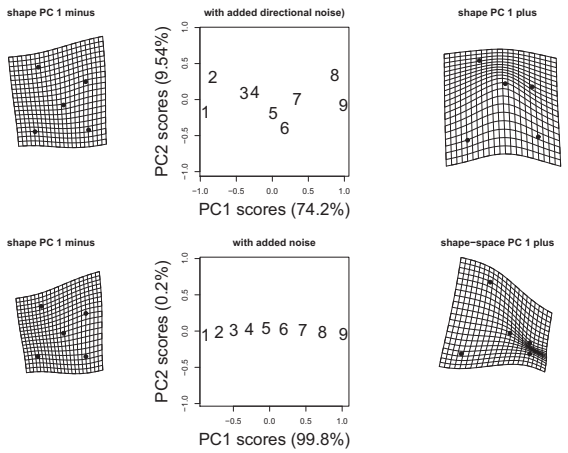
Obrázok: 250 quints generated from a normally distributed sequence of 1000 random numbers—$\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4)$, $\mathbf{X}_i \sim N_8\left(\mu_{\mathbf{X}}, \sigma^2 \mathbf{I}_{8\times 8}\right)$, $\mu_1 = (-1, 0)$, $\mu_2 = (0, 1)$, $\mu_3 = (1, 0)$, $\mu_4 = (0, -1)$, and $\mu_5 = (0, 0)$, $\sigma^2 = 0.001$ (left); 9 quints with different random noise (middle, right)

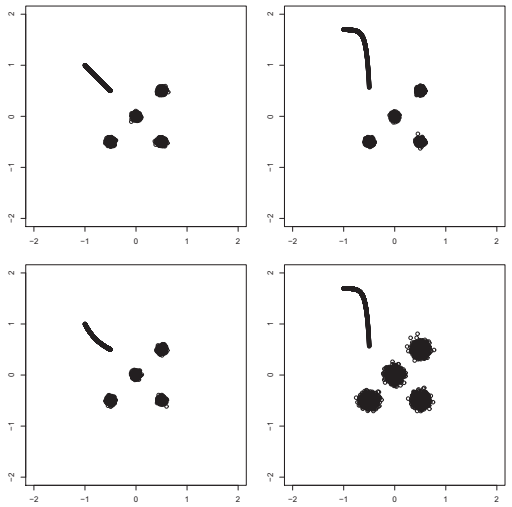




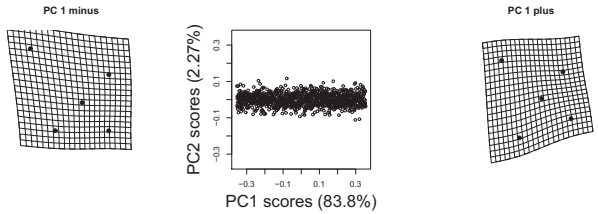
Obrázok: Procrustes form space with $k \times CS$ (first row) and $\ln(k \times CS)$ (second row), $k = 1, 2, 5$

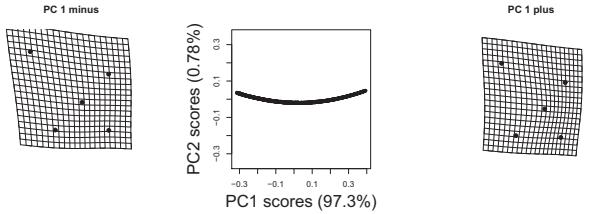




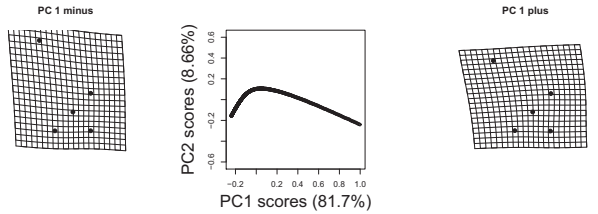


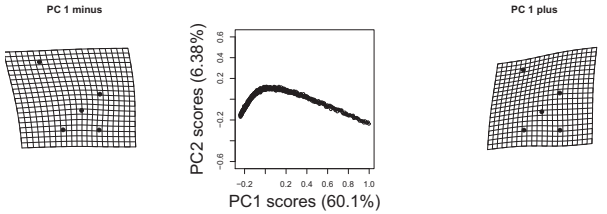
GMM
Simulations—quint examples





GMM
Simulations— quint example





GMM
Key knowledge

1. **form**—information about object geometry that remains after translation and rotation effects are removed
2. **shape**—information about object geometry that remains after translation, rotation, and size effects are removed
3. **object geometry**—2D/3D Cartesian coordinates in $k \times d$ configuration matrix $\mathbf{X}$
4. **shape components**—***affine*** (uniform) $\mathbf{X}_A$, ***non-affine*** (nonuniform) $\mathbf{X}_{NA}$ [***local benging*** and ***global bending***]
5. **biological homology**—biologically correspondent parts of an organism but point locations with respect to deformation TPS model—**landmarks**
6. **geometrical homology**—with respect to some minimization criteria (***bending energy of TPS model***) between source and target configuration— **semilandmarks on curves and surfaces**
7. **vectorization**—Vectorized $\mathbf{X} = (\mathbf{x}^{(1)} \vdots \mathbf{x}^{(2)} \vdots \ldots \vdots \mathbf{x}^{(d)})$ is defined as $Vec(\mathbf{X}) = \mathbf{x} = (\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots \mathbf{x}^{(d)})$, then $\mathbf{X}_S$ is $n \times dk$ matrix of **vectorized Procrustes shape coordinates** $Vec(\mathbf{X}_{P,i}) = \mathbf{x}_{P,i}$ as its rows and its covariance matrix is written as $\mathbf{S}$

Geometric Morphometrics
Generalized Procrustes Analysis—Procrustes k-point registration

Definition (Generalized Procrustes Analysis, GPA)

Procrustes form coordinates $\mathbf{x}_{f,ij} = \boldsymbol{\Gamma}_i(\mathbf{x}_{ij} - \mathbf{t}_i)$, where $\boldsymbol{\Gamma}_i$ is *rotation matrix* and $\mathbf{t}_i$ is *translation*, $\mathbf{x}_{f,ij}$ are rows of $\mathbf{X}_{f,i}$, $i = 1, \ldots n$. Then we say that $\mathbf{X}_i, i = 1, 2, \ldots n$ are in *optimal position* or have ***the best Procrustes fit*** in the sense of 'form' if

$$\arg\inf \sum_{1 \le i \le j \le n} \| \mathbf{X}_{f,i} - \mathbf{X}_{f,j} \|^2 =$$

$$\underbrace{\arg\inf}_{\substack{\boldsymbol{\Gamma}_1, \ldots \boldsymbol{\Gamma}_n \in SO(2) \\ \mathbf{t}_1, \ldots \mathbf{t}_n \in \mathbb{R}^d}} \left\{ \sum_{1 \le i \le j \le n} \left\| \boldsymbol{\Gamma}_i \left(\mathbf{X}_i - \mathbf{1}_k \mathbf{t}_i^T \right)^T - \boldsymbol{\Gamma}_j \left(\mathbf{X}_j - \mathbf{1}_k \mathbf{t}_j^T \right)^T \right\|^2 \right\}$$

Geometric Morphometrics
Generalized Procrustes Analysis—Procrustes k-point registration

Definition (Generalized Procrustes Analysis, GPA)

Procrustes shape coordinates $\mathbf{x}_{P,ij} = c_i \mathbf{\Gamma}_i (\mathbf{x}_{ij} - \mathbf{t}_i)$, where c_i is *scale*, $\mathbf{\Gamma}_i$ is *rotation matrix* and $\mathbf{t}_i$ is *translation*, $\mathbf{x}_{P,ij}$ are rows of $\mathbf{X}_{P,i}$, $i = 1, ..., n$. Then we say that $\mathbf{X}_i, i = 1, 2, ..., n$ are in *optimal position* or have ***the best Procrustes fit*** in the sense of 'shape' if

$$\arg\inf \sum_{1 \le i \le j \le n} \| \mathbf{X}_{P,i} - \mathbf{X}_{P,j} \|^2 =$$

$$\underbrace{\arg\inf}_{\substack{\mathbf{\Gamma}_1, ... \mathbf{\Gamma}_n \in SO(2) \\ \mathbf{t}_1, ... \mathbf{t}_n \in \mathbb{R}^d, c_1, c_2, ... c_n \in \mathbb{R}^+}} \left\{ \sum_{1 \le i \le j \le n} \left\| c_i \mathbf{\Gamma}_i \left(\mathbf{X}_i - \mathbf{1}_k \mathbf{t}_i^T \right)^T - c_j \mathbf{\Gamma}_j \left(\mathbf{X}_j - \mathbf{1}_k \mathbf{t}_j^T \right)^T \right\|^2 \right\}$$

Interpolation TPS Model

Definition (Thin-Plate Spline (TPS))

Consider a **TPS** given by $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \dots f_d(\mathbf{x}))$, where $\mathbf{f}(\mathbf{x}) = \mathbf{c} + \mathbf{A}^T\mathbf{x} + \mathbf{W}^T\mathbf{s}(\mathbf{x})$, $f_m(\mathbf{x}) = c_m + \mathbf{a}_m^T\mathbf{x} + \sum_{j=1}^k w_{jm}\phi_j(\mathbf{x})$, where $m = 1, 2, ...d$, $\mathbf{c} = (c_1, c_2, ...c_d)^T$, $\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, ...\mathbf{a}_d)$, $\mathbf{w}_m = (w_{1m}, w_{2m}, \dots w_{km})^T$, $\mathbf{W} = (\mathbf{w}_1, \mathbf{w}_2, ...\mathbf{w}_d)$, $\mathbf{s}(\mathbf{x})_{k\times 1} = [\phi_1(\mathbf{x}), \dots \phi_k(\mathbf{x})]^T$, continuous **radial (nodal) basis function**

$$\phi(\mathbf{x}) = \begin{cases} \|\mathbf{x}\|_2^2 \log\left(\|\mathbf{x}\|_2^2\right), \forall \|\mathbf{x}\|_2 > 0 & \text{if } d = 2 \\ 0, \forall \|\mathbf{x}\|_2 = 0 & \text{if } d = 2 \\ \|\mathbf{x}\|_2, & \text{if } d = 3 \end{cases}$$

TPS interpolation to the data $(\mathbf{x}_j, \mathbf{y}_j)$ is defined as

$$\begin{pmatrix} \mathbf{Y} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{S} & \mathbf{1}_k & \mathbf{X} \\ \mathbf{1}_k^T & 0 & \mathbf{0} \\ \mathbf{X}^T & \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{W} \\ \mathbf{c}^T \\ \mathbf{A} \end{pmatrix}, \mathbf{L} = \begin{pmatrix} \mathbf{S} & \mathbf{1}_k & \mathbf{X} \\ \mathbf{1}_k^T & 0 & \mathbf{0} \\ \mathbf{X}^T & \mathbf{0} & \mathbf{0} \end{pmatrix},$$

where $\mathbf{Y}_{k\times d} = (\mathbf{y}_1, \dots \mathbf{y}_k)^T$ and $\mathbf{X}_{k\times d} = (\mathbf{x}_1, \dots \mathbf{x}_k)^T$, $(\mathbf{S})_{ij} = \phi_j(\mathbf{x}_i) = \phi(\mathbf{x}_i - \mathbf{x}_j)$, $i, j = 1, 2, ...k$.

Interpolation TPS model

Definition (Thin-Plate Spline (TPS), cont.)

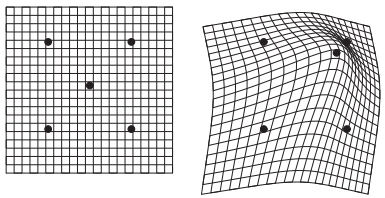
Inverse of $\mathbf{L}$ is equal to

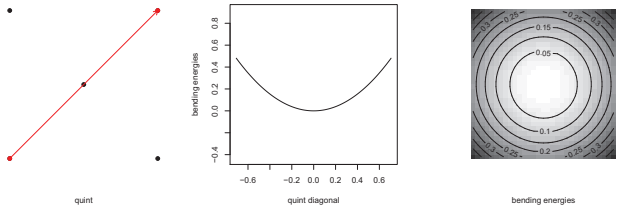
$$\mathbf{L}^{-1} = \begin{pmatrix} \mathbf{L}^{11}_{k\times k} & \mathbf{L}^{12}_{k\times 3} \\ \mathbf{L}^{21}_{3\times k} & \mathbf{L}^{22}_{3\times 3} \end{pmatrix},$$

where

1. **bending energy matrix** equals to $\mathbf{B}_e = \mathbf{L}^{11}_{k\times k}$
2. **bending energy** or **penalty** equals to
   $J(\mathbf{f}) = \sum_{m=1}^d \int \int_{\mathbb{R}^d} \left[\sum_{i,j} \left(\frac{\partial^2 f_m}{\partial x^{(i)} \partial x^{(j)}} \right)^2 \right] dx^{(1)} dx^{(2)} ... dx^{(d)}$,
   with TPS model solution as
   $J(\mathbf{f}) = tr\left(\mathbf{W}^T \mathbf{S} \mathbf{W}\right) = tr(\mathbf{Y}^T \mathbf{B}_e \mathbf{Y})$

Geometric Morphometrics
Bending Energy





Obrázok: **TPS deformation grid**, **bending**, and **bending energy** $J(\mathbf{f})$

Geometric Morphometrics
Affine and non-affine coordinates

Definition (Affine and non-affine coordinates)

Regressing each $k \times d$ matrix $\mathbf{X}_{P,i}$ $(d = 2, 3)$ onto the $\overline{\mathbf{X}}_P$ can be defined by the *MMLRM* (***Multivariate Multiple Linear Regression Model***)

$$\mathbf{X}_{P,i} = \overline{\mathbf{X}}_P \beta_i + \epsilon_i; \widehat{\beta}_i = \left(\overline{\mathbf{X}}_P^T \overline{\mathbf{X}}_P\right)^{-1} \overline{\mathbf{X}}_P^T \mathbf{X}_{P,i}, i = 1, 2, ...n.$$

Let $\widehat{\beta}_i = \left(\widehat{\beta}_{i1} : \widehat{\beta}_{i2}\right)$ for 2D and $\widehat{\beta}_i = \left(\widehat{\beta}_{i1} : \widehat{\beta}_{i2} : \widehat{\beta}_{i3}\right)$ for 3D, then

1. **affine Procrustes coordinates**: $\mathbf{X}_{A,i} = \mathbf{X}_{P,i} \widehat{\beta}_i$
2. **non-affine Procrustes coordinates** (residuals of *MMLRM*): $\mathbf{X}_{NA,i} = \overline{\mathbf{X}}_P + \left(\mathbf{X}_{P,i} - \mathbf{X}_{A,i}\right)$

Relative Warp Analysis
Generalized PCA—from shape space to affine and non-affine subspaces 1

Definition (Relative Warp Analysis (RWA))

If bending energy matrix $\mathbf{B}_e$ is calculated for the mean shape $\overline{\mathbf{X}}_P$, then $dk \times dk$ matrix $\mathbf{B} = \mathbf{I}_{d \times d} \otimes \mathbf{B}_e$. Let ***Generalized covariance matrix with respect to bending energy*** is equal to

$$\mathbf{S}_B^{(\alpha)} = \left(\mathbf{B}^-\right)^{\alpha/2} \mathbf{S} \left(\mathbf{B}^-\right)^{\alpha/2},$$

where $\left(\mathbf{B}^-\right)^{\alpha/2} = \sum_j \widehat{\lambda}_j^{-\alpha/2} \widehat{\gamma}_j^T \widehat{\gamma}_j$ is *Moore-Penrose generalized inverse* of $\mathbf{B}^{\alpha/2}$.The non-zero eigenvalues of $\mathbf{S}_B^{(\alpha)}$ calculated by SVD are $\widehat{l}_j$ and corresponding eigenvectors $\widehat{\mathbf{g}}_j$ (***relative warps, RW***). Then ***RW scores***

$$r_{ij} = \widehat{\mathbf{g}}_j^T \left(\mathbf{B}^-\right)^{\alpha/2} Vec\left(\mathbf{X}_{S,i}\right), i = 1, 2, ...n; j = 1, 2, ...J_d,$$

where J_d is the number of non-zero eigenvalues $(d = 2, 3)$.

Relative Warp Analysis
Generalized PCA—from shape space to affine and non-affine subspaces 2

Definition (Relative Warp Analysis (RWA), cont.)

The effect of the jth *RW* can be viewed by plotting

$$Vec\left(\mathbf{X}_P\left(c, j, \alpha\right)\right) = Vec(\overline{\mathbf{X}}_P) \pm c_j \mathbf{B}^{\alpha/2} \widehat{\mathbf{g}}_j \widehat{l}_j^{1/2}, r_j = c_j \widehat{l}_j^{1/2}$$

for various values of $r_j \in \left\langle 0, \max(|r_{ij}|)\right\rangle$ (or reasonable magnification of $\max(|r_{ij}|)$; alternatively, either $c_j \sim N(0, 1)$ or fixing $c_j = 1$, magnification of $\widehat{l}_j^{1/2}$, standard deviation of RW_j scores), where $\mathbf{B}_e^{\alpha/2} = \sum_j \widehat{\lambda}_j^{\alpha/2} \widehat{\gamma}_j^T$. To emphasize

1. *large scale variability (global bending)*, $\alpha = 1$,
2. *small scale variability (local bending)*, $\alpha = -1$,
3. $\alpha = 0$, then we take $\mathbf{B}^0 = \mathbf{I}$ as the $dk \times dk$ identity matrix and the procedure is equivalent to PCA of Procrustes shape coordinates

Relative Warp Analysis
Generalized PCA—from shape space to affine and non-affine subspaces 3

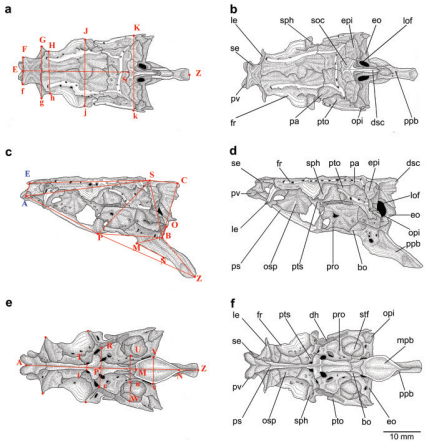
Definition (Relative Warp Analysis (RWA), cont.)

1. ***Affine* contribution** to the variability by performing affine subspace *PCA* on the covariance matrix $\mathbf{S}_A$ of $n \times dk$ matrix $\mathbf{X}_A$ with the rows $Vec\left(\mathbf{X}_{A,i}\right)$, $i = 1, 2, ...n$ (which is equivalent to the *RWA* with $\alpha = 0$)
2. ***Non-affine* contribution** to the variability by performing non-affine subspace *PCA* on the covariance matrix $\mathbf{S}_{NA}$ of $n \times dk$ matrix $\mathbf{X}_{NA}$ with the rows $Vec\left(\mathbf{X}_{NA,i}\right)$, $i = 1, 2, ...n$
3. ***Contribution of (a)symmetry*** by augmenting relabeled and reflected Procrustes configurations to vectorized matrix of Procrustes shape coordinates and performing SVD of $\mathbf{S}_{AS}$
4. ***Size contribution*** by augmenting vectorized matrix of Procrustes shape coordinates by column of **centroid sizes** $\mathbf{x}_{size} = (\ln(CS_1), ..., \ln(CS_n))^T$, where $CS_i = \sqrt{\left(\sum_{j=1}^k \|\mathbf{x}_{ij} - \overline{\mathbf{x}}_i\|_2^2\right)} = \|\mathbf{X}_i\| = tr(\mathbf{X}_i \mathbf{X}_i^T)$, then $n \times (dk + 1)$ matrix of **vectorized form coordinates** $\mathbf{X}_F = (\mathbf{X}_S \dot{:} \mathbf{x}_{size})$, and finally performing SVD of $\mathbf{S}_F$

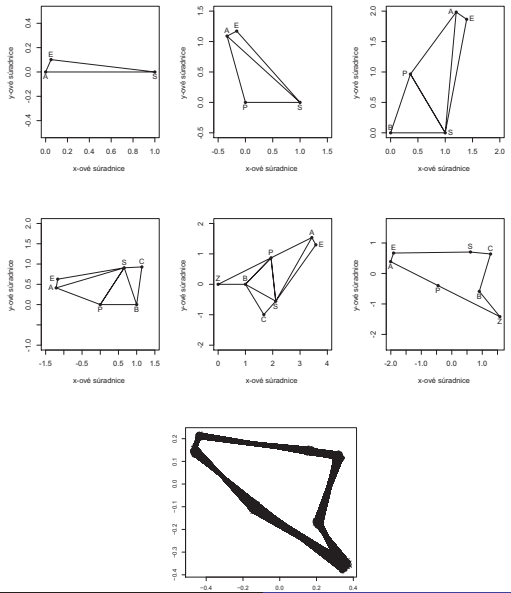
GM vs KM
GM neurokránia rýb z rodu *belica*

- neurocrania–**roaches** *Rutilus rutilus* and *Rutilus virgo* (*Actinopterygii: Cyprinidae*)
- *R. rutilus* ($n_{rr} = 30$) and *R. pigus* neurocrania ($n_{rp} = 50$), 27 measurements

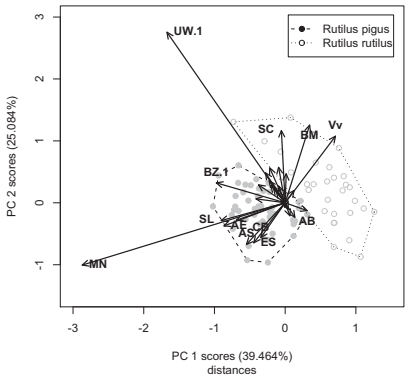
GM vs KM
GM neurokránia rýb z rodu *belica*



GM vs KM
GM neurokránia rýb z rodu *belica*



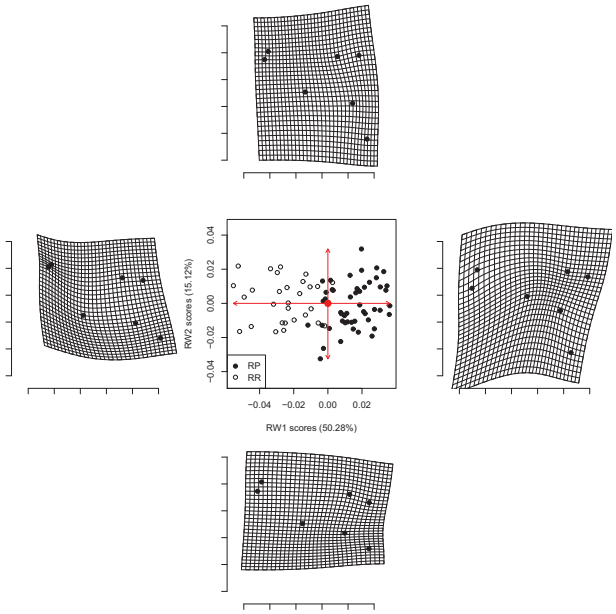
Traditional vs Geometric Morphometrics
Fish Neurocrania–-*Rutilus rutilus* and *R.pigus* (Cyprinidae)



Obrázok: PCA of inter-landmark distances

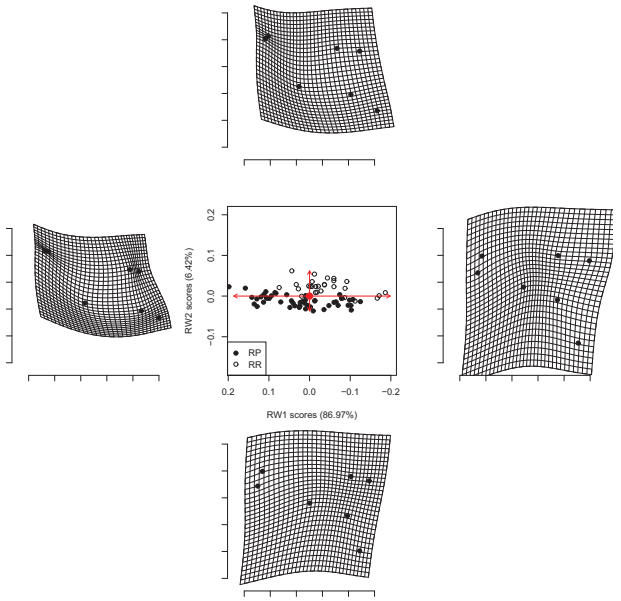
Traditional vs Geometric Morphometrics

Fish Neurocrania–-*Rutilus rutilus* and *R.pigus* (Cyprinidae)—Shape Space PCA



Traditional vs Geometric Morphometrics

Fish Neurocrania–-*Rutilus rutilus* and *R.pigus* (Cyprinidae)—Form Space PCA



Relative Warp Analysis

Generalized PCA—Generelazed PCA for paired data

Definition (RWA for paired data)

1. Let $\mathbf{x}_{P,i} = Vec\left(\mathbf{X}_{P,i}\right), i = 1, 2, ..., n; n = 48$ be a $2k$-vector of ***Procrustes shape coordinates***, where $\mathbf{X}_{P,i} = (\mathbf{x}_{P,i}^{(1)} \vdots \mathbf{x}_{P,i}^{(2)})$, $\mathbf{x}_{P,i}^{(d)} = \left(x_{i1}^{(d)}, x_{i2}^{(d)}, ..., x_{ik}^{(d)}\right), d = 1, 2$
2. Let $\mathbf{x}_{D,i}$ be $2k$-vectors ($k = 22$) of ***matched-pair differences of vectorized Procrustes shape coordinates***, $\mathbf{x}_{D,i} = \mathbf{x}_{P,15,i} - \mathbf{x}_{P,10,i}$, $\mathbf{x}_{P,15,i} = Vec(\mathbf{X}_{P,15,i})$ and $\mathbf{x}_{P,10,i} = Vec(\mathbf{X}_{P,10,i})$
3. $\mathbf{S}_D$ be the ***covariance matrix*** of the data $\mathbf{x}_{D,i}$,
4. $\overline{\mathbf{X}}_{P,10} = (\overline{\mathbf{x}}_{P,10}^{(1)} \vdots \overline{\mathbf{x}}_{P,10}^{(2)}) = (\overline{\mathbf{x}}_1, ..., \overline{\mathbf{x}}_k)^T$ be $k \times 2$ matrix of mean Procrustes shape coordinates $\overline{\mathbf{x}}_j$ of 10-year group, $j = 1, 2, ..., k$, then
5. $$\mathbf{L} = \begin{pmatrix} \mathbf{S} & \mathbf{1}_k & \overline{\mathbf{X}}_{P,10} \\ \mathbf{1}_k^T & 0 & \mathbf{0} \\ \overline{\mathbf{X}}_{P,10}^T & \mathbf{0} & \mathbf{0} \end{pmatrix}, \mathbf{L}^{-1} = \begin{pmatrix} \mathbf{L}_{k\times k}^{11} & \mathbf{L}_{k\times 3}^{12} \\ \mathbf{L}_{3\times k}^{21} & \mathbf{L}_{3\times 3}^{22} \end{pmatrix},$$

Relative Warp Analysis

Generalized PCA—Generelazed PCA for paired data

Definition (RWA for paired data, cont.)

- where $\mathbf{L}$ is symmetric positive definite,
- the inverse of $\mathbf{S}$ exists as long as the landmarks are at least four in number, not all on one straight line, and also not in the same place (coincident); then inverse of $\mathbf{L}$ exists and is equal to $\mathbf{L}^{-1}$
- $\mathbf{S}_{js} = \phi\left(\overline{\mathbf{x}}_j - \overline{\mathbf{x}}_s\right); j, s = 1, 2, ..., k$, $\phi(\mathbf{x}) = \|\mathbf{x}\|_2^2 \log\left(\|\mathbf{x}\|_2^2\right), \forall \|\mathbf{x}\|_2 > 0$, if $\|\mathbf{x}\|_2 = 0$, $\phi(\mathbf{x}) = 0$
- $k \times k$ matrix $\mathbf{B}_e = \mathbf{L}^{11}$ is called *bending energy matrix* of $\overline{\mathbf{X}}_{P,10}$, $2k \times 2k$ matrix $\mathbf{B} = \mathbf{I}_{2\times 2} \otimes \mathbf{B}_e$, and $\mathbf{1}_k^T \mathbf{B}_e = \mathbf{0}$, $\mathbf{X}^T \mathbf{B}_e = \mathbf{0}$, so the rank of the bending energy matrix is $k - 3$
- then $\left(\mathbf{B}^-\right)^{\alpha/2} \mathbf{S}_D \left(\mathbf{B}^-\right)^{\alpha/2}$ is ***generalized covariance matrix of matched-pair differences of vectorized Procrustes shape coordinates***, $\mathbf{x}_{D,i}$
- non-zero ***eigenvalues*** are $\widehat{l}_j$ with corresponding ***eigenvectors*** $\widehat{\mathbf{g}}_j$ (***PC loadings***, RWs)

Definition (RWA for paired data, cont.)

- ***RW scores*** are defined as $r_{ij} = \widehat{\mathbf{g}}_j^T \left(\mathbf{B}^-\right)^{\alpha/2} \mathbf{x}_{D,i}$
- ***the effect of the jth RW*** can be viewed by plotting

$$Vec\left(\mathbf{X}_P\left(c_j, j, \alpha\right)\right) = Vec(\overline{\mathbf{X}}_{P,10}) \pm c_j \mathbf{B}^{\alpha/2} \widehat{\mathbf{g}}_j \widehat{l}_j^{1/2}, r_{ij} = c_j \widehat{l}_j^{1/2}, c_j \in \mathbb{R}^+$$

for various values of $r_{ij} \in \langle 0, \max(|r_{ij}|)\rangle$ (or some magnification of $\max(|r_{ij}|)$; alternatively, fixing $c_j = 1$, magnification of $\widehat{l}_j^{1/2}$ as standard deviation of PC_j scores)

- ***the effect of the linear combination of RW_1 and RW_2*** can be viewed by plotting

$$Vec\left(\mathbf{X}_P\left(c_1, c_2, \alpha\right)\right) = Vec(\overline{\mathbf{X}}_{P,10}) \pm c_1 \mathbf{B}^{\alpha/2} \widehat{\mathbf{g}}_1 \widehat{l}_1^{1/2} \pm c_2 \mathbf{B}^{\alpha/2} \widehat{\mathbf{g}}_2 \widehat{l}_2^{1/2}$$

- a ***PC summary*** of the shape data

$$Vec\left(\mathbf{X}_{P,15,i}\left(\alpha\right)\right) = Vec(\overline{\mathbf{X}}_{P,10}) \pm \mathbf{B}_e^{\alpha/2} \sum_{j=1}^{2} r_{ij} \widehat{\mathbf{g}}_j = Vec(\overline{\mathbf{X}}_{P,10}) + \sum_{j=1}^{2} \widehat{\mathbf{g}}_j \widehat{\mathbf{g}}_j^T \mathbf{x}_{D,i}$$

Definition (RWA for paired data, cont.)

- to find the ***affine component*** we use ***linear regression model***

$$\overline{\mathbf{x}}_{P,10}^{(d)} + \mathbf{x}_{D,i}^{(d)} = \overline{\mathbf{x}}_{P,10}^{(d)} \beta_i^{(d)} + \epsilon_i^{(d)}, d = 1, 2; i = 1, 2, ..., n,$$

- then $\mathbf{x}_{A,i}^{(d)} = \overline{\mathbf{x}}_{P,10}^{(d)} \widehat{\beta}_i^{(d)}$ and $\mathbf{X}_{A,i} = (\mathbf{x}_{A,i}^{(1)} \dot{:} \mathbf{x}_{A,i}^{(2)})$ is the affine component of $\mathbf{X}_{P,i}$
- in ***affine subspace***, $\mathbf{S}_{DA}$ stands for sample covariance matrix of $\mathbf{x}_{DA,i} = Vec(\mathbf{X}_{A,i}) - Vec(\overline{\mathbf{X}}_{P,10})$; then PCA of $\mathbf{S}_{DA}$ is called ***affine-subspace PCA***
- let $\mathbf{X}_{DF} = (\mathbf{X}_D \dot{:} \mathbf{x}_{size})$, be an $n \times (2k+1)$ matrix with the rows equal to $\mathbf{x}_{DF,i} = \left(\mathbf{x}_{D,i}^T, \ln(CS_i)\right)^T, i = 1, 2, ..., n$, and $\mathbf{x}_{size}^T = (\ln(CS_1), ..., \ln(CS_n))$. Let $\mathbf{S}_{DF}$ be the covariance matrix of the data $\mathbf{x}_{DF,i}$; then PCA of $\mathbf{S}_{DF}$ is called ***form-space PCA***
- the first PC represents ***allometry***—shape change during growth

Definition (RWA for paired data, cont.)

Visualization of interpolated shape changes can be done

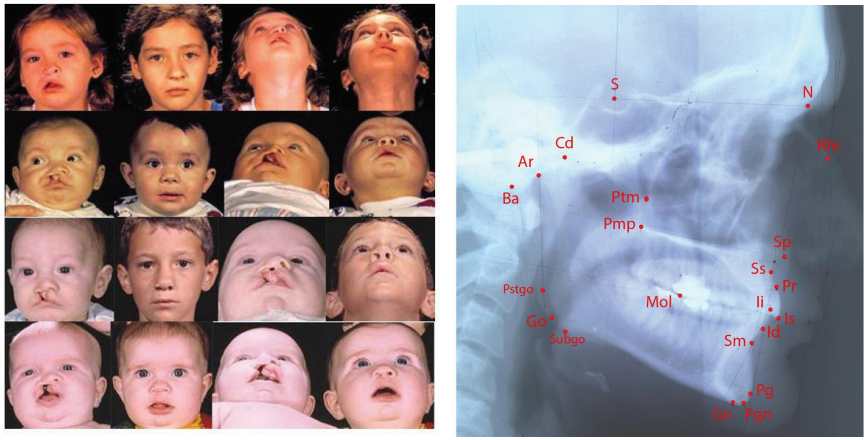
- via ***thin-plate spline (TPS) deformation grids***,
- ***field of vectors*** (within the convex hull of reference shape $\overline{\mathbf{X}}_{P,10}$, where longer vectors show stronger deformation in the specific direction of the shape change) ***superimposed with the grid of gray-scale rectangles with colors corresponding to the Procrustes distances*** (regions showing milder deformation are lighter, regions with stronger deformation are darker; the surface does not show the direction—but only the size—of some shape change)

- Velemínská J., **Katina, S.**, Šmahel, Z., Sedláčková, M., 2006: Analysis of facial skeleton shape in patients with complete unilateral cleft lip and palate: Geometric morphometrics. *Acta Chirurgiae Plasticae*, **48,1**: 26–32
- Velemínská J., Šmahel, Z., **Katina, S.**, 2006: Development prediction of sagittal intermaxillary relations in patients with complete unilateral cleft lip and palate during puberty. *Acta Chirurgiae Plasticae*, **49,2**: 41–46
- **Katina, S.**,2008: Detection of shape outliers with an application to complete unilateral cleft lip and palat in humans. In S. Barber, P.D. Baxter, A. Gusnanto & K.V.Mardia (eds), The Art & Science of Statistical Bioinformatics, pp. 33-37. Leeds, Leeds University Press
- **Katina, S.**,2011: Detection of shape outliers for matched-pair shape data. *Tatra Mountains Mathematical Publication* (accepted)

- 48 boys, complete unilateral cleft of lip and palate (UCLP), without symptoms of other associated malformations, Clinic of Plastic Surgery in Prague
- homogenously operated by the same team of surgeons (cheiloplasty according to Tennison, periosteoplasty without the nasal septum repositioning
- patients monitored during puberty, at the ages of 10 and 15 (born between 1972 and 1978)
- 22 landmarks (x-rays of the patients' heads, under standard conditions, SigmaScan Pro 5 software)

Geometric Morphometrics
2D lateral X-rays—growth after surgery (paired data)

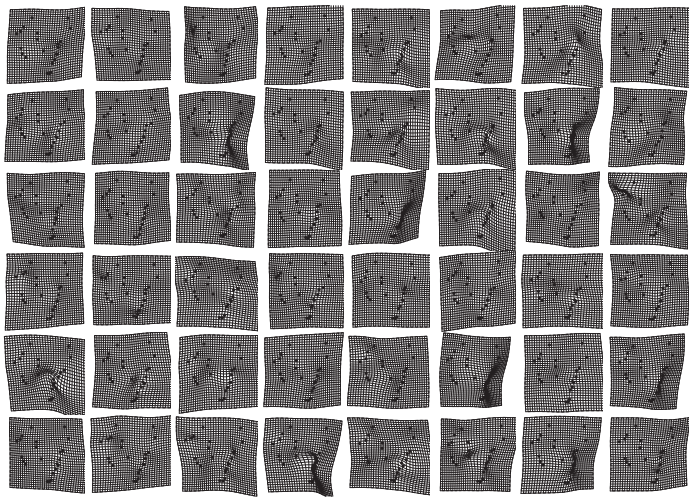
48 boys, 10 – 15yrs old, 22 landmarks



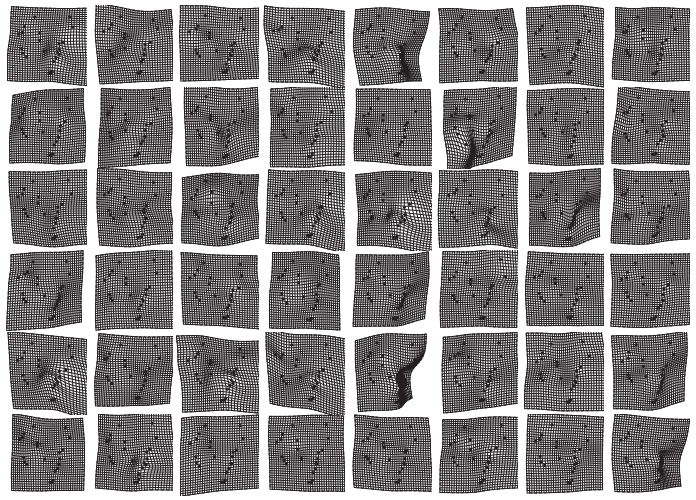
Obrázok: Cleft patients and Design of lateral X-ray (semi)landmarks

[Dpt. of Anthropology, Charles University, Prague, Czech Republic]

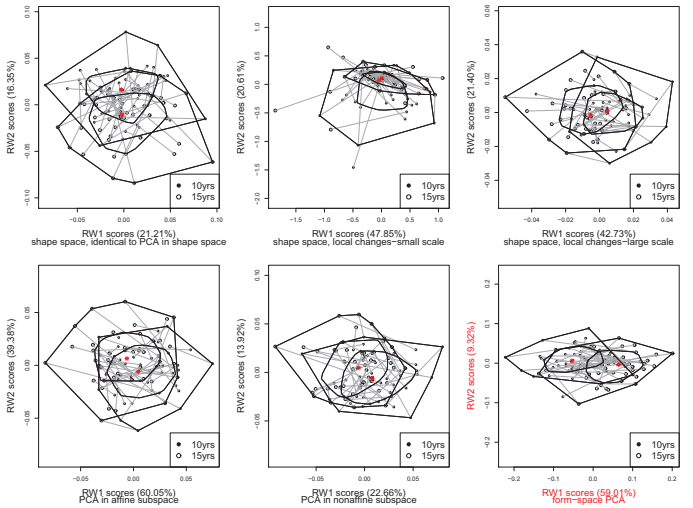
Data—10yrs old boys before operation
2D lateral X-rays—growth after surgery (paired data)



Data—15yrs old boys after operation
2D lateral X-rays—growth after surgery (paired data)

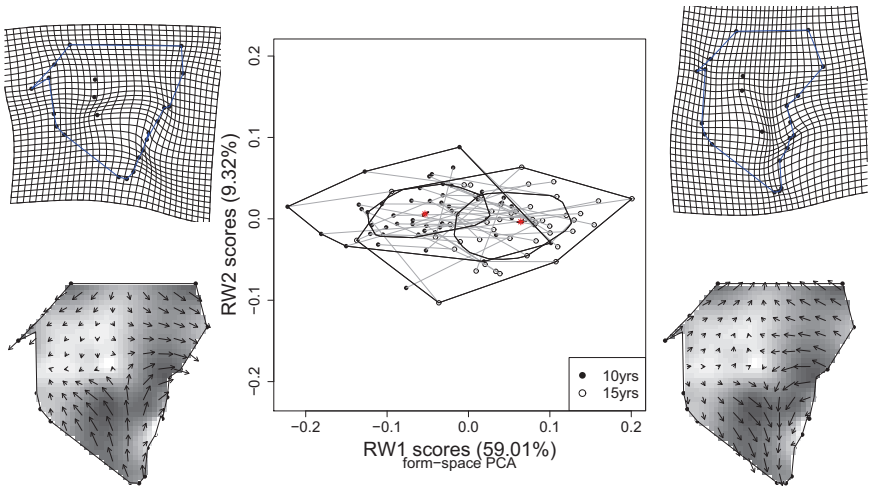


Geometric Morphometrics
2D lateral X-rays—searching biological signal in the data

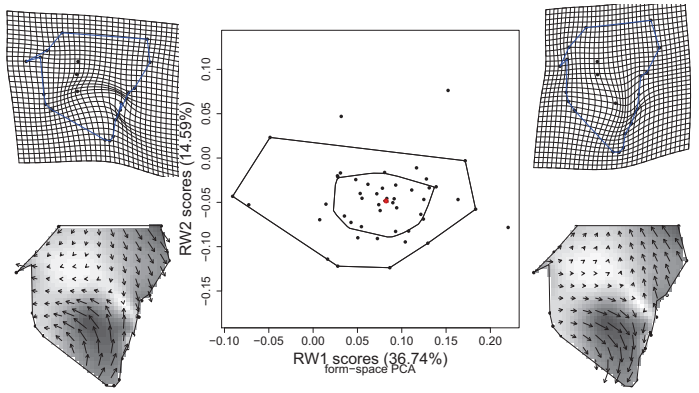


Obrázok: All PCA/RWA models—RW_1,RW_2 subspace

Results of RWA—form space

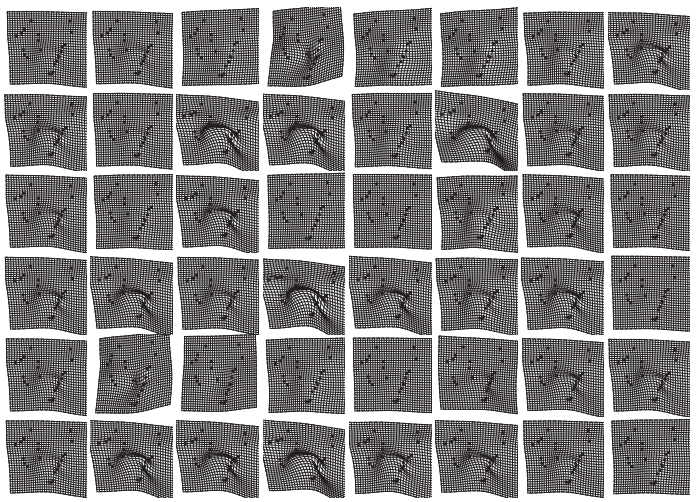


Results of RWA—form space

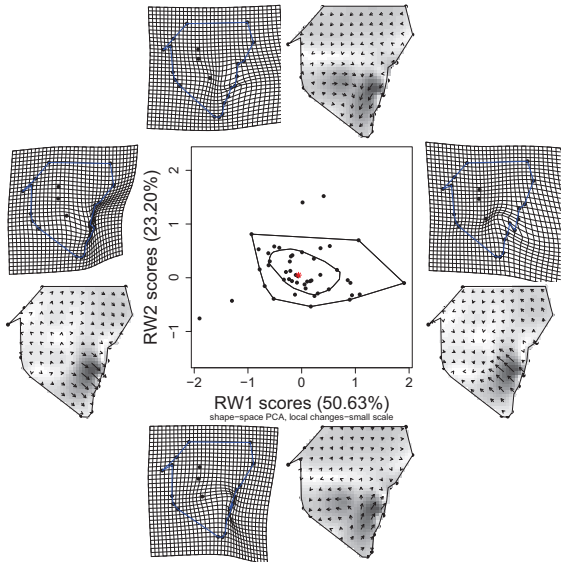


Obrázok: ***Form-space*** space PCA—RWA of $\mathbf{S}_F$ (RW_1,RW_2 subspace)

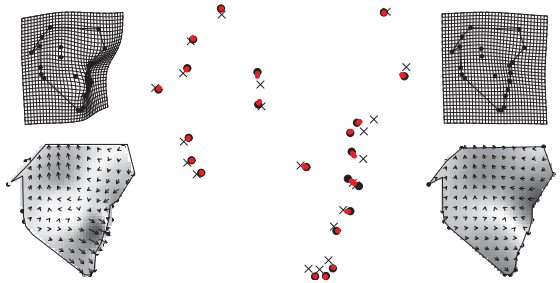
Results of RWA—form space



Results of RWA—bending patterns in small scale



Outlier relaxation using PRM3



Obrázok: *Relaxation in Procrustes shape coordinates*; *TPS deformation grids* and *field of vectors superimposed with the surface of Procrustes distances* of mean shape $\overline{\mathbf{X}}_{P,10}$ to the shape $\mathbf{X}_{P,10,29}$ (left) and to the final relaxed shape $\mathbf{X}_{P,10,29}$ (right); *'curve décolletage'* of the shape $\mathbf{X}_{P,10,29}$ ($\times$–mean shape $\overline{\mathbf{X}}_{P,10}$, big $\bullet$–shape $\mathbf{X}_{P,10,29}$, small $\bullet$–relaxed shapes $\mathbf{X}_{P,10,29}$; middle)

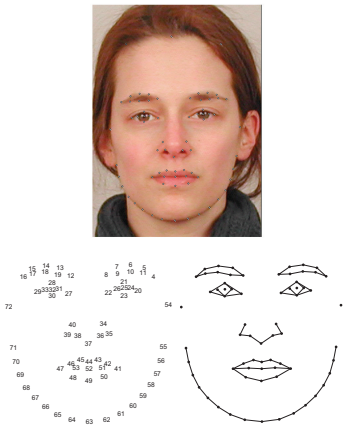
Data—human faces in 2D

- Oberzaucher, E., **Katina, S.**, Holzleitner, I.J., Schmehl, S.F., Mehu-Blantar, I., Grammer, K., 2011: The myth of hidden ovulation: Shape and texture changes in the face during the menstrual cycle. *PNAS* (submitted)
- Pflüger, L.S., Oberzaucher, E., **Katina, S.**, Holzleitner, I.J., Mehu-Blantar The Signal of Fertility. Evidence from a Rural Sample. *Evolution and Human Behaviour* (accepted)
- 20 young women (aged between 19 and 31) who reported to have a regular menstrual cycle and did not take any hormonal contraceptives
- standardized facial photographs–one taken in the **ovulatory** and one in the **luteal phase**
- in a **forced choice task**, 50 male and 50 female subjects were presented with these photographs of each participant–to pick out the **more attractive, healthy, sexy, and likeable**, of the two
- skin patches sized 150 $\times$ 150 pixels from the **cheek** and subjected them to the same forced choice task with slightly modified adjectives
- 46 **landmarks** and 26 **semilandmarks**

Data—human faces in 2D

2D Facial Analysis—two group differences

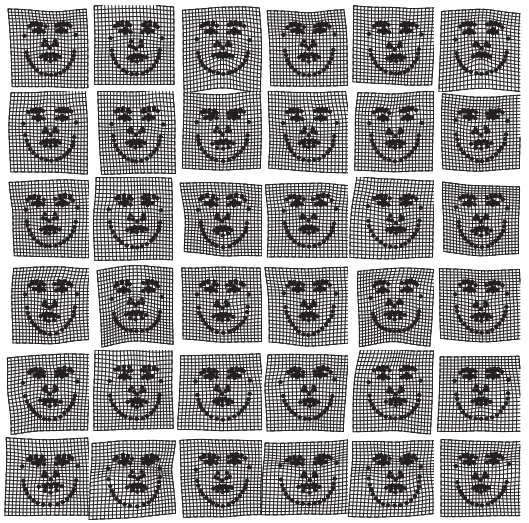
20 young women, $19-31$yrs old, $46+26$ (semi)landmarks



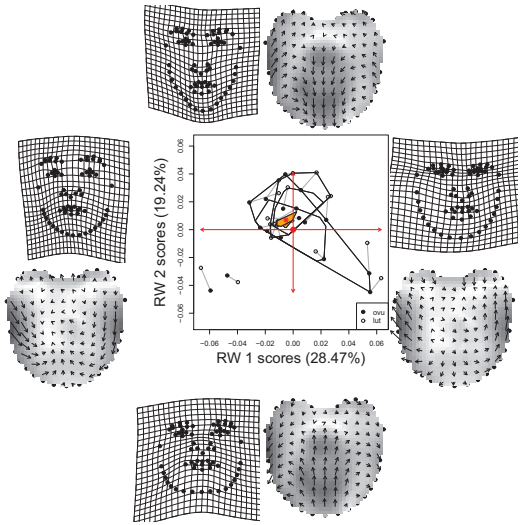
Obrázok: Design of facial (semi)landmarks

[Dpt. of Anthropology, University of Vienna, Vienna, Austria]

Data—human faces in 2D

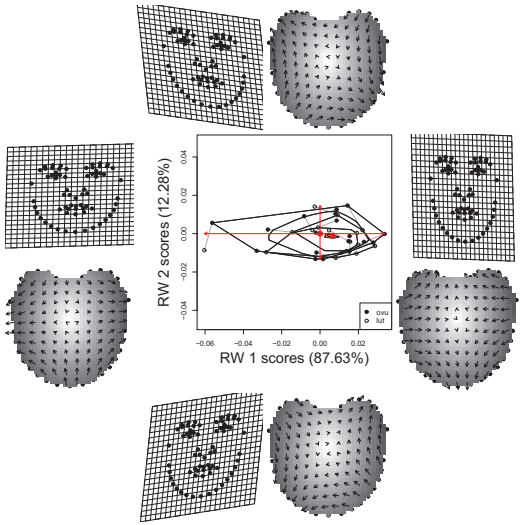


Geometric Morphometrics
2D Facial Analysis—searching for biological signal in the data



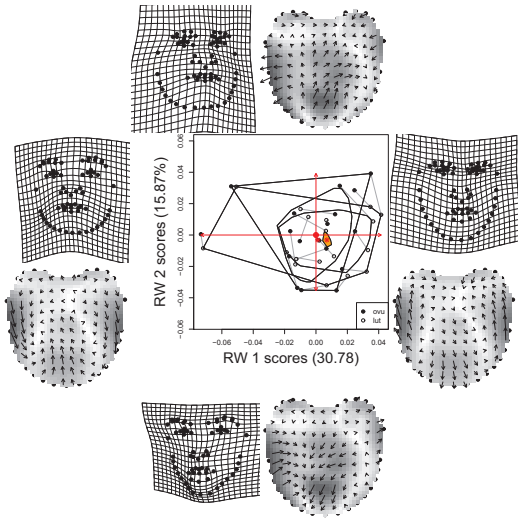
Obrázok: ***Shape*** space PCA—RWA of **S** (RW_1,RW_2 subspace)

Geometric Morphometrics
2D Facial Analysis—searching for biological signal in the data



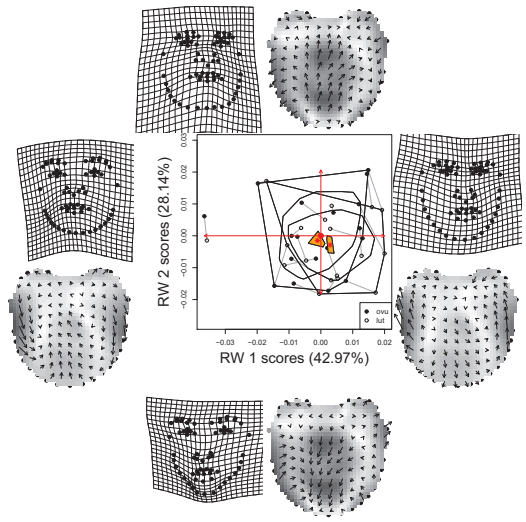
Obrázok: ***Affine*** subspace PCA—RWA of $\mathbf{S}_A$ (RW_1,RW_2 subspace)

Geometric Morphometrics
2D Facial Analysis—searching for biological signal in the data



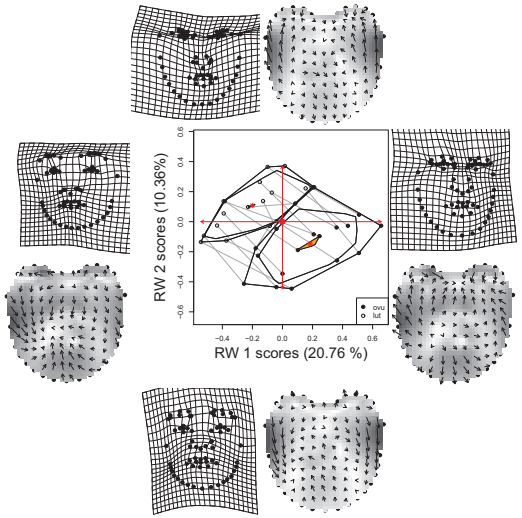
Obrázok: ***Nonaffine*** space PCA—RWA of $\mathbf{S}_{AN}$ (RW_1,RW_2 subspace)

Geometric Morphometrics
2D Facial Analysis—searching for biological signal in the data



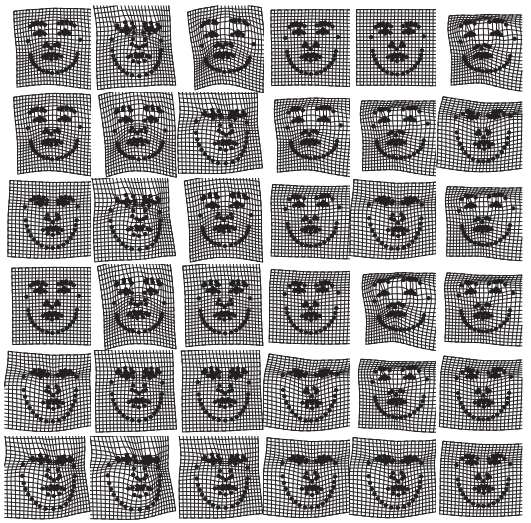
Obrázok: ***Nonaffine*** space PCA—RWA of $\mathbf{S}_B^{(1)}$ (RW_1,RW_2 subspace)

Geometric Morphometrics
2D Facial Analysis—searching biological signal in the data

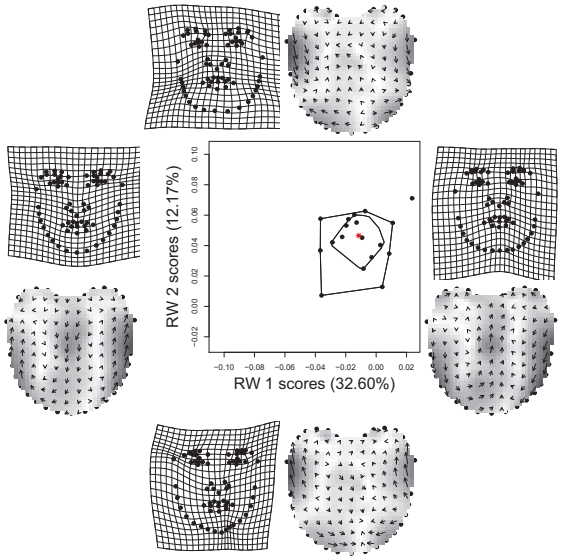


Obrázok: ***Nonaffine*** space PCA—RWA of $\mathbf{S}_B^{(-1)}$ (RW_1,RW_2 subspace)

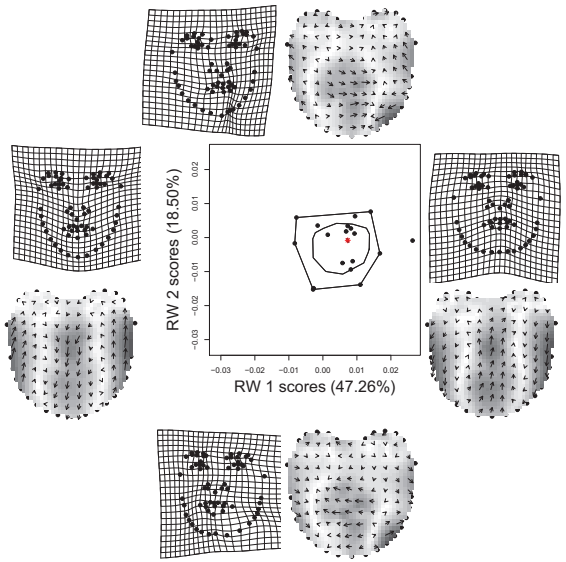
Results of RWA—estimated shapes, RW1 and RW2



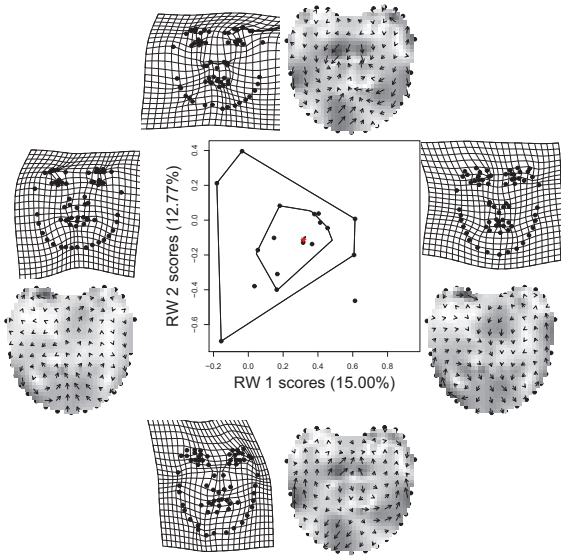
Results of RWA—shape space



Results of RWA—bending patterns with large scale

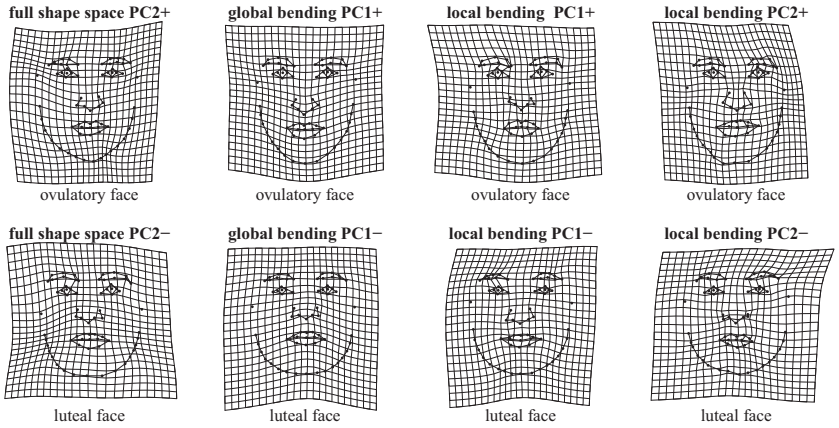


Results of RWA—bending patterns with small scale



Geometric Morphometrics

2D Facial Analysis—searching biological signal in the data



Obrázok: Summary of RWA/PCA analyses in all subspaces of ***paired shape differences*** [statistically significant RWs/PCs]

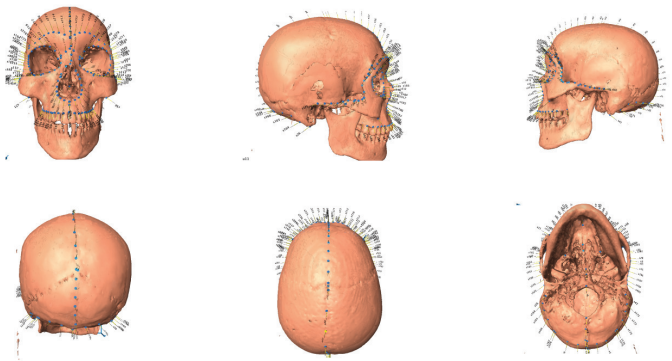
Data—human skulls

3D (semi)landmarks

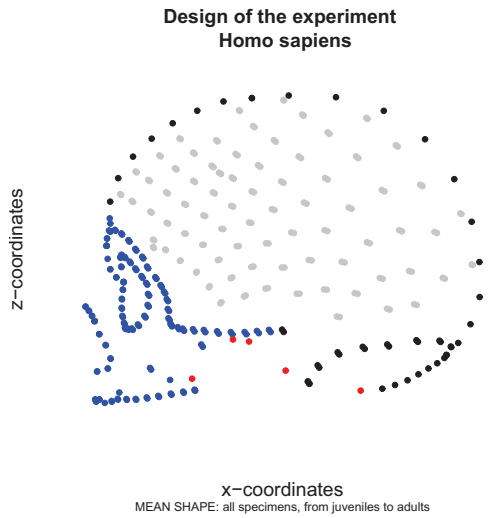
- example re-uses *part of a Vienna data set of* 372 *skulls from various collections*
- 106 human crania (38 adult females, 54 males, 3 juvenile females, 11 juvenile males, 14 unknown sex; from newborns to adults)
- Dept. of Archaeological Biology and Anthropology, Natural History Museum, Vienna, Austria
- Dept. of Anthropology, University of Vienna, Vienna, Austria
- *Weisbach collection* - acquired and exhumed skeletons of soldiers of the Austro-Hungarian monarchy, sex and age of these crania are known from military records
- *Hallstatt collection* from ossuary in Hallstatt, sex and age are known from the church-books
- data – 347 landmarks and semilandmarks – 32 landmark points, 7 ridge curves totalling 161 semilandmarks and 154 surface semilandmarks [5 – base, 184 – face, 158 – neurocranium]
- landmark points on **both sides** of every cranium and semilandmarks (on curves and surface) **on the left side** of every cranium were digitalized using a MicroScribe 3*DX* (Mitteroecker et al, 2004, Gunz, 2005)
- **Katina, S.**, Bookstein, FL., Gunz, P., Schaefer, K., 2007: Was it worth digitizing all those curves? A worked example from craniofacial primatology. *American Journal of Physical Anthropology* Suppl. **44**: 140.

Data—human skulls

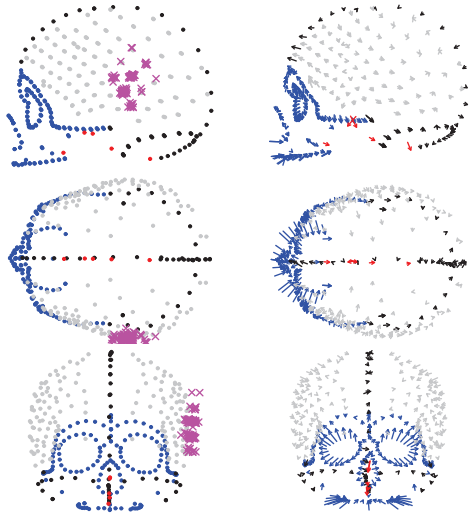
6 norms: norma frontalis, lateralis dex. a sin., occipitalis, verticalis, basilaris



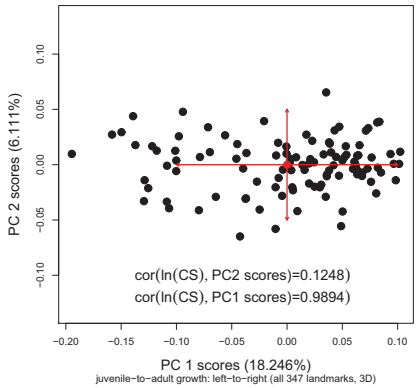
Data—human skulls
(Semi)landmarks of three skull regions



Data—human skulls
(Semi)landmarks of three skull regions and *euryon* variability

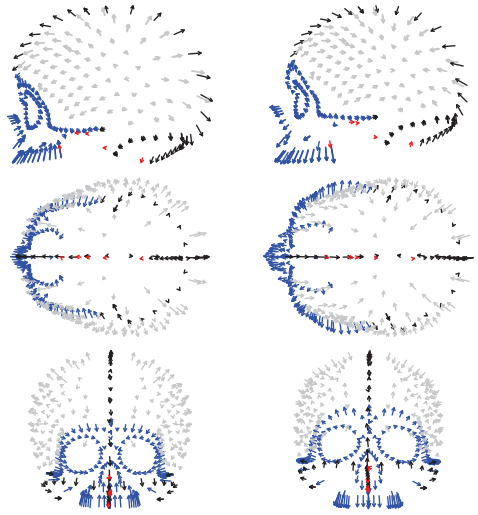


3D Form Space PCA



Obrázok: PC1 and PC2 scores

3D Form Space PCA

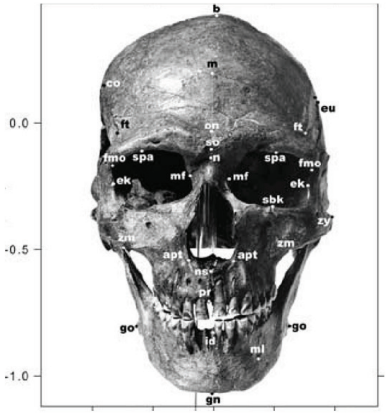


Obrázok: PC1 and PC2

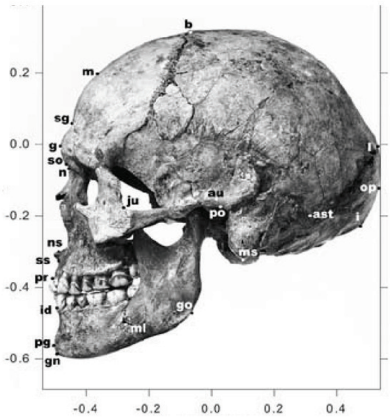
2D Skulls
Předmostí skulls

- **professionally digitised glass plate negatives of fossil skulls** (Předmostí 1 – P1, Předmostí 3 – P3, Předmostí 4 – P4, Předmostí 9 – P9, Předmostí 10 – P10)
- in the accessible norms: frontal, lateral sin., occipital, basal, and vertical views
- the skulls in question are those determined by Matiegka to have been females (P1, P4, P10) and males (P3, P9)
- 17 **landmarks in the right lateral view**
- the recent population collection — ***103 skulls of known sex (51 males and 52 females) and age from the first third of the 20th century***
- **Katina, S.**, Šefčáková, A., Velemínská, J., Bružek, J., Velemínský, P., 2004: A Geometric approach to cranial sexual dimorphism in the upper palaeolithic skulls from Předmostí (Upper Palaeolithic, Czech Republic). *Journal of the National Museum, Natural History Series* 173, 1–4:133–144
- Šefčáková, A., **Katina, S**., 2008: Geometrical analysis of adult skulls from Předmostí, In: Veleminská, J, Bružek, J, (eds), Fossil hominids from Předmostí nr. Přerov : Old documentation and new reading. Academia, Praha, 87 – 101

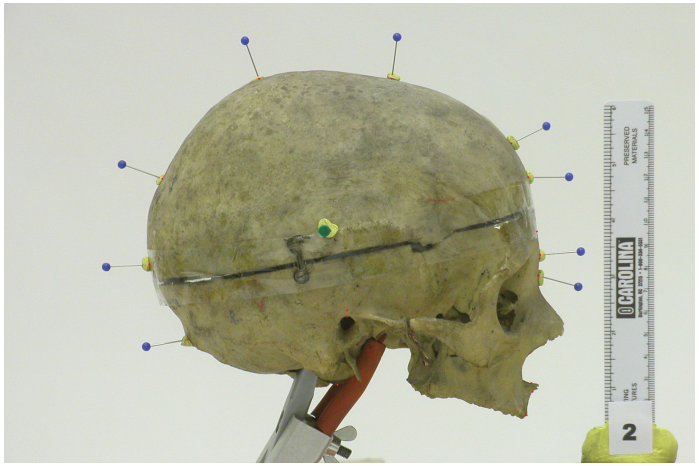
2D Skulls
Norma frontalis



Skulls
Norma lateralis



2D Skulls
Example of skull from Pachner reference sample



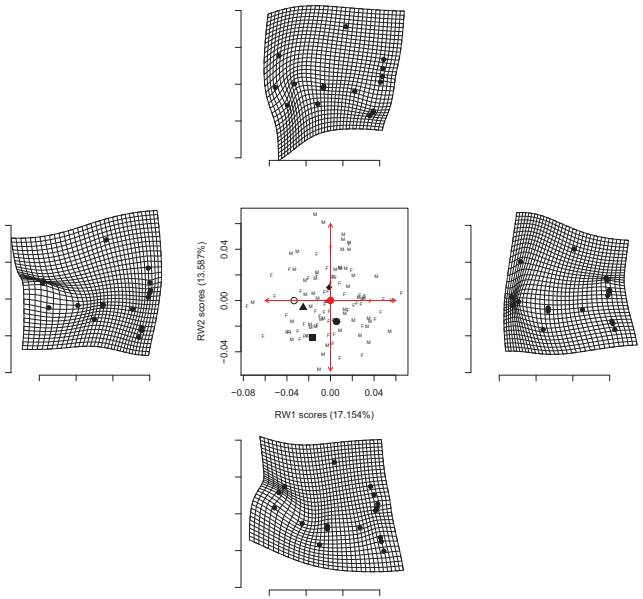
[Pachner collection at the Department of Anthropology and Human Genetics of Charles University in Prague (Czech Republic)]

2D Skulls
PCA Summary

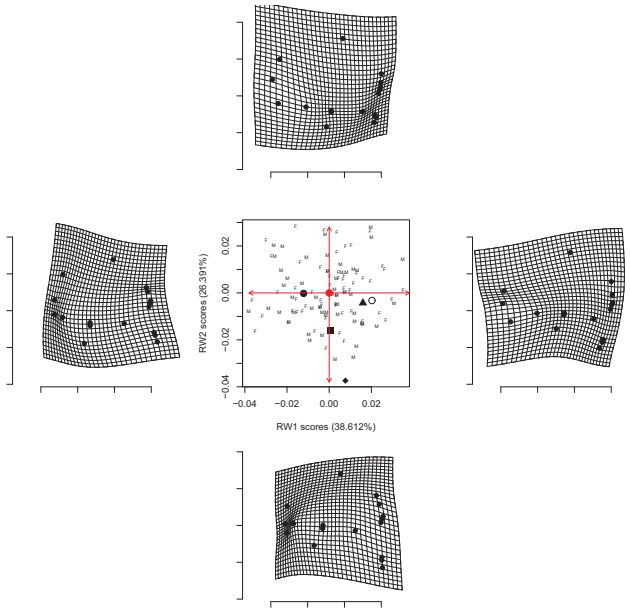
Legenda:
F–Pachner females (n=52), M–Pachner males (n=51),
Předmostí crania–P1 ○, P3 ■, P4 ●, P9 ▲, P10 ♦

1. TPS deformation grids and *RW* scores (*RW*1 and *RW*2) – in ***shape space*** (identical to *PCA* in shape space)
2. TPS deformation grids and *RW* scores (*RW*1 and *RW*2) – in ***shape space for local changes with large scale*** ($\alpha = 1$)
3. TPS deformation grids and PC scores (PC1 and PC2) – in ***form space***
4. TPS deformation grids and PC scores (PC1 and PC2) – in ***form space with*** 95% ***tolerance ellipses for males and females***
5. TPS deformation grids and *RW* scores (*RW*1 and *RW*2) – in ***shape space for local changes with small scale*** ($\alpha = -1$)

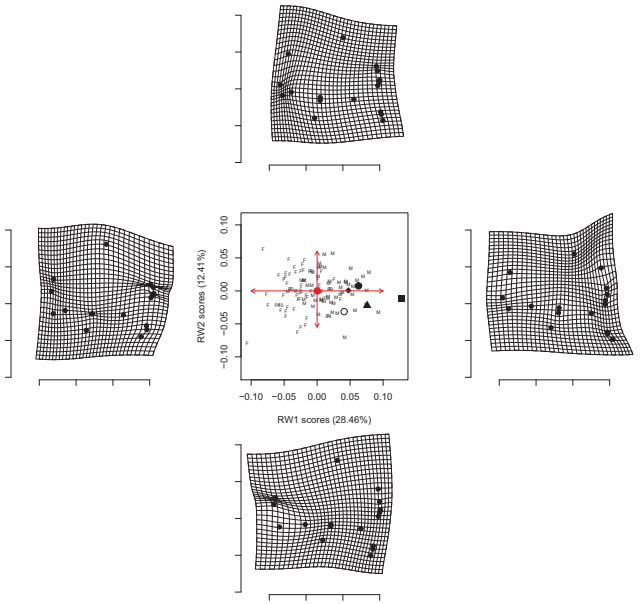
RWA in Paleoanthropology—Shape Space



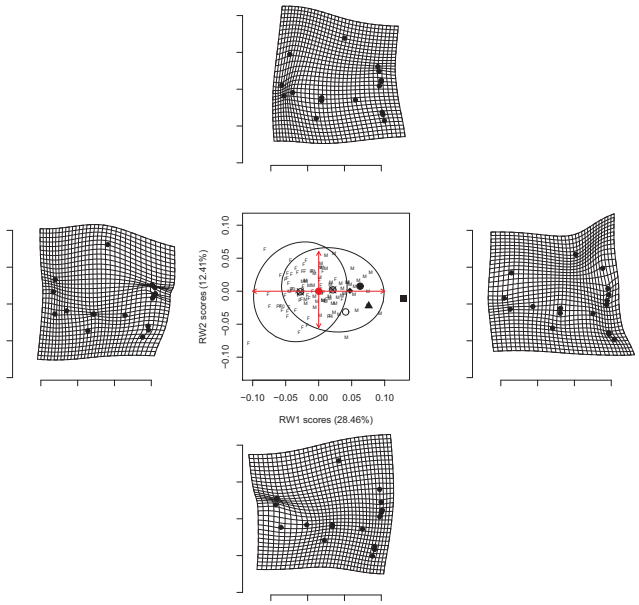
RWA in Paleoanthropology–Global Bending Patterns



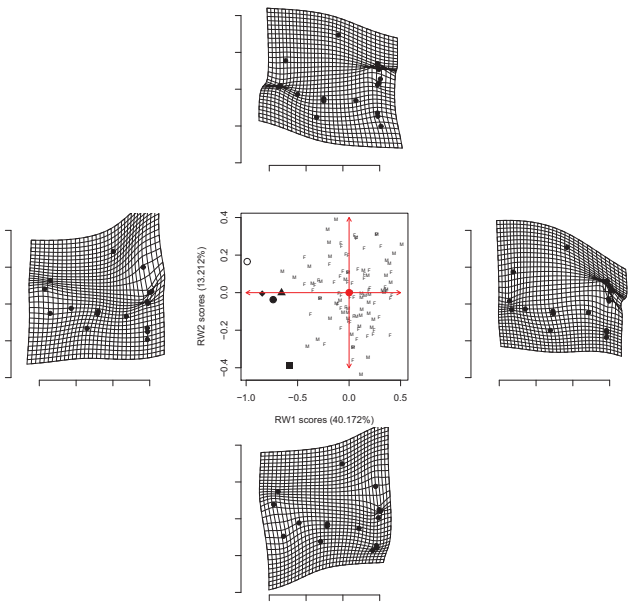
RWA in Paleoanthropology—Form Space



RWA in Paleoanthropology—Form Space



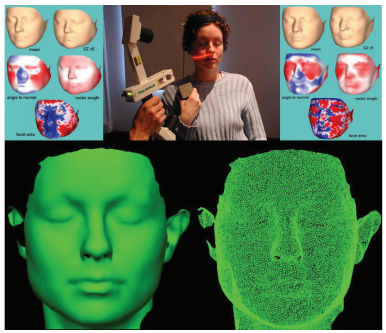
RWA in Paleoanthropology–Local Bending Patterns



3D laser-scan capture

3D facial shape—VCFS data, differences between cases and controls (paired data)

42 pairs of laser-scanned faces, 23 landmarks, 1664 geometrically homologous semilandmarks on curves and surfaces, 59242 mesh-points triangulated with 117386 faces

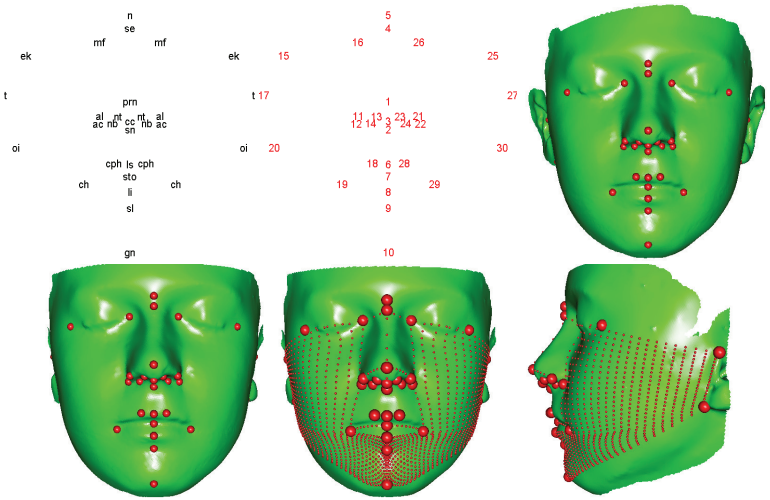


Obrázok: VCFS face, laser-scan, and surface meshes

[Royal College of Surgeons in Ireland, Dublin; Face 3D data]

Geometric Morphometrics

3D facial shape—VCFS data, differences between cases and controls (paired data)



Obrázok: Design of facial (semi)landmarks—***symmetrized mean shape***

3D face—first steps of the analysis
Data analysis 1

Data analysis:

- with respect to the analysis of *object asymmetry* (in our case, facial shape asymmetry), the original coordinates were relabelled and reflected (RR) with respect to *midsagital plane* (MP)
- MP was estimated as an ordinary least square plane of unpaired midsagital landmarks and rotated into (x, y)-plane
- for paired (semi)landmarks, the sign and labels were reversed across the left-hand and right-hand side of the head shape
- the original PSC together with their RR counterparts were jointly submitted to GPA to register both into the same shape space
- both configurations were centered with respect to original and RR Procrustes mean shape, respectively, resulting in original and RR centered PSC
- *fluctuating asymmetry* (FA) expresses how the difference between the original and RR shapes fluctuate in the sample; it is calculated as the sum of squares of individual *asymmetry scores*, i.e. Procrustes distances between original and RR centered PSC of each shape

3D face—first steps of the analysis
Data analysis 2

Data analysis:

- *the asymmetry of the means* (AM) is calculated as the sum of squares of the Procrustes distances between the original and RR Procrustes mean shape; AM multiplied by sample size is called *directional asymmetry* (DA)
- the PSC were adjusted for age and sex by **linear regression model** of the form

  centered $\text{PSC}_{ij} = \text{sex} + \text{age} + \text{sex} : \text{age} + \epsilon_{ij}, i = 1, 2, \ldots 1664; j = 1, 2, 3;$

  for further analysis, residuals of this model were used
- the **direction of case-control difference** was found based on the *projection of "null shape"* to particular PC subspaces; if this fails to negative side of the PC axes cases are on the negative part of the axis as well; if this fails to positive side of the PC axes cases are on the positive part of the axis as well

3D face—first steps of the analysis
Standardized views

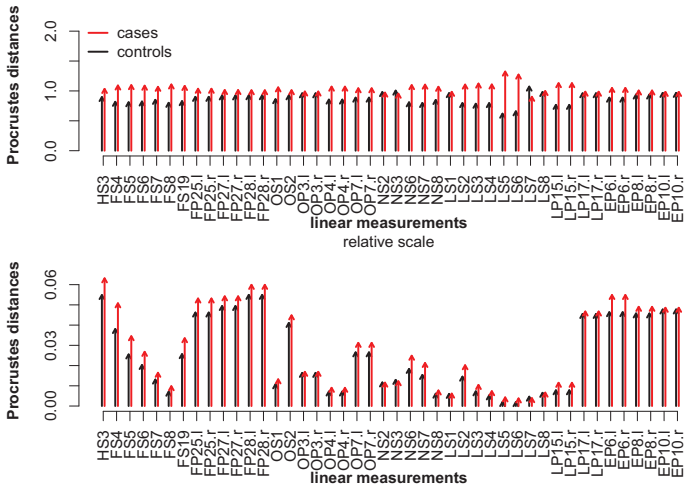
3D face—first steps of the analysis
Data analysis 3

PCA for reversible 3D images:

- 21 RR centered case-control (semi)landmark differences at the same time
- PC scores for original and RR data are equal in absolute values
- in this setting, symmetric and asymmetric PCs are separated which simplifies the interpretation
- the *symmetric PCs* are these where ***PC scores of original and RR data do not have the same sign (they are equal only in absolute value)***
- the *asymmetric PCs* are these where ***PC scores of original and RR data have the same sign (they are equal)***

Geometric Morphometrics

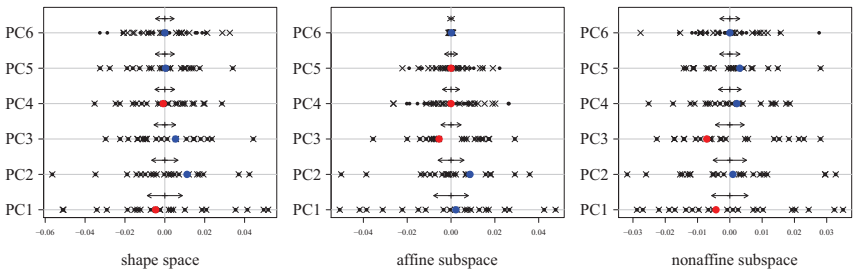
3D facial shape—VCFS data, differences between cases and controls (paired data)



Obrázok: Procrustes shape distances

Geometric Morphometrics

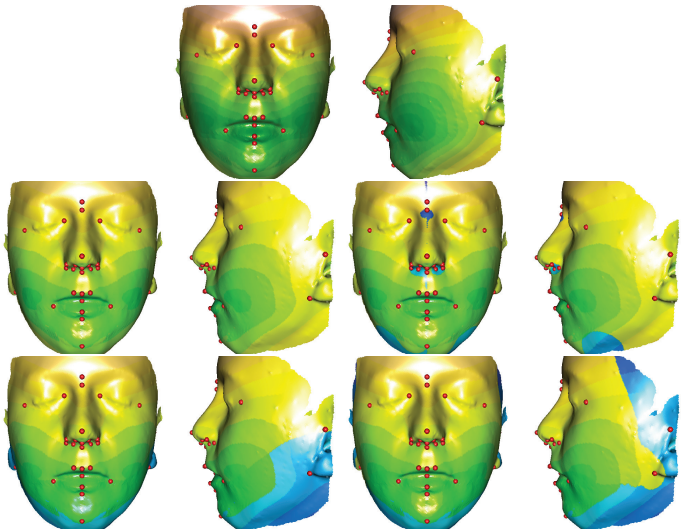
3D facial shape—VCFS data, differences between cases and controls (paired data)



Obrázok: ***PCA of reversible images*** (original, and relabeled and reflected faces) with projection of shape of zero difference—***testing case-control mean difference in particular PC subspace***

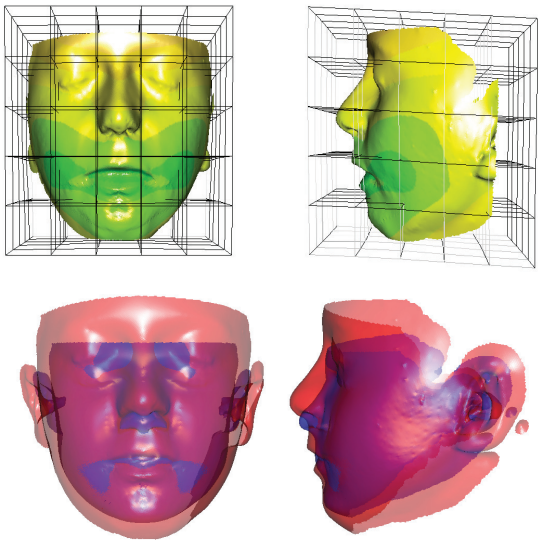
TPS deformations in topo-colors

PCA estimates—control-to-case 3D Euclidean distance, signed distance; *x*-, *y*-, and *z*-axis direction (shape space)



TPS deformations in topo-colors and wireframes

PCA estimates—control-to-case 3D signed Euclidean distance, wireframes, and transparent visualization (shape space)



2-block PLS

- **asymmetric 2-block PLS**—traditional focus of the PLS methods—find the directions in $\mathbf{X}_1$ that best describe $\mathbf{X}_2$ in some way—prediction of dependent variables $\mathbf{X}_2$ from independent variables $\mathbf{X}_1$ (Martens & Naes 1989, Joreskog & Wolf 1982)
- **symmetric 2-block PLS**—low-dimensional linear relationship between two high-dimensional measurement blocks by adapting one single *SVD* (Sampson et al. 1989, Bookstein 1994, McIntosh et al. 1996)

2-block PLS

- Let $\mathbf{S_x} = \frac{1}{n}\mathbf{X}_S^T\mathbf{X}_S$ be the sample covariance matrix and then

$$\mathbf{S_x} = \begin{pmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{pmatrix},$$

where $k_1 + k_2 = k$ $(k_1 < k_2)$ is number of landmarks, k_1 is number of landmarks in the first block and k_2 in the second block, $\mathbf{S}_{bb}$ is $dk_b \times dk_b$ sample covariance matrix of the bth block, $\mathbf{S}_{12} = \mathbf{S}_{21}^T$ is $dk_1 \times dk_2$ **sample cross-block covariance matrix** and is equal to

$$\mathbf{S}_{12} = \frac{1}{n}\mathbf{X}_{S,1}^T\mathbf{X}_{S,2}$$

2-block PLS

- adapting the *SVD* to $\mathbf{S}_{12}$ we get

$$\mathbf{S}_{12} = \widehat{\mathbf{U}}\widehat{\mathbf{\Lambda}}\widehat{\mathbf{V}}^T,$$

where $\widehat{\mathbf{U}}$ is the estimate of $dk_1 \times dk_1$ orthogonal matrix of **left singular vectors** with the columns $\widehat{\gamma}_{1j}$ $(j = 1, 2, \ldots dk_1)$ and $\widehat{\mathbf{V}}$ is the estimate of $dk_2 \times dk_2$ orthogonal matrix of **right singular vectors** with the columns $\widehat{\gamma}_{2j}$ $(j = 1, 2, \ldots dk_2)$ and $\widehat{\mathbf{\Lambda}}$ is the estimate of $dk_1 \times dk_2$ matrix of **singular values** $\widehat{\lambda}_j$ on the diagonal $(j = 1, 2, \ldots dk_1)$.

- *latent variables* (*scores*) are defined as

$$\mathbf{L}_1 = \mathbf{X}_{S,1}\widehat{\mathbf{U}}, \mathbf{L}_2 = \mathbf{X}_{S,2}\widehat{\mathbf{V}}$$

2-block PLS

- covariance between jth column $\mathbf{l}_{1j}$ of $\mathbf{L}_1$ and jth column $\mathbf{l}_{2j}$ of $\mathbf{L}_2$ is

$$Cov\left(\mathbf{l}_{1j}, \mathbf{l}_{2j}\right) = \widehat{\lambda}_j,$$

the maximum for any pair of such linear combination

- each column of $\widehat{\mathbf{U}}$ is proportional to the covariances of the block of $\mathbf{X}_{S,1}$ with the corresponding column of the matrix $\mathbf{L}_2$
- each column of $\widehat{\mathbf{V}}$ is proportional to the covariances of the block of $\mathbf{X}_{S,2}$ with the corresponding column of the matrix $\mathbf{L}_1$

2-block PLS

The additional graphical structure becomes available beyond $\widehat{\mathbf{U}}_{\cdot j}$ and $\widehat{\mathbf{V}}_{\cdot j}$ vectors—**scatter-plots of the latent variable scores** or **TPS grids** (2D), or the **arrows** and **TPS morphs** (3D) of the form, where we visualise

$$Vec(\overline{\mathbf{X}}_{P,1}) \pm c_{1j}\widehat{\mathbf{U}}_{\cdot j}, Vec(\overline{\mathbf{X}}_{P,2}) \pm c_{2j}\widehat{\mathbf{V}}_{\cdot j}$$

for the various values of $c_{1j}, c_{2j} \in \mathbb{R}^+$ (in the range of the particular *SW* scores or reasonable magnification of this range)

2-block PLS

A *SW* summary of the data (from each block separately) in the shape space

$$Vec\left(\mathbf{X}_{P,1,i}\right) = Vec\left(\overline{\mathbf{X}}_{P,1}\right) + \sum_{j=1}^{dk_1} l_{1,ij}\widehat{\mathbf{U}}_{\cdot j},$$

$$Vec\left(\mathbf{X}_{P,2,i}\right) = Vec\left(\overline{\mathbf{X}}_{P,2}\right) + \sum_{j=1}^{dk_2} l_{2,ij}\widehat{\mathbf{V}}_{\cdot j},$$

where $(\mathbf{L}_b)_{ij} = l_{b,ij}$ $(b = 1, 2)$

2-block PLS

SW summary for any q-subset of *SW*s $\{SW_{j_1}, ...SW_{j_q}\}, q \geq 1$ can be written as

$$Vec\left(\mathbf{X}_{P,1,i}\right)_{SW(j_1,...j_q)} = Vec\left(\overline{\mathbf{X}}_{P,1}\right) + \sum_{j_1,...j_q} l_{1,ij}\widehat{\mathbf{U}}_{\cdot j},$$

$$Vec\left(\mathbf{X}_{P,2,i}\right)_{SW(j_1,...j_q)} = Vec\left(\overline{\mathbf{X}}_{P,2}\right) + \sum_{j_1,...j_q} l_{2,ij}\widehat{\mathbf{V}}_{\cdot j}, i = 1, ...n,$$

and then $\mathbf{X}_{P,b}^{SW(j_1,...j_q)}$ are the matrices of all $Vec\left(\mathbf{X}_{P,b,i}\right)_{SW(j_1,...j_q)}$

2-block PLS

- to visualize a **composite shape** (both blocks together) we have to *scale* the singular vectors properly (Mitteroecker & Bookstein 2007)
- **block-wise matrix of common factor scores** $\mathbf{I}_j = \left(\mathbf{I}_{1j} : \mathbf{I}_{2j}\right)$
- necessary scaling factor—eigenvectors from *SVD* of the matrix $\mathbf{I}_j^T\mathbf{I}_j$ are $\widehat{\varphi}_j = \left(\widehat{\varphi}_{1j1}, \widehat{\varphi}_{2j1}\right)^T$
- **composite singular vectors**

$$\mathbf{f}_j = \begin{pmatrix} \widehat{\varphi}_{1j1}\widehat{\mathbf{U}}_{\cdot j} \\ \widehat{\varphi}_{2j1}\widehat{\mathbf{V}}_{\cdot j} \end{pmatrix}$$

2-block PLS

- **composite shape** $Vec(\overline{\mathbf{X}}_P) \pm c_j \mathbf{f}_j^{(sr)}$ for the various values of $c_j \in \mathbb{R}^+$ (in the range of the particular *SW* scores or reasonable magnification of this range)
- let matrix of **composite latent variables** (**composite scores**) be $\mathbf{L}^{(sr)} = \mathbf{X}_S \mathbf{F}^{(sr)}$, $(\mathbf{L}^{(sr)})_{ij} = l_{ij}$, the columns of $\mathbf{F}^{(sr)}$ be $\mathbf{f}_j^{(sr)}$ (Katina 2008)
- *SW* summary of the data in the shape space $Vec\left(\mathbf{X}_{P,i}\right) = Vec\left(\overline{\mathbf{X}}_P\right) + \sum_{j=1}^{dk_1} l_{ij} \mathbf{f}_j^{(sr)}$ and *SW* summary for any q-subset of *SW*s $\{SW_{j_1}, ...SW_{j_q}\}, q \geq 1$ can be written as $Vec\left(\mathbf{X}_{P,i}\right)_{SW(j_1,..j_q)} = Vec\left(\overline{\mathbf{X}}_P\right) + \sum_{j_1,...j_q} l_{ij} \mathbf{f}_j^{(sr)}, i = 1, ...n,$ and then $\mathbf{X}_P^{SW(j_1,..j_q)}$ is the matrix of all $Vec\left(\mathbf{X}_{P,i}\right)_{SW(j_1,..j_q)}$

2-block GPLS

- let $\mathbf{B}_e$ be bending energy matrix of $\overline{\mathbf{X}}_P$

$$\mathbf{B}_e = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix},$$

  where $k_1 + k_2 = k$ $(k_1 < k_2)$ is number of landmarks, k_1 is number of landmarks in the first block and k_2 in the second block, $\mathbf{B}_{bb} = \mathbf{0}$ is the $k_b \times k_b$ bending energy matrix of the bth block, $\mathbf{B}_{12} = \mathbf{B}_{21}^T$ is the $k_1 \times k_2$ **cross-block bending energy matrix**
- $dk \times dk$ matrix $\mathbf{B} = \mathbf{I}_{d \times d} \otimes \mathbf{B}_e, d = 2, 3$

2-block GPLS

- let $\mathbf{S}_B = (\mathbf{B}^-)^{\alpha/2} \widehat{\Sigma}_S^{(12)} (\mathbf{B}^-)^{\alpha/2}$
- then

$$\mathbf{S}_B = \begin{pmatrix} \mathbf{S}_{11}^{(B)} & \mathbf{S}_{12}^{(B)} \\ \mathbf{S}_{21}^{(B)} & \mathbf{S}_{22}^{(B)} \end{pmatrix},$$

- let $\mathbf{S}_{12}^{(B)}$ be **weighted cross-block covariance matrix**, $\left(\mathbf{B}_e^-\right)^{\alpha/2} = \sum_j \widehat{\lambda}_j^{-\alpha/2} \widehat{\gamma}_j^T \widehat{\gamma}_j$ (Moore-Penrose generalized inverse of $\mathbf{B}_e^{\alpha/2}$)

2-block GPLS

- **large scale variability**, $\alpha = 1$,
- **small scale variability**, $\alpha = -1$,
- $\alpha = 0$, then $\mathbf{B}_e^0 = \mathbf{I}$, the $k \times k$ identity matrix
- then SVD of

$$\mathbf{S}_{12}^{(B)} = \widehat{\mathbf{U}} \widehat{\mathbf{\Lambda}} \widehat{\mathbf{V}}^T,$$

- let $\mathbf{L}_B^{(sr)} = \mathbf{X}_S^{(B)} \mathbf{F}_B^{(sr)}$ be the matrix of **weighted composite latent variables (composite scores)**, $(\mathbf{L}_B^{(sr)})_{ij} = l_{B,ij}$, let the columns of $\mathbf{F}_B^{(sr)}$ be $\mathbf{f}_{B,j}^{(sr)}$

2-block GPLS

Then a *SW* summary of the data in the shape space is

$$Vec\left(\mathbf{X}_{P,i}\right) = Vec\left(\overline{\mathbf{X}}_P\right) + \mathbf{B}^{\alpha/2}\sum_{j=1}^{dk_1} l_{B,ij}\mathbf{f}_{B,j}^{(sr)}$$

and *SW* summary for any q-subset of *SW*s $\{SW_{j_1}, ...SW_{j_q}\}, q \geq 1$ can be written as

$$Vec\left(\mathbf{X}_{P,i}\right)_{SW(j_1,...j_q)} = Vec\left(\overline{\mathbf{X}}_P\right) + \mathbf{B}^{\alpha/2}\sum_{j_1,...j_q} l_{B,ij}\mathbf{f}_{B,j}^{(sr)}, i = 1, ...n,$$

and then $\mathbf{X}_P^{SW(j_1,...j_q)}$ is the matrix of all $Vec\left(\mathbf{X}_{P,i}\right)_{SW(j_1,...j_q)}$

Two shape subspaces and two different 2-block GPLS

Following Katina (2008)

- **affine** contribution to the variability—**affine subspace PLS** on $n \times k_b$ matrices $\mathbf{X}_{A,b}$ with the rows $\mathbf{x}_{A,bi}$, $i = 1, 2, ...n; b = 1, 2$
- **non-affine** contribution to the variability—**non-affine subspace PLS** on $n \times k_b$ matrices $\mathbf{X}_{NA,b}$ with the rows $\mathbf{x}_{NA,bi}$, $i = 1, 2, ...n$

Two different 2-block GPLS

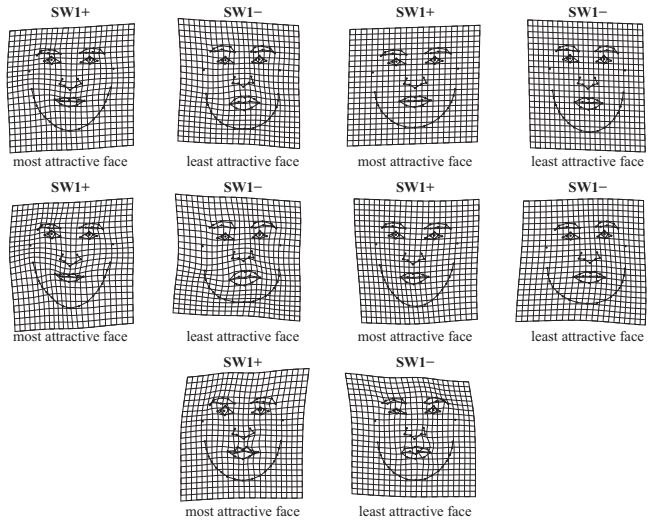
- if we have **two shape blocks**—Procrustes shape coordinates are pre-multiplied with $\left(\mathbf{B}_e^-\right)^{\alpha/2}$ of $\overline{\mathbf{X}}_P$ (Procrustes mean of the composite shape)
- if we have **one shape block and one block of external variables**—Procrustes shape coordinates of the shape block are pre-multiplied with $\left(\mathbf{B}_e^-\right)^{\alpha/2}$ of $\overline{\mathbf{X}}_{P_1}$(Procrustes mean shape)

Symmetric GPLS summary

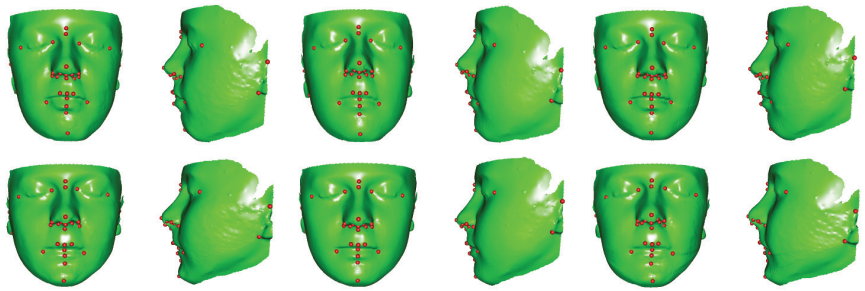
- **two shape blocks**
- **one shape block** and **one block of external variables**
- **shape space**
- **affine subspace**
- **non-affine subspace**
- **non-afine subspace with global and local bending**
- **one or more external variables**

Results of GPLS

TPS grids of shape block vs "attractiveness" block in all shape subspaces



Results of GPLS
3D warps of shape block vs SOFA scores in three different shape subspaces



Statistical inference in shape analysis
Outline

For one-, two-sample, and paired hypotheses about shapes, there are the following tests

1. ***one-sample Hotelling T^2 test***, ***one-sample Goodall F test***
2. ***two-independent sample Hotelling T^2 test***, ***modification of Nel-Van der Merwe test for the multivariate Behrens-Fisher problem***, and ***two independent sample Goodall F test***,
3. ***paired Hotelling T^2 test*** and ***paired Goodall F test***
4. ***Mardia test of object symmetry***

Moore-Penrose generalized inverse of symmetric square matrix $\mathbf{A}$, let say $\mathbf{A}^-$, is inverse, where following equation holds $\mathbf{A}^-\mathbf{A}\mathbf{A}^- = \mathbf{A}$, so

$$\mathbf{A}^- = \sum_{j=1}^{s} \lambda_j^{-1} \gamma_j \gamma_j^T,$$

where γ_j are ***eigenvectors*** of matrix $\mathbf{A}$ corresponding to ***eigenvalues*** $\lambda_j > 0$, where $j = 1, 2...s \leq kd$.

Statistical inference in shape analysis
One-sample Multivariate Inference

Definition (One-sample tests)

Let $Vec(\mathbf{X}_{P,i}), i = 1, 2, ...n$, be the random sample from population with vectorized ***Procrustes mean shape*** $Vec\,(\boldsymbol{\mu}_P)$ estimated by $\overline{\mathbf{x}}_P$ and ***covariance matrix*** $\boldsymbol{\Sigma}_P$ estimated by $\mathbf{S}_\mathbf{X}$. Let

$$Vec(\mathbf{X}_{P,i}) \sim N_{dk}\left(Vec\left(\boldsymbol{\mu}_P\right), \boldsymbol{\Sigma}_P\right), i = 1, ...n.$$

The null hypothesis is defined as: the Procrustes mean shape $\boldsymbol{\mu}_P$ is equal to the Procrustes mean shape $\boldsymbol{\mu}_0$, so $H_0 : \boldsymbol{\mu}_P = \boldsymbol{\mu}_0, H_1 : \boldsymbol{\mu}_P \neq \boldsymbol{\mu}_0$. If H_0 holds, ***Hotelling T^2 test statistic*** is equal to

$$F_H = \frac{n-s}{s} T_H^2 \sim F_{s,n-s},$$

where $s = min(dk, n-1)$, and

$$T_H^2 = \left(\overline{\mathbf{x}}_P - Vec\left(\boldsymbol{\mu}_0\right)\right)^T \mathbf{S}_\mathbf{X}^- \left(\overline{\mathbf{x}}_P - Vec\left(\boldsymbol{\mu}_0\right)\right) = \sum_{j=1}^{s} \frac{\widehat{r}_{j0}^2}{\widehat{\lambda}_j}$$

is square of ***Mahalanobis distance*** between $\overline{\mathbf{x}}_P$ and $Vec\,(\boldsymbol{\mu}_0)$, where $\mathbf{S}_\mathbf{X}^-$ is Moore-Penrose generalized inverse of $\mathbf{S}_\mathbf{X}$;

Statistical inference in shape analysis
One-sample Multivariate Inference

Definition (One-sample tests; cont.)

$\widehat{r}_{j0} = \widehat{\gamma}_j^T \left(\overline{\mathbf{x}}_P - Vec\left(\boldsymbol{\mu}_0\right)\right)$ is the jth PC score for the difference $\left(\overline{\mathbf{x}}_P - Vec\left(\boldsymbol{\mu}_0\right)\right), j = 1, 2, ..., s$. High values of $\widehat{r}_{j0}^2/\widehat{\lambda}_j$ indicates the direction of high shape variability associated with $\overline{\mathbf{x}}_P$ in jth PC. The test statistic T_H^2 can be modified with respect to any subset of PCs as

$$T_H^2 = \left(\overline{\mathbf{x}}_P - Vec\left(\boldsymbol{\mu}_0\right)\right)^T \left(\mathbf{S}_\mathbf{X}^{PC(j_1,...j_q)}\right)^- \left(\overline{\mathbf{x}}_P - Vec\left(\boldsymbol{\mu}_0\right)\right) = \sum_{j_1,..j_q} \frac{\widehat{r}_{j0}^2}{\widehat{\lambda}_j}$$

where $\mathbf{S}_\mathbf{X}^{PC(j_1,...j_q)} = \sum_{PC(j_1,...j_q)} \widehat{\lambda}_j \widehat{\gamma}_j \widehat{\gamma}_j^T$ is the covariance matrix estimated by any q-subset of PCs $\left\{PC_{j_1}, PC_{j_2}, ..., PC_{j_q}\right\}; q \geq 1$.
If covariance matrix $\boldsymbol{\Sigma}_P = \sigma^2\mathbf{I}$ and if H_0 holds, ***Goodall test statistic***

$$F_G = n\left(n-1\right) \frac{d_F^2\left(\overline{\mathbf{X}}_P, \boldsymbol{\mu}_0\right)}{\sum_{i=1}^{n} d_F^2\left(\mathbf{X}_{P,i}, \overline{\mathbf{X}}_P\right)} \sim F_{s,n-s},$$

which is the special case of Hotelling T^2 under the isotropy.

Definition (Two-sample tests)

Let $Vec(\mathbf{X}_{P,ji}), i = 1,2,...n_j$,be the random sample from population $j, j = 1,2$, with vectorized ***Procrustes mean shape*** $Vec\left(\boldsymbol{\mu}_{P,j}\right)$ estimated by $\overline{\mathbf{x}}_{P,j}$ and ***covariance matrix*** $\boldsymbol{\Sigma}_P$ estimated by ***common covariance matrix*** $\mathbf{S}_U = \left(n_1\mathbf{S}_{\mathbf{X},1} + n_2\mathbf{S}_{\mathbf{X},2}\right)/\left(n_1 + n_2 - 2\right)$, where ***sample covariance matrices*** $\mathbf{S}_{\mathbf{X},j} = \frac{1}{n}\mathbf{X}_{P,j}^T\mathbf{H}\mathbf{X}_{P,j}$, $\mathbf{X}_{P,j}$ is $n_j \times (dk)$ matrix of $Vec(\mathbf{X}_{P,ji})$ as the rows. Let

$$Vec(\mathbf{X}_{P,ji}) \sim N_{dk}\left(Vec\left(\boldsymbol{\mu}_{P,j}\right), \boldsymbol{\Sigma}_P\right); j = 1,2; i = 1,...n.$$

The null hypothesis is defined as: the Procrustes mean shape $\boldsymbol{\mu}_{P,1}$ is equal to the Procrustes mean shape $\boldsymbol{\mu}_{P,2}$, so $H_0 : \boldsymbol{\mu}_{P,1} = \boldsymbol{\mu}_{P,2}, H_1 : \boldsymbol{\mu}_{P,1} \neq \boldsymbol{\mu}_{P,2}$. If H_0 holds, ***Hotelling T^2 test statistic*** is equal to

$$F_H = \frac{n_1 n_2\left(n_1 + n_2 - s - 1\right)}{\left(n_1 + n_2\right)\left(n_1 + n_2 - 2\right)s}T_H^2 \sim F_{s,n_1+n_2-s-1},$$

where $s = min(dk, n_1 + n_2 - 2)$, and

$$T_H^2 = \left(\overline{\mathbf{x}}_{P,1} - \overline{\mathbf{x}}_{P,2}\right)^T \mathbf{S}_U^{-}\left(\overline{\mathbf{x}}_{P,1} - \overline{\mathbf{x}}_{P,2}\right) = \sum_{j=1}^{s}\frac{\widehat{r}_{j0}^2}{\widehat{\lambda}_j},$$

Definition (Two-sample tests; cont.)

$\widehat{r}_{j0} = \widehat{\gamma}_j^T\left(\overline{\mathbf{x}}_{P,1} - \overline{\mathbf{x}}_{P,2}\right)$ is the jth PC score for the difference $\left(\overline{\mathbf{x}}_P - \overline{\mathbf{x}}_{P,2}\right)$, $j = 1,2,...,s$. High values of $\widehat{r}_{j0}^2/\widehat{\lambda}_j$ indicates the direction of high shape variability associated with observed group difference $\overline{\mathbf{x}}_{P,1} - \overline{\mathbf{x}}_{P,2}$ in jth PC. The test statistic T_H^2 can be modified with respect to any subset of PCs as

$$T_H^2 = \left(\overline{\mathbf{x}}_{P,1} - \overline{\mathbf{x}}_{P,2}\right)^T\left(\mathbf{S}_U^{PC(j_1,...j_q)}\right)^{-}\left(\overline{\mathbf{x}}_{P,1} - \overline{\mathbf{x}}_{P,2}\right) = \sum_{j_1,...j_q}\frac{\widehat{r}_{j0}^2}{\widehat{\lambda}_j}$$

where $\mathbf{S}_U^{PC(j_1,...j_q)} = \sum_{PC(j_1,...j_q)}\widehat{\lambda}_j\widehat{\gamma}_j\widehat{\gamma}_j^T$ is the covariance matrix estimated by any q-subset of PCs $\left\{PC_{j_1}, PC_{j_2}, ..., PC_{j_q}\right\}; q \geq 1$.
If covariance matrix $\boldsymbol{\Sigma}_{P,j} = \sigma^2\mathbf{I}$ and if H_0 holds, ***Goodall test statistic***

$$F_G = \frac{n_1+n_2-2}{n_1^{-1}+n_2^{-1}}\frac{d_F^2\left(\overline{\mathbf{x}}_{P,1}, \overline{\mathbf{x}}_{P,2}\right)}{\sum_{i=1}^{n_1} d_F^2\left(\mathbf{x}_{P,1i}, \overline{\mathbf{x}}_{P,1}\right) + \sum_{i=1}^{n_2} d_F^2\left(\mathbf{x}_{P,2i}, \overline{\mathbf{x}}_{P,2}\right)} \sim F_{s,(n_1+n_2-2)s},$$

which is the special case of Hotelling T^2 under the isotropy.

Definition (Paired tests)

Let $Vec(\mathbf{X}_{P,ji}), j = 1,2, i = 1,2,...n$, be the random sample from population with vectorized ***Procrustes mean shapes*** $Vec\left(\boldsymbol{\mu}_{P,j}\right)$ estimated by $\overline{\mathbf{x}}_{P,j}$ and ***covariance matrices*** $\boldsymbol{\Sigma}_{P,j}$ estimated by $\mathbf{S}_{\mathbf{X},j}$. Let

$$Vec(\mathbf{X}_{P,ji}) \sim N_{dk}\left(Vec\left(\boldsymbol{\mu}_{P,j}\right), \boldsymbol{\Sigma}_{P,j}\right), j = 1,2, i = 1,...n.$$

Let $Vec(\mathbf{X}_{D,i}) = Vec(\mathbf{X}_{P,1i} - \mathbf{X}_{P,2i})$, $i = 1,2,...,n$, be a random sample of the coordinate differences of one object with coordinates measured two times and then $Vec(\mathbf{X}_{D,i}) \sim N_{dk}\left(Vec\left(\boldsymbol{\mu}_D\right), \boldsymbol{\Sigma}_D\right)$. The estimates of parameters are $\overline{\mathbf{x}}_D$ and $\mathbf{S}_D$.
The null hypothesis is defined as: the Procrustes mean shape $\boldsymbol{\mu}_D$ is equal to the Procrustes mean shape $\boldsymbol{\mu}_0$, so $H_0 : \boldsymbol{\mu}_D = \boldsymbol{\mu}_0, H_1 : \boldsymbol{\mu}_D \neq \boldsymbol{\mu}_0$. If H_0 holds, ***Hotelling T^2 test statistic*** is equal to

$$F_H = \frac{n-s}{s}T_H^2 \sim F_{s,n-s},$$

where $s = min(dk, n-1)$, and

Definition (Paired tests; cont.)

$$T_H^2 = \left(\overline{\mathbf{x}}_D - Vec\left(\boldsymbol{\mu}_0\right)\right)^T\mathbf{S}_D^{-}\left(\overline{\mathbf{x}}_D - Vec\left(\boldsymbol{\mu}_0\right)\right) = \sum_{j=1}^{s}\frac{\widehat{r}_{j0}^2}{\widehat{\lambda}_j};$$

$\widehat{r}_{j0} = \widehat{\gamma}_j^T\left(\overline{\mathbf{x}}_D - Vec\left(\boldsymbol{\mu}_0\right)\right)$ is the jth PC score for the difference $\left(\overline{\mathbf{x}}_D - Vec\left(\boldsymbol{\mu}_0\right)\right), j = 1,2,...,s$. High values of $\widehat{r}_{j0}^2/\widehat{\lambda}_j$ indicates the direction of high shape variability associated with $\overline{\mathbf{x}}_D$ in jth PC. The test statistic T_H^2 can be modified with respect to any subset of PCs as
$T_H^2 = \left(\overline{\mathbf{x}}_D - Vec\left(\boldsymbol{\mu}_0\right)\right)^T\left(\mathbf{S}_D^{PC(j_1,...j_q)}\right)^{-}\left(\overline{\mathbf{x}}_D - Vec\left(\boldsymbol{\mu}_0\right)\right) = \sum_{j_1,...j_q}\frac{\widehat{r}_{j0}^2}{\widehat{\lambda}_j}$, where $\mathbf{S}_D^{PC(j_1,...j_q)} = \sum_{PC(j_1,...j_q)}\widehat{\lambda}_j\widehat{\gamma}_j\widehat{\gamma}_j^T$ is the covariance matrix estimated by any q-subset of PCs $\left\{PC_{j_1}, PC_{j_2}, ..., PC_{j_q}\right\}; q \geq 1$.
If covariance matrix $\boldsymbol{\Sigma}_D = \sigma^2\mathbf{I}$ and if H_0 holds, ***Goodall test statistic***

$$F_G = n\left(n-1\right)\frac{d_F^2\left(\overline{\mathbf{X}}_D, \boldsymbol{\mu}_0\right)}{\sum_{i=1}^{n} d_F^2\left(\mathbf{X}_{D,i}, \overline{\mathbf{X}}_P\right)} \sim F_{s,n-s},$$

which is the special case of Hotelling T^2 under the isotropy.

Statistical inference in shape analysis
Confidence and Tolerance Ellipsoids

Definition (Confidence and Tolerance Ellipsoids)

If $k > 1$, then the generalization of $(1-\alpha)100\%$ ***confidence interval (CI)*** for $\boldsymbol{\mu}$ is $(1-\alpha)100\%$ ***confidence set (CS)*** for $\boldsymbol{\mu}$

$$CS = \left\{ \boldsymbol{\mu}_0 : \left(\overline{\mathbf{x}} - \boldsymbol{\mu}_0\right)^T \mathbf{S}^{-1} \left(\overline{\mathbf{x}} - \boldsymbol{\mu}_0\right) \leq \frac{(n-1)\,k}{(n-k)n} F_{k,n-k}(\alpha) \right\}.$$

Then $\Pr[CS \cap \{\boldsymbol{\mu}\} \neq \emptyset] = 1-\alpha$. We can calculate ***realization of*** $(1-\alpha)\,\%$ CS. It is ***confidence ellipsoid (CE)*** centered in $\overline{\mathbf{x}}$. The direction of ellipsoid-axes is parallel to eigenvectors $\widehat{\gamma}_j$ of $\mathbf{S}$ ($\widehat{\lambda}_j$ are particular eigenvalues). The length of ellipsoid-axes visualized from the center $\overline{\mathbf{x}}$ is equal to

$$\pm\sqrt{\widehat{\lambda}_j \frac{(n-1)\,k}{(n-k)\,n} F_{k,n-k}(1-\alpha)}, j = 1,2,...k.$$

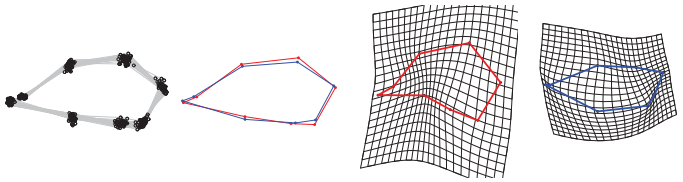
These CEs (in one-, two-sample, and paired case) can be applied to: ***(semi)landmark coordinates*** and ***PC scores***. Multiplying $F_{k,n-k}(\alpha)$ by n we get ***tolerance ellipsoid (TE)***.

Example 27
PCA a testovanie hypotéz

Example (DÚ 9)

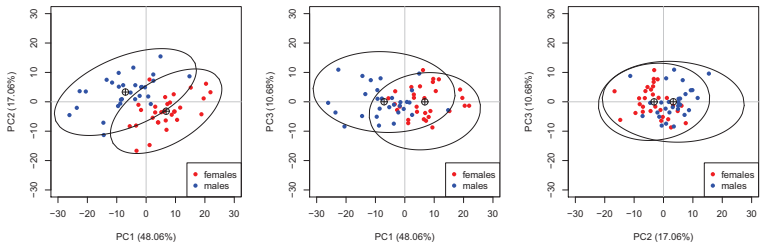
Majme dáta `gorf.dat` a `gorm.dat`, ktoré sú v knižnici `shapes` a predstavujú súradnice $k = 8$ landmarkov na lebkách $n = 30$ samíc a $n = 29$ samcov goríl (*Gorilla gorilla*). Pokrač. príkladu 7.

9.1) Registrujte súradnice landmarkov `gorf.dat` a `gorm.dat` do spoločného tvarového priestoru pomocou GPA a aplikujte algoritmus výpočtu rotácie do smeru najväčšej variability z DÚ7. Použite funkciu `procGPA(...)$rotated` (GPA, kde výstupom je pole rozmeru $8 \times 2 \times 59$ **procrustovských tvarových súradníc**). Vypočítajte priemerné procrustovské súradnice pre samice a samcov, deformujte súradnice samíc na samcov a naopak, extrapolujte $3\times$.

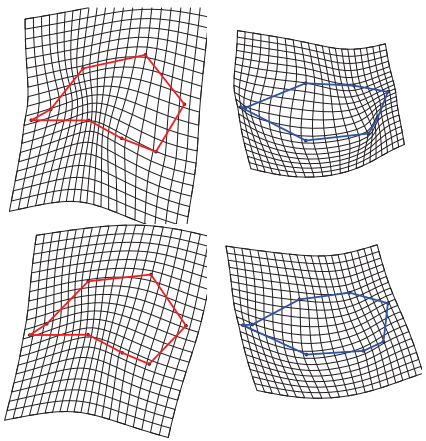


Example 27
PCA a testovanie hypotéz

9.2) Vypočítajte ***vlastné čísla*** a ***vlastné vektory*** kovariančnej matice $\mathbf{S}_\mathbf{X}$ centrovaných procrustovských tvarových súradníc. Použite funkciu `eigen()`. Skontrolujte, či majú všetky vlastné vektory jednotkovú dĺžku. Škálujte vlastné čísla ich sumou, vynásobte 100 (zaokrúhlite na dve desatinné miesta) a kumulatívne ich zosumujte. Použite fukncie `sum()` a `cumsum()`. Zobrazte skóre PC_i vs PC_j, $j = 1,2,3$; $i < j$ v rozptylových grafoch (rozsahy všetkých grafov škálujte rovnako) spolu s 95% ***tolerančnými elipsoidmi***. Vypočítajte priemerné procrustovské súradnice pre samice a samcov v podpriestore PC1 (***spätnou projekciou skóre do tvarového priestoru*** – viď. slajdy o klasickej alebo zovšeobecnenej PCA), extrapolujte $3\times$. Porovnajte obrázky s (9.1) a interpretujte použitím matematicko-štatistického pojmového aparátu.



Example 27
PCA a testovanie hypotéz



Obrázok: TPS deformácie samcov na samice a naopak samíc na samcov; priemerné procrustovské tvary (horný riadok), odhadnuté priemerné procrustovské tvary v podpriestore PC1 (dolný riadok); extrapolované $3\times$