ON THE GEOMETRY OF MULTIVARIATE L_1 OBJECTIVE FUNCTIONS

Hannu Oja

May 12, 2003

University of Jyväskylä, FINLAND

http://www.maths.jyu.fi/~ojahannu/

1 INTRODUCTION

- Objective function I: ave $\{V(0, y_{i_1}, ..., y_{i_k})\}$

  – mean difference

- Objective function II: ave $\left\{V(y_{j_1}, ..., y_{j_{k+1}})\right\}$

  – mean deviance

- Data based hyperplanes going through the origin, normals to hyperplanes, (angular) distance between data vectors and between normals

- Data based hyperplanes, normals,

  distance between data vectors and between hyperplanes

- On location and scatter estimates

- Optimal location tests based on hyperplanes

2 L_1 OBJECTIVE FUNCTIONS

- $Y = \{y_1, \ldots, y_n\}$ - a set of k-variate residuals

- Volume of the simplex with vertices $y_{j_1}, ..., y_{j_{k+1}}$ (multiplied by $k!$)

$$V(y_{j_1}, ..., y_{j_{k+1}}) = \text{abs}\left\{\det\begin{pmatrix} 1 & \cdots & 1 \\ y_{j_1} & \cdots & y_{j_{k+1}} \end{pmatrix}\right\}$$

- Objective function I: ave $\{V(0, y_{i_1}, ..., y_{i_k})\}$ $\Longrightarrow$

  median-type estimates, sign tests [univariate case: ave $\{|y_i|\}$]

- Objective function II: ave $\{V(y_{j_1}, ..., y_{j_{k+1}})\}$ $\Longrightarrow$

  Hodges-Lehmann-type estimates, rank tests [univariate case: ave $\{|y_j - y_i|\}$]

- One sample, two samples, several samples affine equivariant/invariant multivariate estimates and tests;

  see Oja(1983), Brown and Hettmansperger (1987,1989), ..., Oja (1999)

- Depth of y: ave $\{V(y, y_{i_1}, ..., y_{i_k})\}$ $\Longrightarrow$

  median-type estimates, sign tests [univariate case: ave $\{|y_i|\}$]

- L_1 objective functions, zonotopes and lift zonotopes:

$$\text{ave}\{V(0, y_{i_1}, ..., y_{i_k})\} = 2^{-k}\text{ave}\{[-y_i, y_i]\}$$

and

$$\text{ave}\{V(y_{j_1}, y_{j_2}, ..., y_{j_{k+1}})\} = 2^{-(k+1)}\text{ave}\left\{\left[-\begin{pmatrix}1\\ y_i\end{pmatrix}, \begin{pmatrix}1\\ y_i\end{pmatrix}\right]\right\}$$

$\Longrightarrow$ another story .. .

3 DATA BASED HYPERPLANES I

- $Y = \{y_1, \ldots, y_n\}$ - a k-variate (centered) data set
- $I = (i_1, ..., i_{k-1});\ \ 1 \le i_1 < ... < i_{k-1} \le n$
- **hyperplane** (going through the origin and $k-1$ data points)

$$H(I) = \{y\ :\ \det\left(y_{i_1}\quad ...\quad y_{i_{k-1}}\quad y\right) = 0\}$$

- **normal** to hyperplane

$$e(I): \qquad \det\left(y_{i_1}\quad ...\quad y_{i_{k-1}}\quad y\right) = e^T(I)y$$

- Define $k \times k$ matrices

$$V(I) = e(I)e^T(I)\ \ \text{and}\ \ V_i = y_i y_i^T$$

- a measure of **angular closeness** of two data vectors y_i and $y_{i'}$:

$$\delta_{i,i'} = \text{ave}_I\left\{sgn\left(y_i^T\ V(I)\ y_{i'}\right)\right\} \quad \in (-1,1)$$

- **NOTE.** Randles' interdirections (1989): $(1-\delta_{i,i'})/2\ \in (0,1)$
- a measure of **angular closeness** of two normals $e(I)$ and $e(I')$:

$$\Delta(I,I') = \text{ave}_i\left\{sgn\left(e^T(I)\ V_i\ e(I')\right)\right\} \quad \in (-1,1)$$

4 OBJECTIVE FUNCTION ave $\left\{V(0, y_{i_1}, ..., y_{i_k})\right\}$ – MEAN DEVIANCE

- $I = (i_1, ..., i_{k-1});\ \ 1 \le i_1 < ... < i_{k-1} \le n$

- $V(0, y_{i_1}, ..., y_{i_{k-1}}, y_i)$

  $= \text{abs}\{\det\begin{pmatrix} y_{i_1} & ... & y_{i_{k-1}} & y_i\end{pmatrix}\}$

  $= \text{abs}\{e^T(I)y_i\}$

- $\text{ave}\left\{\text{abs}\{e^T(I)y_i\}\right\}$

  $= \text{ave}\left\{\text{sign}(e^T(I)y_i) \cdot e^T(I)y_i\right\}$

- Use $\text{ave}_{i,I}\{\cdot\} = \text{ave}_i\{\text{ave}_I\{\cdot\}\} = \text{ave}_I\{\text{ave}_i\{\cdot\}\}$ ...

- ... and write

$$\text{ave}\left\{\text{abs}\{e^T(I)y_i\}\right\} \;=\; \text{ave}_i\{S_i^T y_i\} \;=\; \text{ave}_I\{h^T(I)e(I)\}$$

- y_i and $e(I)$ given before. S_i and $h(I)$??

- Interpretation: S_i is the **multivariate sign vector** of y_i as,

  in the univariate case,

$$\text{ave}\{|y_i|\} = \text{ave}\{sgn(y_i)y_i\}$$

5 ave $\left\{V(0, y_{i_1}, ..., y_{i_{k-1}}, y_i)\right\} = \mathbf{ave}\{S_i^T y_i\} = \mathbf{ave}\{h^T(I)e(I)\}$

- all vectors

$$y_i, S_i, h(I), e(I)$$

  carry information about the covariance structure (scatter) !!!

- $\text{ave}\{e(I)e^T(I)\} \;=\; \text{a constant} \;\times\; \left[\text{ave}\{y_i y_i^T\}\right]^{-1}$

- **Sign Covariance Matrix (SCM)** (Ollila, Oja, Croux, 2003)

$$SCM \;=\; \text{ave}\{S_i S_i^T\} \;=\; \text{ave}\left\{\Delta(I, I')\; e(I)e^T(I')\right\}$$

- **Zonoid Covariance Matrix (ZCM)** (Koshevoy, Möttönen, Oja, 2003)

$$ZCM \;=\; \text{ave}\{h(I)h^T(I)\} \;=\; \text{ave}\left\{\delta_{i,i'}\; y_i y_{i'}^T\right\}$$

- the use of SCM in multivariate regression:

  Ollila, Hettmansperger and Oja (2002)

6 DATA BASED HYPERPLANES II

- $Y = \{y_1, \ldots, y_n\}$ - a k-variate data set

- $\tilde{Y} = \{\tilde{y}_1, \ldots, \tilde{y}_n\}$ where $\tilde{y}_i = (1, y_i^T)^T$

  - a $k+1$-variate lift data set

- $J = (j_1, ..., j_k)$; $1 \le j_1 < ... < j_k \le n$

- **hyperplane** (going through k data points $y_{j_1}, ..., y_{j_k}$)

$$\tilde{H}(J) = \{y \ : \ \det\left(\tilde{y}_{j_1} \quad ... \quad \tilde{y}_{j_k} \quad \tilde{y}\right) = 0\}$$

- Define $d(J) = (d_0(J), d_1^T(J))^T$ implicitly by

$$\det\left(\tilde{y}_{j_1} \quad ... \quad \tilde{y}_{j_k} \quad \tilde{y}\right) = d^T(J)\tilde{y}$$

  $(d_1(J))$ is orthogonal to $\tilde{H}(J))$

- Define $(k+1) \times (k+1)$ matrices

$$\tilde{V}(J) = d(J)d^T(J) \quad \text{and} \quad \tilde{V}_j = \tilde{y}_j\tilde{y}_j^T$$

- a measure of **closeness** of two data points y_j and $y_{j'}$:

$$\tilde{\delta}_{j,j'} = \text{ave}_J \left\{ sgn\left(\tilde{y}_j^T \ \tilde{V}(J) \ \tilde{y}_{j'}\right)\right\} \quad \in (-1, 1)$$

- **NOTE.** 'Lift interdirections' : $(1 - \tilde{\delta}_{j,j'})/2 \ \in (0, 1)$

- a measure of **closeness** between two hyperplanes $d(J)$ and $d(J')$:

$$\tilde{\Delta}(J, J') = \text{ave}_j \left\{ sgn\left(d^T(J) \ \tilde{V}_j \ d(J')\right)\right\} \quad \in (-1, 1)$$

7 OBJECTIVE FUNCTION ave $\left\{V(y_{j_1},...,y_{j_{k+1}})\right\}$ – MEAN DIFFERENCE

- $J = (j_1, ..., j_k);\;\; 1 \le j_1 < ... < j_k \le n$

- $V(y_{j_1}, ..., y_{j_k}, y_j)$

  $= \text{abs}\{\det\left(\tilde{y}_{j_1} \;\; ... \;\; \tilde{y}_{j_k} \;\; \tilde{y}_j\right)\}$

  $= \text{abs}\{d^T(J)\tilde{y}_j\}$

- ave $\left\{\text{abs}\{d^T(J)\tilde{y}_j\}\right\}$

  $= \text{ave}\left\{\text{sign}(d^T(J)\tilde{y}_j)\cdot d^T(J)\tilde{y}_j\right\}$

- Use $\text{ave}_{j,J}\{\cdot\} = \text{ave}_j\{\text{ave}_J\{\cdot\}\} = \text{ave}_J\{\text{ave}_j\{\cdot\}\}$ ...

- ... and write

$$\text{ave}\left\{\text{abs}\{d^T(J)\tilde{y}_j\}\right\} \;=\; \text{ave}_i\{\tilde{R}_j^T\tilde{y}_j\} \;=\; \text{ave}_J\{g^T(J)d(J)\}$$

- $\tilde{y}_j$ and $d(J)$ given before. $\tilde{R}_j$ and $g(J)$??

- Interpretation: $\tilde{R}_j$ is called the **multivariate lift rank vector** of y_j as, in the univariate case,

$$\text{ave}\{|y_j - y_i|\} = \text{ave}\left\{sgn(y_i - y_j)(-y_j \;\; 1)\begin{pmatrix}1\\ y_i\end{pmatrix}\right\}$$

and

$$\tilde{R}_j = \text{ave}_j\left\{sgn(y_j - y_i)\begin{pmatrix}-y_i\\ 1\end{pmatrix}\right\}$$

(the second component is the regular univariate centered rank.

8 $\textbf{ave}\left\{V(y_{j_1},...,y_{j_k},y_j)\right\} = \textbf{ave}\{\tilde{R}_j^T\tilde{y}_j\} = \textbf{ave}\{g^T(J)d(J)\}$

- all vectors

  $$\tilde{y}_j, \tilde{R}_j, g(J), d(J)$$

  carry information about location and scatter !!!

- $$\text{ave}\{\tilde{y}_j\tilde{y}_j^T\} \;=\; \begin{pmatrix} 1 & \bar{y}^T \\ \bar{y} & S+\bar{y}\bar{y}^T \end{pmatrix}$$

- $$\text{ave}\{d(J)d^T(J)\} \;=\; \text{a constant} \;\times\; \left[\text{ave}\{\tilde{y}_j\tilde{y}_j^T\}\right]^{-1}$$

- **Lift Rank Covariance Matrix (LRCM)** (Ollila, Oja, Koivunen, 2003)

  $$LRCM \;=\; \text{ave}\{\tilde{R}_j\tilde{R}_j^T\} \;=\; \text{ave}\left\{\tilde{\Delta}(J,J')\; d(J)d^T(J')\right\}$$

- **Lift Zonoid Covariance Matrix (LZCM)**

  $$LZCM \;=\; \text{ave}\{g(J)g^T(J)\} \;=\; \text{ave}\left\{\tilde{\delta}_{j,j'}\; \tilde{y}_j\tilde{y}_{j'}^T\right\}$$

9 Optimal one-sample location tests

- $Y = \{y_1, ..., y_n\}$ is a random sample from an elliptical distribution

  with symmetry center μ

- null hypothesis $H_0 : \mu = 0$

- interdirections

$$p_{ij} = \text{ave}_I \left\{ 1_{y_i^T V(I) y_j > 0} \right\}$$

- lift interdirections between y_i and $-y_i$

$$r_i = \text{ave}_I \left\{ 1_{\tilde{y}_i^T \tilde{V}(I) \widetilde{(-y_i)} > 0} \right\}$$

- R_i is the rank of r_i among $\{r_1, ..., r_n\}$

- **rank scores test statistic**

$$Q \;=\; nk \;\text{ave}_{i,j} \left\{ a\left(\frac{R_i}{n+1}\right) a\left(\frac{R_j}{n+1}\right) \; \cos(\pi p_{ij}) \right\}$$

  where $a(u)$ is the score function ($\int a^2(u)du = 1$).

  van der Waerden test: $a(u) = \chi_{k,u}/\sqrt{k}$

- See Hallin and Paindaveine (2002), Oja and Paindaveine (2003, in preparation)

Table 1: Objective function $\text{ave}\left\{V(0, y_{i_1}, ..., y_{i_{k-1}}, y_i)\right\} = \text{ave}\{S_i^T y_i\} = \text{ave}\{h^T(I)e(I)\}$.

	original space $\mathbb{R}^k$	dual space $\mathbb{R}^k$
index	i, i'	I, I'
vector	y_i	$e(I)$
scatter matrix	$\text{ave}\left\{y_i y_i^T\right\}$	$\text{ave}\left\{e(I)e^T(I)\right\}$
hyperplane	$\{y \ : \ e^T(I)y = 0\}$	$\{e \ : \ y_i^T e = 0\}$
ang. dist.	$\delta_{i,i'} =$ $\text{ave}_I\left\{sgn\left(y_i^T V(I) y_{i'}\right)\right\}$	$\Delta(I, I') =$ $\text{ave}_i\left\{sgn\left(e^T(I) V_i e(I')\right)\right\}$
sign	$S_i =$ $\text{ave}_I\left\{sgn(e^T(I)y_i) \cdot e(I)\right\}$	$h(I) =$ $\text{ave}_i\left\{sgn(e^T(I)y_i) \cdot y_i)\right\}$
SCM&ZCM	$\text{ave}\{S_i S_i^T\} =$ $\text{ave}\left\{\Delta(I, I') \ e(I)e^T(I')\right\}$	$\text{ave}\{h(I)h^T(I)\} =$ $\text{ave}\left\{\delta_{i,i'} \ y_i y_{i'}^T\right\}$
zonotope	$\text{ave}\{[-y_i, y_i]\}$	$\text{ave}\{[-e(I), e(I)]\}$
facet centre	$h(I)$	S_i, etc.

Table 2: Objective function ave $\{V(y_{j_1},...,y_{j_k},y_j)\} = \text{ave}\{\tilde{R}_j^T\tilde{y}_j\} = \text{ave}\{g^T(J)d(J)\}$.

	original space $\mathbb{R}^{k+1}$	dual space $\mathbb{R}^{k+1}$
index	j, j'	J, J'
vector	$\tilde{y}_j$	$d(J)$
scatter matrix	ave $\left\{\tilde{y}_j\tilde{y}_j^T\right\}$	ave $\left\{d(J)d^T(J)\right\}$
hyperplane	$\{y \ : \ d^T(J)\tilde{y} = 0\}$	$\{d \ : \ \tilde{y}_j^T d = 0\}$
distance	$\tilde{\delta}_{j,j'} =$ $\text{ave}_J \left\{sgn\left(\tilde{y}_j^T\tilde{V}(J)\tilde{y}_{j'}\right)\right\}$	$\tilde{\Delta}(J,J') =$ $\text{ave}_j \left\{sgn\left(d^T(J)\tilde{V}_j d(J')\right)\right\}$
lift rank	$\tilde{R}_j =$ $\text{ave}_J \left\{sgn\left(d^T(J)\tilde{y}_j\right)\cdot d(J)\right\}$	$g(J) =$ $\text{ave}_j \left\{sgn\left(d^T(J)\tilde{y}_j\right)\cdot \tilde{y}_j\right\}$
LRCM & LZCM	$\text{ave}\{\tilde{R}_j\tilde{R}_j^T\} =$ $\text{ave}\left\{\tilde{\Delta}(J,J')d(J)d^T(J')\right\}$	$\text{ave}\{g(J)g^T(J)\} =$ $\text{ave}\left\{\tilde{\delta}_{j,j'}\tilde{y}_j\tilde{y}_{j'}^T\right\}$
lift zonotope	$\text{ave}\{[-\tilde{y}_j,\tilde{y}_j]\}$	$\text{ave}\{[-d(J),d(J)]\}$
facet centre	$g(J)$	$\tilde{R}_j$, etc.