Robust ℓ_1 Approaches to Computing the Geometric Median and Principal and Independent Components

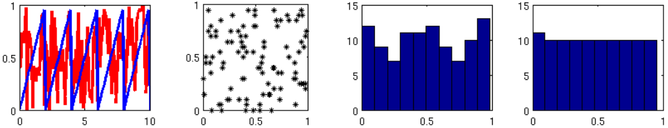
Stephen Keeling and Karl Kunisch
Institute for Mathematics and Scientific Computing
Karl Franzens University of Graz, Austria

Mathematical Image Processing Section
GAMM 2014, Erlangen
March 12, 2014

Graphical Demonstration of PCA/ICA

Sources Z, Measurements Y, sphered Y_s, separated X_s

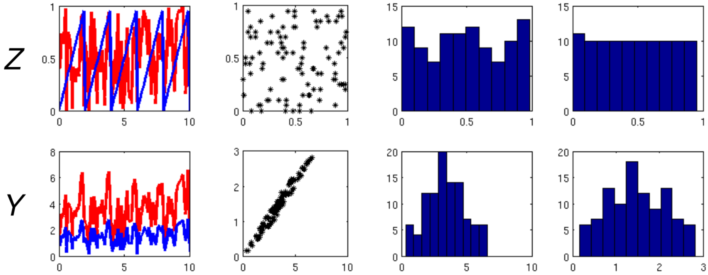
Z



Time Scatter Histograms

Graphical Demonstration of PCA/ICA

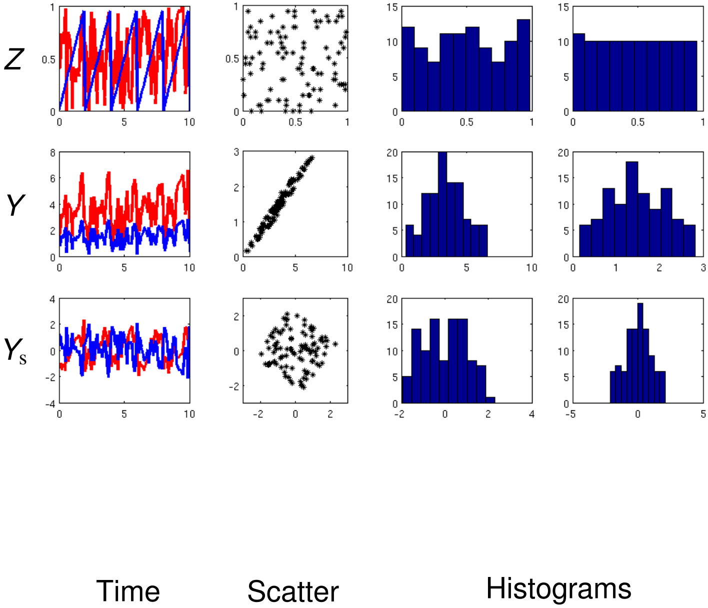
Sources Z, Measurements Y, sphered Y_s, separated X_s



Time Scatter Histograms

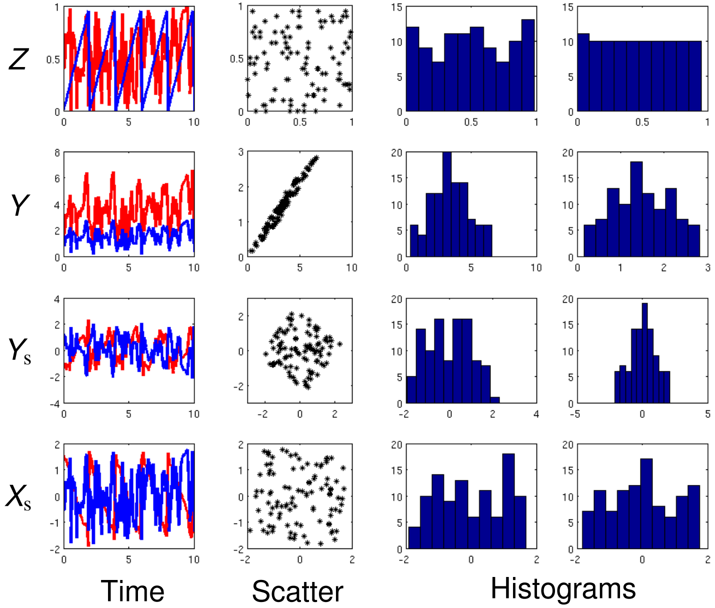
Graphical Demonstration of PCA/ICA

Sources Z, Measurements Y, sphered Y_s, separated X_s



Graphical Demonstration of PCA/ICA

Sources Z, Measurements Y, sphered Y_s, separated X_s



Formulation of PCA/ICA

- Rows of Z are unknown samples of sources which are independent and not Gauß distributed.

$$Z = \begin{bmatrix} z_1(t_1) & z_1(t_2) & \cdots & z_1(t_n) \\ \vdots & \vdots & & \vdots \\ z_m(t_1) & z_m(t_2) & \cdots & z_m(t_n) \end{bmatrix}$$

Formulation of PCA/ICA

- Rows of Z are unknown samples of sources which are independent and not Gauß distributed.

$$Z = \begin{bmatrix} z_1(t_1) & z_1(t_2) & \cdots & z_1(t_n) \\ \vdots & \vdots & & \vdots \\ z_m(t_1) & z_m(t_2) & \cdots & z_m(t_n) \end{bmatrix}$$

- Rows of Y are measured samples of unknown mixtures of the sources

$$Y = AZ$$

no longer independent and now more Gauß distributed.

Formulation of PCA/ICA

- Rows of Z are unknown samples of sources which are independent and not Gauß distributed.

$$Z = \begin{bmatrix} z_1(t_1) & z_1(t_2) & \cdots & z_1(t_n) \\ \vdots & \vdots & & \vdots \\ z_m(t_1) & z_m(t_2) & \cdots & z_m(t_n) \end{bmatrix}$$

- Rows of Y are measured samples of unknown mixtures of the sources

$$Y = AZ$$

  no longer independent and now more Gauß distributed.

- Goal is to undo the trend toward Gaußianity to recover the sources

$$X = WY$$

  with $W = U\Lambda^{-\frac{1}{2}}V^{\mathrm{T}} \approx A^{-1}$ but unavoidable ambiguities.

Formulation of PCA/ICA

Summary of Steps:

- Centering:

$$Y_{\mathrm{c}} = Y - \overline{Y}$$

Formulation of PCA/ICA

Summary of Steps:

- Centering:

$$Y_{\mathrm{c}} = Y - \overline{Y}$$

- Rotation:

$$K = \frac{1}{n} Y_{\mathrm{c}} Y_{\mathrm{c}}^{\mathrm{T}}, \quad KV = V\Lambda, \quad Y_{\mathrm{r}} = V^{\mathrm{T}} Y_{\mathrm{c}}$$

Formulation of PCA/ICA

Summary of Steps:

- Centering:

$$Y_{\mathrm{c}} = Y - \overline{Y}$$

- Rotation:

$$K = \frac{1}{n} Y_{\mathrm{c}} Y_{\mathrm{c}}^{\mathrm{T}}, \quad KV = V\Lambda, \quad Y_{\mathrm{r}} = V^{\mathrm{T}} Y_{\mathrm{c}}$$

- Scaling:

$$Y_{\mathrm{s}} = \Lambda^{-\frac{1}{2}} Y_{\mathrm{r}}$$

Formulation of PCA/ICA

Summary of Steps:

- Centering:

$$Y_{\mathrm{c}} = Y - \overline{Y}$$

- Rotation:

$$K = \frac{1}{n} Y_{\mathrm{c}} Y_{\mathrm{c}}^{\mathrm{T}}, \quad KV = V\Lambda, \quad Y_{\mathrm{r}} = V^{\mathrm{T}} Y_{\mathrm{c}}$$

- Scaling:

$$Y_{\mathrm{s}} = \Lambda^{-\frac{1}{2}} Y_{\mathrm{r}}$$

- Rotation:

$$X_{\mathrm{s}} = U Y_{\mathrm{s}}, \quad U^{\mathrm{T}} = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m\}$$

  where each $\boldsymbol{u}_k$ minimizes Gaußianity.

Formulation of PCA/ICA

Summary of Steps:

- Centering:

$$Y_{\mathrm{c}} = Y - \overline{Y}$$

- Rotation:

$$K = \frac{1}{n} Y_{\mathrm{c}} Y_{\mathrm{c}}^{\mathrm{T}}, \quad KV = V\Lambda, \quad Y_{\mathrm{r}} = V^{\mathrm{T}} Y_{\mathrm{c}}$$

- Scaling:

$$Y_{\mathrm{s}} = \Lambda^{-\frac{1}{2}} Y_{\mathrm{r}}$$

- Rotation:

$$X_{\mathrm{s}} = U Y_{\mathrm{s}}, \quad U^{\mathrm{T}} = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m\}$$

  where each $\boldsymbol{u}_k$ minimizes Gaußianity.

  For example, kurtosis

$$\mathcal{K}(\boldsymbol{x}) = M_4(\boldsymbol{x}) - 3M_2^2(\boldsymbol{x})$$

  satisfies $\mathcal{K}(\boldsymbol{n}) = 3\sigma^4 - 3\sigma^4 = 0$ for $\boldsymbol{n} \sim N(\mu, \sigma^2)$.

Formulation of PCA/ICA

Summary of Steps:

- Centering:

$$Y_{\mathrm{c}} = Y - \overline{Y}$$

- Rotation:

$$K = \frac{1}{n} Y_{\mathrm{c}} Y_{\mathrm{c}}^{\mathrm{T}}, \quad KV = V\Lambda, \quad Y_{\mathrm{r}} = V^{\mathrm{T}} Y_{\mathrm{c}}$$

- Scaling:

$$Y_{\mathrm{s}} = \Lambda^{-\frac{1}{2}} Y_{\mathrm{r}}$$

- Rotation:

$$X_{\mathrm{s}} = U Y_{\mathrm{s}}, \quad U^{\mathrm{T}} = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m\}$$

  where each $\boldsymbol{u}_k$ minimizes Gaußianity.

  For example, kurtosis

$$\mathcal{K}(\boldsymbol{x}) = M_4(\boldsymbol{x}) - 3M_2^2(\boldsymbol{x})$$

  satisfies $\mathcal{K}(\boldsymbol{n}) = 3\sigma^4 - 3\sigma^4 = 0$ for $\boldsymbol{n} \sim N(\mu, \sigma^2)$.

  So $J(\boldsymbol{u}) = [\mathcal{K}(Y_{\mathrm{s}}^{\mathrm{T}} \boldsymbol{u})]^2$ may be maximized with $\boldsymbol{u}_k^{\mathrm{T}} \boldsymbol{u}_l = \delta_{kl}$.

Formulation of PCA/ICA

(PCA) Let the data be so decomposed,

$$Y_{\mathrm{c}} = Y - \overline{Y}, \quad K = \frac{1}{n} Y_{\mathrm{c}} Y_{\mathrm{c}}^{\mathrm{T}}, \quad KV = V\Lambda, \quad Y_{\mathrm{s}} = \Lambda^{-\frac{1}{2}} V^{\mathrm{T}} Y_{\mathrm{c}}$$

Let $\Lambda = \mathrm{diag}\{\lambda_1, \ldots, \lambda_m\}$ with $\lambda_1 \geq \cdots \geq \lambda_m$. With $P \in \mathbb{R}^{r \times m}$, $r < m$, $P_{i,j} = \delta_{i,j}$, the data Y are so projected to its r strongest principal components,

$$Y \approx Y_P = \overline{Y} + V\Lambda^{\frac{1}{2}} P^{\mathrm{T}} P Y_{\mathrm{s}} = \overline{Y} + \frac{1}{n}(PY_{\mathrm{s}})^{\mathrm{T}}(PY_{\mathrm{s}})$$

Formulation of PCA/ICA

(PCA) Let the data be so decomposed,

$$Y_{\mathrm{c}} = Y - \overline{Y}, \quad K = \frac{1}{n} Y_{\mathrm{c}} Y_{\mathrm{c}}^{\mathrm{T}}, \quad KV = V\Lambda, \quad Y_{\mathrm{s}} = \Lambda^{-\frac{1}{2}} V^{\mathrm{T}} Y_{\mathrm{c}}$$

Let $\Lambda = \mathrm{diag}\{\lambda_1, \ldots, \lambda_m\}$ with $\lambda_1 \geq \cdots \geq \lambda_m$. With $P \in \mathbb{R}^{r \times m}$, $r < m$, $P_{i,j} = \delta_{i,j}$, the data Y are so projected to its r strongest principal components,

$$Y \approx Y_P = \overline{Y} + V\Lambda^{\frac{1}{2}} P^{\mathrm{T}} P Y_{\mathrm{s}} = \overline{Y} + \frac{1}{n}(P Y_{\mathrm{s}})^{\mathrm{T}}(P Y_{\mathrm{s}})$$

(ICA) Let the data be further so decomposed,

$$X_{\mathrm{s}} = U Y_{\mathrm{s}}$$

With $Q \in \mathbb{R}^{r \times m}$, $r < m$, $Q_{i,j} = \delta_{q_i,j}$, the data Y are so projected to the r independent components $\{q_1, \ldots, q_r\}$,

$$Y \approx Y_Q = \overline{Y} + V\Lambda^{\frac{1}{2}} U^{\mathrm{T}} Q^{\mathrm{T}} Q X_{\mathrm{s}} = \overline{Y} + \frac{1}{n}(Q X_{\mathrm{s}})^{\mathrm{T}}(Q X_{\mathrm{s}})$$

ℓ_2 and ℓ_1 Formulations of Centering

Given data $\boldsymbol{x} = \langle 0, 1, \ldots, 1 \rangle \in \mathbb{R}^m$,

ℓ_2 and ℓ_1 Formulations of Centering

Given data $\boldsymbol{x} = \langle 0, 1, \ldots, 1 \rangle \in \mathbb{R}^m$,

$$\begin{aligned} \text{mean}(\boldsymbol{x}) &= \operatorname*{argmin}_{\mu} \sum_{j=1}^{m} (\mu - x_j)^2 \\ &= \operatorname*{argmin}_{\mu} \left[(\mu - 0)^2 + (m-1)(\mu - 1)^2 \right] = (m-1)/m \end{aligned}$$

ℓ_2 and ℓ_1 Formulations of Centering

Given data $\boldsymbol{x} = \langle 0, 1, \ldots, 1 \rangle \in \mathbb{R}^m$,

$$\begin{aligned}
\text{mean}(\boldsymbol{x}) &= \operatorname*{argmin}_{\mu} \sum_{j=1}^{m} (\mu - x_j)^2 \\
&= \operatorname*{argmin}_{\mu} \left[(\mu - 0)^2 + (m-1)(\mu - 1)^2 \right] = (m-1)/m \\
\text{median}(\boldsymbol{x}) &= \operatorname*{argmin}_{\mu} \sum_{j=1}^{m} |\mu - x_j| \\
&= \operatorname*{argmin}_{0 \le \mu \le 1} [(\mu - 0) + (m-1)(1-\mu)] = 1 \quad \text{(robust!)}
\end{aligned}$$

ℓ_2 and ℓ_1 Formulations of Centering

Given data $\boldsymbol{x} = \langle 0, 1, \ldots, 1 \rangle \in \mathbb{R}^m$,

$$\begin{aligned}
\text{mean}(\boldsymbol{x}) &= \underset{\mu}{\operatorname{argmin}} \sum_{j=1}^{m} (\mu - x_j)^2 \\
&= \underset{\mu}{\operatorname{argmin}} \left[(\mu - 0)^2 + (m-1)(\mu - 1)^2 \right] = (m-1)/m \\
\text{median}(\boldsymbol{x}) &= \underset{\mu}{\operatorname{argmin}} \sum_{j=1}^{m} |\mu - x_j| \\
&= \underset{0 \le \mu \le 1}{\operatorname{argmin}} \left[(\mu - 0) + (m-1)(1 - \mu) \right] = 1 \quad \text{(robust!)}
\end{aligned}$$

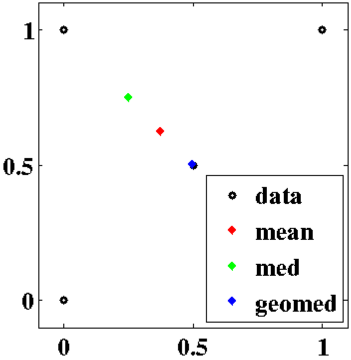
Generalization for higher dimensional data,
$Y = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n\} \in \mathbb{R}^{m \times n}$,

$$\text{geometric median}(Y) = \underset{\boldsymbol{\mu} \in \mathbb{R}^m}{\operatorname{argmin}} M(\boldsymbol{\mu}), \quad M(\boldsymbol{\mu}) = \sum_{j=1}^{n} \|\boldsymbol{\mu} - \boldsymbol{y}_j\|_{\ell_2}$$

ℓ_2 and ℓ_1 Formulations of Centering

The data:

$$Y = \begin{bmatrix} 0 & \frac{1}{2} & 1 & 0 \\ 0 & \frac{1}{2} & 1 & 1 \end{bmatrix}, \quad \bar{Y} = \underset{\mu \in \mathbb{R}^m}{\operatorname{argmin}} M(\mu)$$

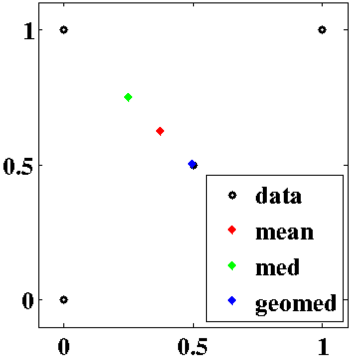


Data with Averages

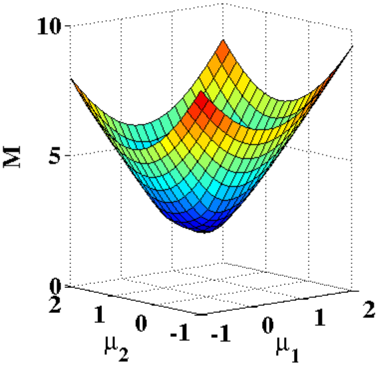
ℓ_2 and ℓ_1 Formulations of Centering

The data:

$$Y = \begin{bmatrix} 0 & \frac{1}{2} & 1 & 0 \\ 0 & \frac{1}{2} & 1 & 1 \end{bmatrix}, \quad \bar{Y} = \underset{\boldsymbol{\mu} \in \mathbb{R}^m}{\operatorname{argmin}} M(\boldsymbol{\mu})$$



Data with Averages



Geometric Median Landscape

ℓ_2 and ℓ_1 Formulations of Sphering

Sphered data:

$$Y_{\mathrm{s}} = \Lambda^{-\frac{1}{2}} V^{\mathrm{T}} Y_{\mathrm{c}}, \quad Y_{\mathrm{c}} = Y - \bar{Y} = \{\boldsymbol{y}_j^{\mathrm{c}}\}$$

where

$$\Lambda = \mathrm{diag}\{\lambda_i\}, \quad V = \{\hat{\boldsymbol{v}}_i\}, \qquad V^{\mathrm{T}} V = I.$$

ℓ_2 and ℓ_1 Formulations of Sphering

Sphered data:

$$Y_{\mathrm{s}} = \Lambda^{-\frac{1}{2}} V^{\mathrm{T}} Y_{\mathrm{c}}, \quad Y_{\mathrm{c}} = Y - \bar{Y} = \{\boldsymbol{y}_j^{\mathrm{c}}\}$$

where

$$\Lambda = \mathrm{diag}\{\lambda_i\}, \quad V = \{\hat{\boldsymbol{v}}_i\}, \qquad V^{\mathrm{T}} V = I.$$

ℓ_2 formulation:

$$\begin{aligned} \hat{\boldsymbol{v}}_1 &= \operatorname*{argmin}_{\|\hat{\boldsymbol{v}}\|_{\ell_2}=1} \tilde{H}(\hat{\boldsymbol{v}}) \\ \lambda_1 &= \|\hat{\boldsymbol{v}}^{\mathrm{T}} Y_{\mathrm{c}}/\sqrt{n}\|_{\ell_2}^2 \end{aligned} \quad \text{where} \quad \tilde{H}(\hat{\boldsymbol{v}}) = \sum_{j=1}^{n} \|(\hat{\boldsymbol{v}}\hat{\boldsymbol{v}}^{\mathrm{T}} - I)\boldsymbol{y}_j^{\mathrm{c}}\|_{\ell_2}^2$$

ℓ_2 and ℓ_1 Formulations of Sphering

Sphered data:

$$Y_{\mathrm{s}} = \Lambda^{-\frac{1}{2}} V^{\mathrm{T}} Y_{\mathrm{c}}, \quad Y_{\mathrm{c}} = Y - \bar{Y} = \{\boldsymbol{y}_j^{\mathrm{c}}\}$$

where

$$\Lambda = \mathrm{diag}\{\lambda_i\}, \quad V = \{\hat{\boldsymbol{v}}_i\}, \qquad V^{\mathrm{T}} V = I.$$

ℓ_2 formulation:

$$\begin{aligned} \hat{\boldsymbol{v}}_1 &= \operatorname*{argmin}_{\|\hat{\boldsymbol{v}}\|_{\ell_2}=1} \tilde{H}(\hat{\boldsymbol{v}}) \\ \lambda_1 &= \|\hat{\boldsymbol{v}}^{\mathrm{T}} Y_{\mathrm{c}}/\sqrt{n}\|_{\ell_2}^2 \end{aligned} \qquad \text{where} \quad \tilde{H}(\hat{\boldsymbol{v}}) = \sum_{j=1}^{n} \|(\hat{\boldsymbol{v}}\hat{\boldsymbol{v}}^{\mathrm{T}} - I)\boldsymbol{y}_j^{\mathrm{c}}\|_{\ell_2}^2$$

and similarly in orthogonal complements for $\{(\hat{\boldsymbol{v}}_i, \lambda_i)\}_{i>1}$.

ℓ_2 and ℓ_1 Formulations of Sphering

Sphered data:

$$Y_{\mathrm{s}} = \Lambda^{-\frac{1}{2}} V^{\mathrm{T}} Y_{\mathrm{c}}, \quad Y_{\mathrm{c}} = Y - \bar{Y} = \{\boldsymbol{y}_j^{\mathrm{c}}\}$$

where

$$\Lambda = \mathrm{diag}\{\lambda_i\}, \quad V = \{\hat{\boldsymbol{v}}_i\}, \qquad V^{\mathrm{T}} V = I.$$

ℓ_2 formulation:

$$\begin{aligned} \hat{\boldsymbol{v}}_1 &= \operatorname*{argmin}_{\|\hat{\boldsymbol{v}}\|_{\ell_2}=1} \tilde{H}(\hat{\boldsymbol{v}}) \\ \lambda_1 &= \|\hat{\boldsymbol{v}}^{\mathrm{T}} Y_{\mathrm{c}}/\sqrt{n}\|_{\ell_2}^2 \end{aligned} \quad \text{where} \quad \tilde{H}(\hat{\boldsymbol{v}}) = \sum_{j=1}^{n} \|(\hat{\boldsymbol{v}}\hat{\boldsymbol{v}}^{\mathrm{T}} - I)\boldsymbol{y}_j^{\mathrm{c}}\|_{\ell_2}^2$$

and similarly in orthogonal complements for $\{(\hat{\boldsymbol{v}}_i, \lambda_i)\}_{i>1}$.

ℓ_1 formulation:

$$\begin{aligned} \hat{\boldsymbol{v}}_1 &= \operatorname*{argmin}_{\|\hat{\boldsymbol{v}}\|_{\ell_2}=1} H(\hat{\boldsymbol{v}}) \\ \lambda_1 &= \|\hat{\boldsymbol{v}}^{\mathrm{T}} Y_{\mathrm{c}}/\sqrt{n}\|_{\ell_1}^2 \end{aligned} \quad \text{where} \quad H(\hat{\boldsymbol{v}}) = \sum_{j=1}^{n} \|(\hat{\boldsymbol{v}}\hat{\boldsymbol{v}}^{\mathrm{T}} - I)\boldsymbol{y}_j^{\mathrm{c}}\|_{\ell_2}$$

ℓ_2 and ℓ_1 Formulations of Sphering

Sphered data:

$$Y_\mathrm{s} = \Lambda^{-\frac{1}{2}} V^\mathrm{T} Y_\mathrm{c}, \quad Y_\mathrm{c} = Y - \bar{Y} = \{\boldsymbol{y}_j^\mathrm{c}\}$$

where

$$\Lambda = \mathrm{diag}\{\lambda_i\}, \quad V = \{\hat{\boldsymbol{v}}_i\}, \qquad V^\mathrm{T} V = I.$$

ℓ_2 formulation:

$$\begin{aligned} \hat{\boldsymbol{v}}_1 &= \underset{\|\hat{\boldsymbol{v}}\|_{\ell_2}=1}{\mathrm{argmin}}\, \tilde{H}(\hat{\boldsymbol{v}}) \\ \lambda_1 &= \|\hat{\boldsymbol{v}}^\mathrm{T} Y_\mathrm{c}/\sqrt{n}\|_{\ell_2}^2 \end{aligned} \quad \text{where} \quad \tilde{H}(\hat{\boldsymbol{v}}) = \sum_{j=1}^{n} \|(\hat{\boldsymbol{v}}\hat{\boldsymbol{v}}^\mathrm{T} - I)\boldsymbol{y}_j^\mathrm{c}\|_{\ell_2}^2$$

and similarly in orthogonal complements for $\{(\hat{\boldsymbol{v}}_i, \lambda_i)\}_{i>1}$.

ℓ_1 formulation:

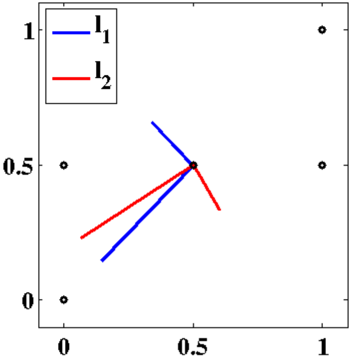
$$\begin{aligned} \hat{\boldsymbol{v}}_1 &= \underset{\|\hat{\boldsymbol{v}}\|_{\ell_2}=1}{\mathrm{argmin}}\, H(\hat{\boldsymbol{v}}) \\ \lambda_1 &= \|\hat{\boldsymbol{v}}^\mathrm{T} Y_\mathrm{c}/\sqrt{n}\|_{\ell_1}^2 \end{aligned} \quad \text{where} \quad H(\hat{\boldsymbol{v}}) = \sum_{j=1}^{n} \|(\hat{\boldsymbol{v}}\hat{\boldsymbol{v}}^\mathrm{T} - I)\boldsymbol{y}_j^\mathrm{c}\|_{\ell_2}$$

and similarly in orthogonal complements for $(\hat{\boldsymbol{v}}_i, \lambda_i)_{i>1}$.

ℓ_2 and ℓ_1 Formulations of Sphering

The data:

$$Y = \begin{bmatrix} 0 & \frac{1}{2} & 1 & 0 & 1 \\ 0 & \frac{1}{2} & 1 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \quad \bar{Y} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}, \quad Y_{\mathrm{c}} = Y - \bar{Y}, \quad Y_{\mathrm{s}} = \Lambda^{-\frac{1}{2}} V^{\mathrm{T}} Y_{\mathrm{c}}$$

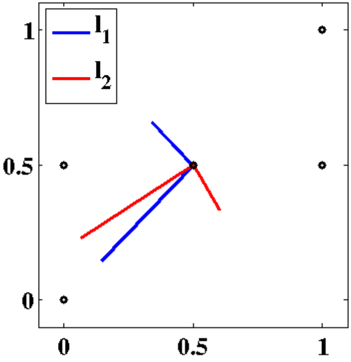


Data with Axes

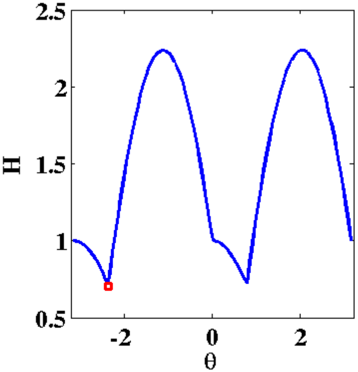
ℓ_2 and ℓ_1 Formulations of Sphering

The data:

$$Y = \begin{bmatrix} 0 & \frac{1}{2} & 1 & 0 & 1 \\ 0 & \frac{1}{2} & 1 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \quad \bar{Y} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}, \quad Y_{\mathrm{c}} = Y - \bar{Y}, \quad Y_{\mathrm{s}} = \Lambda^{-\frac{1}{2}} V^{\mathrm{T}} Y_{\mathrm{c}}$$



Data with Axes



Rotation Landscape

ℓ_2 and ℓ_1 Formulations of Independence

Independent data:

$$X_s = UY_s \quad \text{where} \quad U = \{\hat{\boldsymbol{u}}_i\}, \qquad U^T U = I.$$

ℓ_2 and ℓ_1 Formulations of Independence

Independent data:

$$X_\mathrm{s} = U Y_\mathrm{s} \quad \text{where} \quad U = \{\hat{\boldsymbol{u}}_i\}, \qquad U^\mathrm{T} U = I.$$

ℓ_2 formulation:

$$\hat{\boldsymbol{u}}_1 = \underset{\|\hat{\boldsymbol{u}}\|_{\ell_2}=1}{\operatorname{argmax}} K_\nu^2(\hat{\boldsymbol{u}}^\mathrm{T} Y_\mathrm{s}) \quad \text{where} \quad K_\nu(\nu) = 0 \quad \text{for} \quad \nu \sim N(0, \cdot)$$

ℓ_2 and ℓ_1 Formulations of Independence

Independent data:

$$X_s = U Y_s \quad \text{where} \quad U = \{\hat{\mathbf{u}}_i\}, \qquad U^T U = I.$$

ℓ_2 formulation:

$$\hat{\mathbf{u}}_1 = \underset{\|\hat{\mathbf{u}}\|_{\ell_2}=1}{\text{argmax}}\, K_\nu^2(\hat{\mathbf{u}}^T Y_s) \quad \text{where} \quad K_\nu(\nu) = 0 \quad \text{for} \quad \nu \sim N(0,\cdot), \text{ e.g.,}$$

$$\begin{aligned} K_p(\mathrm{x}) &= \mathrm{E}[|\mathrm{x}|^p] - \kappa_p \mathrm{E}[\mathrm{x}^2]^{p/2} \\ K_e(\mathrm{x}) &= \mathrm{E}[-\tfrac{1}{2}\mathrm{x}^2] - 1/\sqrt{1 + \mathrm{E}[\mathrm{x}^2]} \end{aligned} \qquad \kappa_p = \begin{cases} (p-1)!!, & p \text{ even} \\ \sqrt{\frac{2}{\pi}}(p-1)!!, & p \text{ odd} \end{cases}$$

ℓ_2 and ℓ_1 Formulations of Independence

Independent data:

$$X_\mathrm{s} = U Y_\mathrm{s} \quad \text{where} \quad U = \{\hat{\boldsymbol{u}}_i\}, \qquad U^\mathrm{T} U = I.$$

ℓ_2 formulation:

$$\hat{\boldsymbol{u}}_1 = \operatorname*{argmax}_{\|\hat{\boldsymbol{u}}\|_{\ell_2}=1} K_\nu^2(\hat{\boldsymbol{u}}^\mathrm{T} Y_\mathrm{s}) \quad \text{where} \quad K_\nu(\nu) = 0 \quad \text{for} \quad \nu \sim N(0,\cdot), \text{ e.g.,}$$

$$\begin{aligned} K_p(\mathrm{x}) &= \mathrm{E}[|\mathrm{x}|^p] - \kappa_p \mathrm{E}[\mathrm{x}^2]^{p/2} \\ K_\mathrm{e}(\mathrm{x}) &= \mathrm{E}[-\tfrac{1}{2}\mathrm{x}^2] - 1/\sqrt{1+\mathrm{E}[\mathrm{x}^2]} \end{aligned} \qquad \kappa_p = \begin{cases} (p-1)!!, & p \text{ even} \\ \sqrt{\frac{2}{\pi}}(p-1)!!, & p \text{ odd} \end{cases}$$

and similarly in orthogonal complements for $\{\hat{\boldsymbol{u}}_i\}_{i>1}$.

ℓ_2 and ℓ_1 Formulations of Independence

Independent data:

$$X_{\rm s} = U Y_{\rm s} \quad \text{where} \quad U = \{\hat{\boldsymbol{u}}_i\}, \qquad U^{\rm T} U = I.$$

ℓ_2 formulation:

$$\hat{\boldsymbol{u}}_1 = \underset{\|\hat{\boldsymbol{u}}\|_{\ell_2}=1}{\text{argmax}}\, K_\nu^2(\hat{\boldsymbol{u}}^{\rm T} Y_{\rm s}) \quad \text{where} \quad K_\nu(\nu) = 0 \quad \text{for} \quad \nu \sim N(0,\cdot), \text{ e.g.,}$$

$$\begin{aligned} K_p(\mathrm{x}) &= \mathrm{E}[|\mathrm{x}|^p] - \kappa_p \mathrm{E}[\mathrm{x}^2]^{p/2} \\ K_{\rm e}(\mathrm{x}) &= \mathrm{E}[-\tfrac{1}{2}\mathrm{x}^2] - 1/\sqrt{1+\mathrm{E}[\mathrm{x}^2]} \end{aligned} \qquad \kappa_p = \begin{cases} (p-1)!!, & p \text{ even} \\ \sqrt{\frac{2}{\pi}}(p-1)!!, & p \text{ odd} \end{cases}$$

and similarly in orthogonal complements for $\{\hat{\boldsymbol{u}}_i\}_{i>1}$.

ℓ_1 formulation:

$$\hat{\boldsymbol{u}}_1 = \underset{\|\hat{\boldsymbol{u}}\|_{\ell_2}=1}{\text{argmax}}\, G(\hat{\boldsymbol{u}}) \quad \text{where} \quad G(\hat{\boldsymbol{u}}) = \|\hat{\boldsymbol{u}}^{\rm T} Y_{\rm s}\|_{\ell_1}$$

ℓ_2 and ℓ_1 Formulations of Independence

Independent data:

$$X_\mathrm{s} = U Y_\mathrm{s} \quad \text{where} \quad U = \{\hat{\boldsymbol{u}}_i\}, \qquad U^\mathrm{T} U = I.$$

ℓ_2 formulation:

$$\hat{\boldsymbol{u}}_1 = \underset{\|\hat{\boldsymbol{u}}\|_{\ell_2}=1}{\operatorname{argmax}} K_\nu^2(\hat{\boldsymbol{u}}^\mathrm{T} Y_\mathrm{s}) \quad \text{where} \quad K_\nu(\nu) = 0 \quad \text{for} \quad \nu \sim N(0,\cdot), \text{ e.g.,}$$

$$\begin{aligned} K_p(\mathrm{x}) &= \mathrm{E}[|\mathrm{x}|^p] - \kappa_p \mathrm{E}[\mathrm{x}^2]^{p/2} \\ K_\mathrm{e}(\mathrm{x}) &= \mathrm{E}[-\tfrac{1}{2}\mathrm{x}^2] - 1/\sqrt{1 + \mathrm{E}[\mathrm{x}^2]} \end{aligned} \qquad \kappa_p = \begin{cases} (p-1)!!, & p \text{ even} \\ \sqrt{\frac{2}{\pi}}(p-1)!!, & p \text{ odd} \end{cases}$$

and similarly in orthogonal complements for $\{\hat{\boldsymbol{u}}_i\}_{i>1}$.

ℓ_1 formulation:

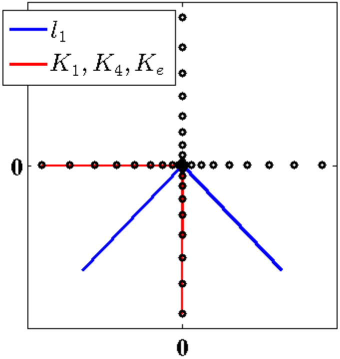
$$\hat{\boldsymbol{u}}_1 = \underset{\|\hat{\boldsymbol{u}}\|_{\ell_2}=1}{\operatorname{argmax}} G(\hat{\boldsymbol{u}}) \quad \text{where} \quad G(\hat{\boldsymbol{u}}) = \|\hat{\boldsymbol{u}}^\mathrm{T} Y_\mathrm{s}\|_{\ell_1}$$

and similarly in orthogonal complements for $\{\hat{\boldsymbol{u}}_i, \lambda_i\}_{i>1}$.

ℓ_2 and ℓ_1 Formulations of Independence

The data: $r = 20, k = 2, \sigma(t) = \text{sign}(t), i = -r, \ldots, r,$

$$Y = \begin{bmatrix} -1 & \ldots & \sigma(\frac{i}{r})|\frac{i}{r}|^k & \ldots & 1 & 0 & \ldots & & \ldots & 0 \\ 0 & \ldots & & \ldots & 0 & -1 & \ldots & \sigma(\frac{i}{r})|\frac{i}{r}|^k & \ldots & 1 \end{bmatrix}$$

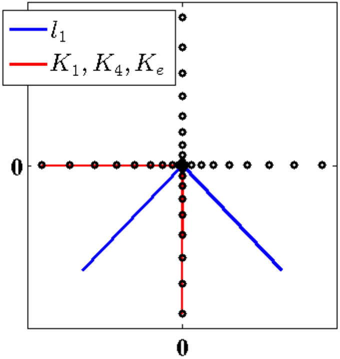


Data with Axes

ℓ_2 and ℓ_1 Formulations of Independence

The data: $r = 20, k = 2, \sigma(t) = \text{sign}(t), i = -r, \ldots, r,$

$$Y = \begin{bmatrix} -1 & \ldots & \sigma(\frac{i}{r})|\frac{i}{r}|^k & \ldots & 1 & 0 & \ldots & & \ldots & 0 \\ 0 & \ldots & & \ldots & 0 & -1 & \ldots & \sigma(\frac{i}{r})|\frac{i}{r}|^k & \ldots & 1 \end{bmatrix}$$



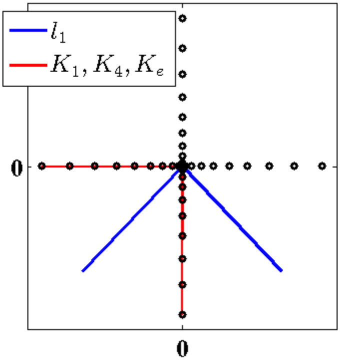
Data with Axes

$$\mathrm{P}(\mathrm{x}_1(\theta) = \alpha \text{ and } \mathrm{x}_2(\theta) = \beta) \overset{?}{=} \mathrm{P}(\mathrm{x}_1(\theta) = \alpha) \cdot \mathrm{P}(\mathrm{x}_2(\theta) = \beta), \quad \forall \alpha, \beta \in \mathbb{R}$$

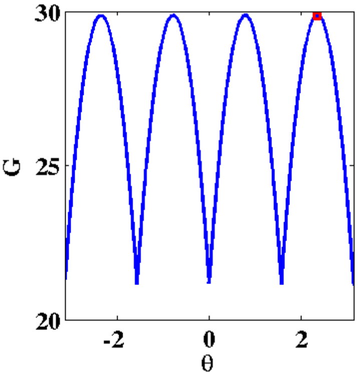
ℓ_2 and ℓ_1 Formulations of Independence

The data: $r = 20, k = 2, \sigma(t) = \text{sign}(t), i = -r, \ldots, r,$

$$Y = \begin{bmatrix} -1 & \ldots & \sigma(\frac{i}{r})|\frac{i}{r}|^k & \ldots & 1 & 0 & \ldots & & \ldots & 0 \\ 0 & \ldots & & \ldots & 0 & -1 & \ldots & \sigma(\frac{i}{r})|\frac{i}{r}|^k & \ldots & 1 \end{bmatrix}$$



Data with Axes



Rotation Landscape

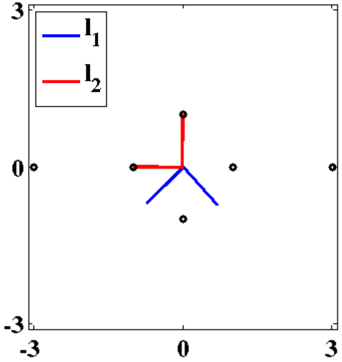
$$\mathrm{P}(\mathrm{x}_1(\theta) = \alpha \text{ and } \mathrm{x}_2(\theta) = \beta) \stackrel{?}{=} \mathrm{P}(\mathrm{x}_1(\theta) = \alpha)\cdot\mathrm{P}(\mathrm{x}_2(\theta) = \beta), \quad \forall \alpha, \beta \in \mathbb{R}$$

ℓ_2 and ℓ_1 Formulations of Independence

The data:

$$Y_x = \begin{bmatrix} +1 & +1 & +1 & +1 & +1 & -1 & -1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$Y_y = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} Y_x, \quad Y_{\mathrm{o}} = \begin{bmatrix} 3 & -3 \\ 0 & 0 \end{bmatrix}, \quad Y = \begin{bmatrix} Y_x & Y_y & Y_{\mathrm{o}} \end{bmatrix}$$



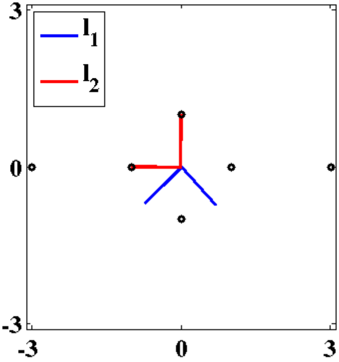
Data with Axes

ℓ_2 and ℓ_1 Formulations of Independence

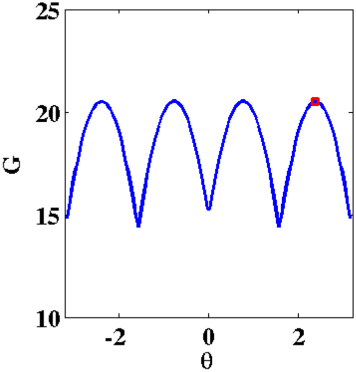
The data:

$$Y_x = \begin{bmatrix} +1 & +1 & +1 & +1 & +1 & -1 & -1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$Y_y = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} Y_x, \quad Y_\text{o} = \begin{bmatrix} 3 & -3 \\ 0 & 0 \end{bmatrix}, \quad Y = \begin{bmatrix} Y_x & Y_y & Y_\text{o} \end{bmatrix}$$



Data with Axes



Rotation Landscape

Numerical Optimization and Convergence

Geometric Median: $Y = \{\boldsymbol{y}_j\}$, $\boldsymbol{y}_j, \boldsymbol{d}_{j,l} \in \mathbb{R}^m$, $\boldsymbol{\mu}_l \in \mathbb{R}^m$, $\tau > 0$,

$$\boldsymbol{d}_{j,l+1} = \frac{\boldsymbol{d}_{j,l} + \tau(\boldsymbol{\mu}_l - \boldsymbol{y}_j)}{1 + \tau\|\boldsymbol{\mu}_l - \boldsymbol{y}_j\|_{\ell_2}}, \quad j = 1, \ldots, n$$

$$\boldsymbol{\mu}_{l+1} = \sum_{j=1}^{n} \frac{(\boldsymbol{d}_{j,l} - \tau \boldsymbol{y}_j)}{1 + \tau\|\boldsymbol{\mu}_l - \boldsymbol{y}_j\|_{\ell_2}} \Big/ \sum_{j=1}^{n} \frac{-\tau}{1 + \tau\|\boldsymbol{\mu}_l - \boldsymbol{y}_j\|_{\ell_2}}.$$

Numerical Optimization and Convergence

Geometric Median: $Y = \{\boldsymbol{y}_j\}$, $\boldsymbol{y}_j, \boldsymbol{d}_{j,l} \in \mathbb{R}^m$, $\boldsymbol{\mu}_l \in \mathbb{R}^m$, $\tau > 0$,

$$\boldsymbol{d}_{j,l+1} = \frac{\boldsymbol{d}_{j,l} + \tau(\boldsymbol{\mu}_l - \boldsymbol{y}_j)}{1 + \tau\|\boldsymbol{\mu}_l - \boldsymbol{y}_j\|_{\ell_2}}, \quad j = 1, \dots, n$$

$$\boldsymbol{\mu}_{l+1} = \sum_{j=1}^{n} \frac{(\boldsymbol{d}_{j,l} - \tau\boldsymbol{y}_j)}{1 + \tau\|\boldsymbol{\mu}_l - \boldsymbol{y}_j\|_{\ell_2}} \Big/ \sum_{j=1}^{n} \frac{-\tau}{1 + \tau\|\boldsymbol{\mu}_l - \boldsymbol{y}_j\|_{\ell_2}}.$$

Theorem: Suppose the columns $\{\boldsymbol{y}_j\}_{j=1}^n = Y \in \mathbb{R}^{m\times n}$ are not colinear. Let $\{\boldsymbol{d}_1^\star, \dots, \boldsymbol{d}_n^\star, \boldsymbol{\mu}^\star\}$ satisfy the necessary optimality condition for the geometric median merit function,

$$\boldsymbol{\mu}^\star - \boldsymbol{y}_j = \|\boldsymbol{\mu}^\star - \boldsymbol{y}_j\|_{\ell_2} \boldsymbol{d}_j^\star, \quad \|\boldsymbol{d}_j^\star\|_{\ell_2} \le 1, \quad j = 1, \dots, n, \quad \sum_{j=1}^{n} \boldsymbol{d}_j^\star = \boldsymbol{0}.$$

with $\boldsymbol{\mu}^\star \notin \{\boldsymbol{y}_j\}_{j=1}^n$. Then $\{\boldsymbol{d}_1^\star, \dots, \boldsymbol{d}_n^\star, \boldsymbol{\mu}^\star\}$ is a fixed point for the iteration which is locally asymptotically stable for τ large enough.

Numerical Optimization and Convergence

Given previous calculations, $Y_1 = Y_c, S_1 = \mathbb{R}^m$

$$V_k = \{\hat{\boldsymbol{v}}_i\}_{i=1}^{k}, \quad Y_k = (I - V_{k-1}V_{k-1}^{\mathrm{T}})Y_{\mathrm{c}}, \quad S_k = \mathcal{R}(V_{k-1})^{\perp}.$$

Sphering: $Y_k = \{\boldsymbol{y}_{j,k}\}$, $\boldsymbol{y}_{j,k}, \boldsymbol{d}_{j,l} \in S_k$, $\hat{\mathbf{v}}_l \in S_k$, $\tau, \rho > 0$

$$\boldsymbol{d}_{j,l+1} = \frac{(\hat{\mathbf{v}}_l\hat{\mathbf{v}}_l^{\mathrm{T}} - I)(\tau\boldsymbol{y}_{j,k} - \boldsymbol{d}_{j,l})}{1 + \tau\|(\hat{\mathbf{v}}_l\hat{\mathbf{v}}_l^{\mathrm{T}} - I)\boldsymbol{y}_{j,k}\|_{\ell_2}}, \quad j = 1, \ldots, n$$

$$\hat{\mathbf{v}}_{l+1} = \frac{\mathbf{v}_{l+1}}{\|\mathbf{v}_{l+1}\|_{\ell_2}} \quad \text{with} \quad \mathbf{v}_{l+1} = \hat{\mathbf{v}}_l - \rho\sum_{j=1}^{n}\frac{(\hat{\mathbf{v}}_l^{\mathrm{T}}\boldsymbol{y}_{j,k})(\tau\boldsymbol{y}_{j,k} - \boldsymbol{d}_{j,l})}{1 + \tau\|(\hat{\mathbf{v}}_l\hat{\mathbf{v}}_l^{\mathrm{T}} - I)\boldsymbol{y}_{j,k}\|_{\ell_2}}.$$

Numerical Optimization and Convergence

Given previous calculations, $Y_1 = Y_{\rm c}, S_1 = \mathbb{R}^m$

$$V_k = \{\hat{\boldsymbol{v}}_i\}_{i=1}^k, \quad Y_k = (I - V_{k-1}V_{k-1}^{\rm T})Y_{\rm c}, \quad S_k = \mathcal{R}(V_{k-1})^\perp.$$

Sphering: $Y_k = \{\boldsymbol{y}_{j,k}\}$, $\boldsymbol{y}_{j,k}, \boldsymbol{d}_{j,l} \in S_k$, $\hat{\mathbf{v}}_l \in S_k$, $\tau, \rho > 0$

$$\boldsymbol{d}_{j,l+1} = \frac{(\hat{\mathbf{v}}_l\hat{\mathbf{v}}_l^{\rm T} - I)(\tau\boldsymbol{y}_{j,k} - \boldsymbol{d}_{j,l})}{1 + \tau\|(\hat{\mathbf{v}}_l\hat{\mathbf{v}}_l^{\rm T} - I)\boldsymbol{y}_{j,k}\|_{\ell_2}}, \quad j = 1, \dots, n$$

$$\hat{\mathbf{v}}_{l+1} = \frac{\mathbf{v}_{l+1}}{\|\mathbf{v}_{l+1}\|_{\ell_2}} \quad \text{with} \quad \mathbf{v}_{l+1} = \hat{\mathbf{v}}_l - \rho\sum_{j=1}^n \frac{(\hat{\mathbf{v}}_l^{\rm T}\boldsymbol{y}_{j,k})(\tau\boldsymbol{y}_{j,k} - \boldsymbol{d}_{j,l})}{1 + \tau\|(\hat{\mathbf{v}}_l\hat{\mathbf{v}}_l^{\rm T} - I)\boldsymbol{y}_{j,k}\|_{\ell_2}}.$$

Theorem: Let $\{\boldsymbol{d}_1^\star, \dots, \boldsymbol{d}_n^\star, \hat{\boldsymbol{v}}^\star\}$ satisfy for $1 \le j \le n$, $\|\boldsymbol{d}_j^\star\|_{\ell_2} \le 1$,

$$(\hat{\boldsymbol{v}}^\star\hat{\boldsymbol{v}}^{\star\rm T} - I)\boldsymbol{y}_{j,k} = \|(\hat{\boldsymbol{v}}^\star\hat{\boldsymbol{v}}^{\star\rm T} - I)\boldsymbol{y}_{j,k}\|_{\ell_2}\boldsymbol{d}_j^\star, \quad \hat{\boldsymbol{v}}^{\star\rm T}\boldsymbol{d}_j^\star = 0, \quad \sum_{j=1}^n(\hat{\boldsymbol{v}}^{\star\rm T}\boldsymbol{y}_{j,k})\boldsymbol{d}_j^\star = \boldsymbol{0}$$

with $\hat{\boldsymbol{v}}^\star \in S_k$, $\|\hat{\boldsymbol{v}}^\star\|_{\ell_2} = 1$, and suppose

$$\hat{\boldsymbol{v}}^{\star\rm T}\boldsymbol{y}_{j,k} \neq 0, \quad (\hat{\boldsymbol{v}}^\star\hat{\boldsymbol{v}}^{\star\rm T} - I)\boldsymbol{y}_{j,k} \neq \boldsymbol{0}, \quad j = 1, \dots, n.$$

Then $\{\boldsymbol{d}_1^\star, \dots, \boldsymbol{d}_n^\star, \hat{\boldsymbol{v}}^\star\}$ is a fixed point of the iteration which is locally asymptotically stable for τ large enough and ρ small enough.

Numerical Optimization and Convergence

Given previous calculations, $Y_1 = Y_s,\, T_1 = \mathbb{R}^m$

$$U_l = \{\hat{\boldsymbol{u}}_i\}_{i=1}^{l}, \quad Y_l = (I - U_{l-1}^{\mathrm{T}} U_{k-1}) Y_{\mathrm{s}}, \quad T_l = \mathcal{R}(U_{l-1}^{\mathrm{T}})^{\perp}.$$

Independence: $\hat{\mathbf{u}}_k \in T_l$, $\tau > 0$,

$$\hat{\mathbf{u}}_{k+1} = \frac{\mathbf{u}_{k+1}}{\|\mathbf{u}_{k+1}\|_{\ell_2}} \quad \text{with} \quad \mathbf{u}_{k+1} = \hat{\mathbf{u}}_k + \tau \left[Y_l\, \sigma(Y_l^{\mathrm{T}} \hat{\mathbf{u}}_k) - \hat{\mathbf{u}}_k \|\hat{\mathbf{u}}_k^{\mathrm{T}} Y_l\|_{\ell_1} \right]$$

$$\text{where} \quad \begin{cases} \sigma(t) & = & \mathrm{sign}(t), & t \in \mathbb{R} \\ \sigma(\boldsymbol{v}) & = & \{\sigma(v_j)\}_{j=1}^{n}, & \boldsymbol{v} = \{v_j\}_{j=1}^{n} \in \mathbb{R}^m. \end{cases}$$

Numerical Optimization and Convergence

Given previous calculations, $Y_1 = Y_{\mathrm{s}}$, $T_1 = \mathbb{R}^m$

$$U_l = \{\hat{\boldsymbol{u}}_i\}_{i=1}^{l}, \quad Y_l = (I - U_{l-1}^{\mathrm{T}} U_{k-1}) Y_{\mathrm{s}}, \quad T_l = \mathcal{R}(U_{l-1}^{\mathrm{T}})^{\perp}.$$

Independence: $\hat{\mathbf{u}}_k \in T_l$, $\tau > 0$,

$$\hat{\mathbf{u}}_{k+1} = \frac{\mathbf{u}_{k+1}}{\|\mathbf{u}_{k+1}\|_{\ell_2}} \quad \text{with} \quad \mathbf{u}_{k+1} = \hat{\mathbf{u}}_k + \tau \left[Y_l\, \sigma(Y_l^{\mathrm{T}} \hat{\mathbf{u}}_k) - \hat{\mathbf{u}}_k \|\hat{\mathbf{u}}_k^{\mathrm{T}} Y_l\|_{\ell_1} \right]$$

$$\text{where} \quad \begin{cases} \sigma(t) & = & \mathrm{sign}(t), & t \in \mathbb{R} \\ \sigma(\boldsymbol{v}) & = & \{\sigma(v_j)\}_{j=1}^{n}, & \boldsymbol{v} = \{v_j\}_{j=1}^{n} \in \mathbb{R}^m. \end{cases}$$

Theorem: Let $\hat{\boldsymbol{u}}^\star \in T_l$ with $\|\hat{\boldsymbol{u}}^\star\|_{\ell_2} = 1$ satisfy

$$Y_l \sigma(Y_l^{\mathrm{T}} \hat{\boldsymbol{u}}^\star) = \|Y_l^{\mathrm{T}} \hat{\boldsymbol{u}}^\star\|_{\ell_1} \hat{\boldsymbol{u}}^\star$$

with $\mathcal{S} = \{j : \hat{\mathbf{e}}_j^{\mathrm{T}} Y_l^{\mathrm{T}} \hat{\boldsymbol{u}}^\star = 0\} = \emptyset$. Then $\hat{\boldsymbol{u}}^\star$ is a fixed point of the iteration which is locally asymptotically stable for τ small enough.

Application to DCE-MRI sequences

For each time $t = 1, \ldots, T$, let the matrix of pixel values,

$$I(t) = \{I_{i,j}(t)\}_{1 \leq i,j \leq N}$$

be an image in a DCE-MRI sequence.

Application to DCE-MRI sequences

For each time $t = 1, \ldots, T$, let the matrix of pixel values,

$$I(t) = \{I_{i,j}(t)\}_{1 \leq i,j \leq N}$$

be an image in a DCE-MRI sequence. Eliminate noise? motion?

Application to DCE-MRI sequences

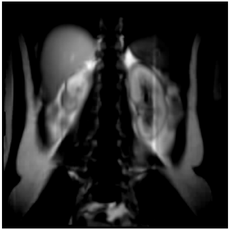
For each time $t = 1, \ldots, T$, let the matrix of pixel values,

$$I(t) = \{I_{i,j}(t)\}_{1 \le i,j \le N}$$

be an image in a DCE-MRI sequence. Eliminate noise? motion?

With $n = T = 134$ and $m = N^2 = 400^2$ the images are represented as long vectors, and PCA/ICA is carried out with

$$Y = \left\{ \langle I_{1,1}(t), \ldots, I_{N,1}(t), I_{1,2}(t), \ldots, I_{N,2}(t), \ldots, I_{1,N}(t), \ldots, I_{N,N}(t) \rangle^{\mathrm{T}} : t = 1, \ldots, T \right\}$$



Application to DCE-MRI sequences

For each time $t = 1, \ldots, T$, let the matrix of pixel values,

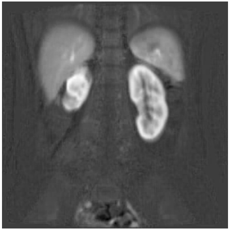
$$I(t) = \{I_{i,j}(t)\}_{1 \leq i,j \leq N}$$

be an image in a DCE-MRI sequence. Eliminate noise? motion?

With $n = T = 134$ and $m = N^2 = 400^2$ the images are represented as long vectors, and PCA/ICA is carried out with

$$Y = \left\{ \langle I_{1,1}(t), \ldots, I_{N,1}(t), I_{1,2}(t), \ldots, I_{N,2}(t), \ldots, I_{1,N}(t), \ldots, I_{N,N}(t) \rangle^{\mathrm{T}} : t = 1, \ldots, T \right\}$$

To the left is the first column of V (displayed as an image),



Application to DCE-MRI sequences

For each time $t = 1, \ldots, T$, let the matrix of pixel values,

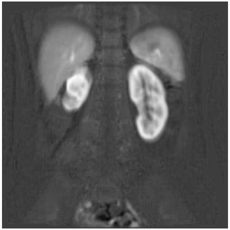
$$I(t) = \{I_{i,j}(t)\}_{1 \leq i,j \leq N}$$

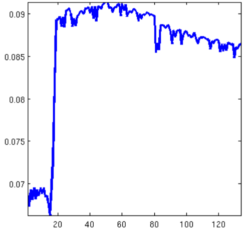
be an image in a DCE-MRI sequence. Eliminate noise? motion?

With $n = T = 134$ and $m = N^2 = 400^2$ the images are represented as long vectors, and PCA/ICA is carried out with

$$Y = \left\{ \langle I_{1,1}(t), \ldots, I_{N,1}(t), I_{1,2}(t), \ldots, I_{N,2}(t), \ldots, I_{1,N}(t), \ldots, I_{N,N}(t) \rangle^{\mathrm{T}} : t = 1, \ldots, T \right\}$$

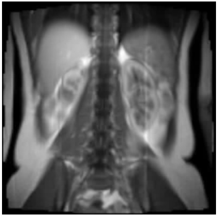
To the left is the first column of V (displayed as an image), and to the right is the first row of Y_{s}.



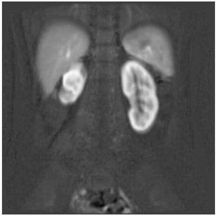


Application to DCE-MRI sequences

Basis: $\hat{\boldsymbol{v}}_i = V\hat{\boldsymbol{e}}_i, \quad i = 1, 2, \quad (\hat{\boldsymbol{e}}_i)_j = \delta_{ij},$



$\bar{Y}$

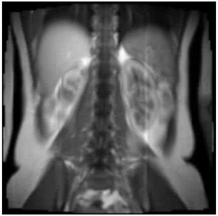


$\hat{\boldsymbol{v}}_1$

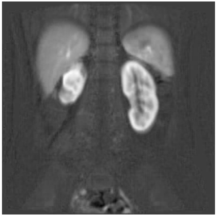
$\hat{\boldsymbol{v}}_2$

Application to DCE-MRI sequences

Basis: $\hat{\boldsymbol{v}}_i = V\hat{\boldsymbol{e}}_i, \quad i = 1, 2, \quad (\hat{\boldsymbol{e}}_i)_j = \delta_{ij},$



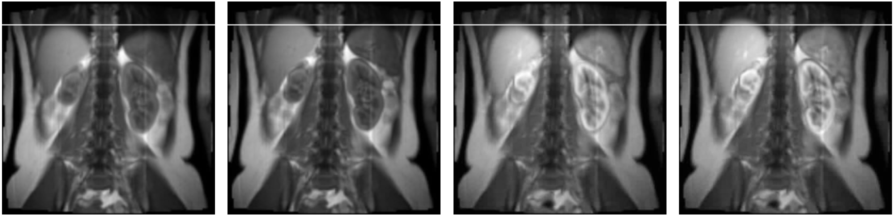
$\bar{Y}$



$\hat{\boldsymbol{v}}_1$

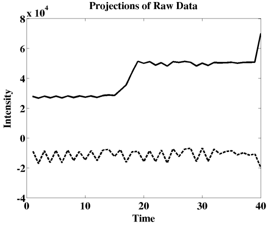
$\hat{\boldsymbol{v}}_2$

e.g., $\bar{Y} + V\Lambda^{\frac{1}{2}}\boldsymbol{e}_{ij}, \quad i, j = \pm 1, \quad \boldsymbol{e}_{ij} = (i, j, 0, \ldots, 0)^{\mathrm{T}}$

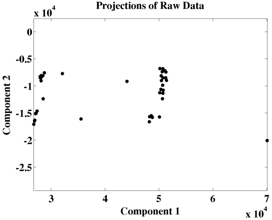


Application to DCE-MRI sequences

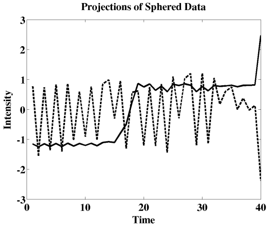
Decomposition with ℓ_2 PCA/ICA and outlier at $t = 40$:



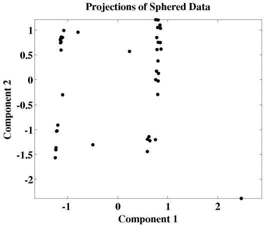
$\hat{\boldsymbol{v}}_i^{\mathrm{T}} Y$ temporal



$\hat{\boldsymbol{v}}_i^{\mathrm{T}} Y$ phase



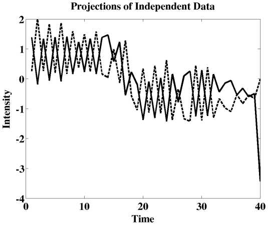
$\hat{\boldsymbol{v}}_i^{\mathrm{T}} Y_{\mathrm{s}}$ temporal

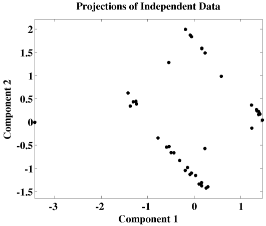


$\hat{\boldsymbol{v}}_i^{\mathrm{T}} Y_{\mathrm{s}}$ phase

Application to DCE-MRI sequences

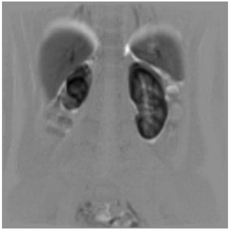
Decomposition with ℓ_2 PCA/ICA and outlier at $t = 40$:





$\hat{\boldsymbol{v}}_i^{\mathrm{T}} X_{\mathrm{s}}$ temporal

$\hat{\boldsymbol{v}}_i^{\mathrm{T}} X_{\mathrm{s}}$ phase



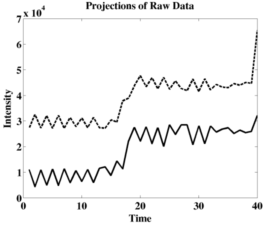
X_1

X_2

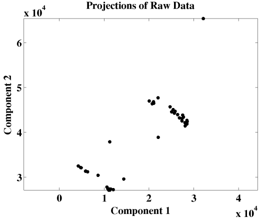
$$X_i = V\Lambda^{\frac{1}{2}} U^{\mathrm{T}} Q_i X_{\mathrm{s}},\ Q_i = \mathrm{diag}\{\hat{\boldsymbol{e}}_i\},\ i = 1, 2$$

Application to DCE-MRI sequences

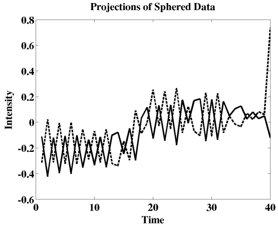
Decomposition with ℓ_1 PCA/ICA and outlier at $t = 40$:



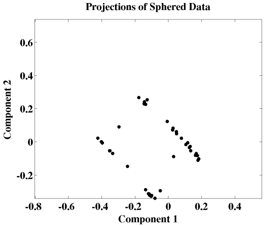
$\hat{\boldsymbol{v}}_i^{\mathrm{T}} Y$ temporal



$\hat{\boldsymbol{v}}_i^{\mathrm{T}} Y$ phase



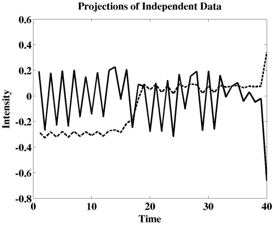
$\hat{\boldsymbol{v}}_i^{\mathrm{T}} Y_{\mathrm{s}}$ temporal



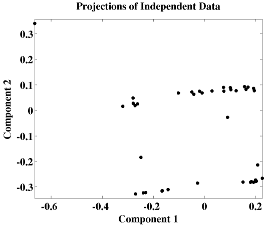
$\hat{\boldsymbol{v}}_i^{\mathrm{T}} Y_{\mathrm{s}}$ phase

Application to DCE-MRI sequences

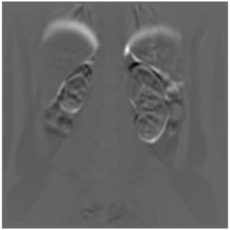
Decomposition with ℓ_1 PCA/ICA and outlier at $t = 40$:



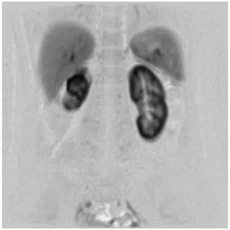
$\hat{\boldsymbol{v}}_i^{\mathrm{T}} X_{\mathrm{s}}$ temporal



$\hat{\boldsymbol{v}}_i^{\mathrm{T}} X_{\mathrm{s}}$ phase



X_1



X_2

$$X_i = V \Lambda^{\frac{1}{2}} U^{\mathrm{T}} Q_i X_{\mathrm{s}}, \quad Q_i = \mathrm{diag}\{\hat{\boldsymbol{e}}_i\}, \quad i = 1, 2$$

Application to DCE-MRI sequences

With motion component removed:

Original | Iterative Template | PCA/ICA

where Iterative Template sequence is given by iterating:

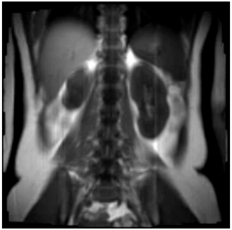
$$
\begin{aligned}
I_{\mathrm{time}}(\boldsymbol{x},\cdot) &= \operatorname*{argmin}_{I} \int_0^T \left\{ |I - I_{\mathrm{stat}}|^2 + \alpha |\partial_t I|^2 \right\} dt \quad (\text{init:} \quad I_{\mathrm{stat}} \leftarrow I_{\mathrm{orig}}) \\
\boldsymbol{u}_{\mathrm{space}}(\cdot, t) &= \operatorname*{argmin}_{\boldsymbol{u}} \int_\Omega \left\{ |I_{\mathrm{orig}} \circ (\mathrm{Id} + \boldsymbol{u}) - I_{\mathrm{time}}|^2 + \mu |\nabla \boldsymbol{u}^{\mathrm{T}} + \nabla \boldsymbol{u}|^2 \right\} d\boldsymbol{x} \\
I_{\mathrm{stat}} &\leftarrow I_{\mathrm{orig}} \circ (\mathrm{Id} + \boldsymbol{u}_{\mathrm{space}})
\end{aligned}
$$

Application to DCE-MRI sequences

With motion component removed:

Original

Iterative Template



PCA/ICA

where Iterative Template sequence is given by iterating:

$$
\begin{aligned}
I_{\text{time}}(\boldsymbol{x}, \cdot) &= \underset{I}{\operatorname{argmin}} \int_0^T \left\{ |I - I_{\text{stat}}|^2 + \alpha |\partial_t I|^2 \right\} dt \quad (\text{init:} \quad I_{\text{stat}} \leftarrow I_{\text{orig}}) \\
\boldsymbol{u}_{\text{space}}(\cdot, t) &= \underset{\boldsymbol{u}}{\operatorname{argmin}} \int_\Omega \left\{ |I_{\text{orig}} \circ (\text{Id} + \boldsymbol{u}) - I_{\text{time}}|^2 + \mu |\nabla \boldsymbol{u}^{\mathrm{T}} + \nabla \boldsymbol{u}|^2 \right\} d\boldsymbol{x} \\
I_{\text{stat}} &\leftarrow I_{\text{orig}} \circ (\text{Id} + \boldsymbol{u}_{\text{space}})
\end{aligned}
$$

References

- J.P. Brooks, J.H. Dulá and E.L. Boone, *A pure L_1-norm principal comoponent analysis*, Computational Statistics and Data Analysis, Vol. 61, pp. 83 – 98, May 2013.
- A. Hyvärinen, J. Karhunen and E. Oja, *Independent Component Analysis*, New York, Wiley, 2001.
- I.T. Jolliffe, *Principal Component Analysis*, New York, Springer, 1986.
- N. Kwak, *Principal component analysis based on $L1$-norm maximization*, IEEE Transactions on Pattern Analysis and Machine Intelligence Vol. 30, pp. 1672 – 1680, 2008.
- E. Weiszfeld and F. Plastria, *On the point for which the sum of the distances to n given points is minimum*, Ann. Oper. Res., Vol. 167, pp. 7 – 41, 2009. (1937)
- V. Zarzoso, *Robust Independent Component Analysis by Iterative Maximization of the Kurtosis Contrast with Algebraic Optimal Step Size*, IEEE Trans. on Neural Networks, Vol. 21, No. 2, pp. 248 – 261, Feb. 2010.

References

- J.P. Brooks, J.H. Dulá and E.L. Boone, *A pure L_1-norm principal comoponent analysis*, Computational Statistics and Data Analysis, Vol. 61, pp. 83 – 98, May 2013.
- A. Hyvärinen, J. Karhunen and E. Oja, *Independent Component Analysis*, New York, Wiley, 2001.
- I.T. Jolliffe, *Principal Component Analysis*, New York, Springer, 1986.
- N. Kwak, *Principal component analysis based on L1-norm maximization*, IEEE Transactions on Pattern Analysis and Machine Intelligence Vol. 30, pp. 1672 – 1680, 2008.
- E. Weiszfeld and F. Plastria, *On the point for which the sum of the distances to n given points is minimum*, Ann. Oper. Res., Vol. 167, pp. 7 – 41, 2009. (1937)
- V. Zarzoso, *Robust Independent Component Analysis by Iterative Maximization of the Kurtosis Contrast with Algebraic Optimal Step Size*, IEEE Trans. on Neural Networks, Vol. 21, No. 2, pp. 248 – 261, Feb. 2010.

Thank You!