Robust Clustering for Satellite Images Time-Series

Antoine Collas[1]
Jean-Philippe Ovarlez[2], Guillaume Ginolhac[3], Chengfang Ren[4], Arnaud Breloy[5]

[1]2ème année - SONDRA
[2]Directeur de thèse - ONERA/SONDRA
[3]Directeur de thèse - LISTIC
[4]Co-encadrant - SONDRA
[5]Co-encadrant - LEME

14 janvier 2020

ONERA
THE FRENCH AEROSPACE LAB

retour sur innovation

Plan

Context of the PhD.

In the last few years many images have been taken from the earth with different technologies (SAR, multi-spectral/hyperspectral imaging, ...).

Problematics

The objective is to develop clustering methods specific to these new data. More particularly we focus on 2 specific topics :

- Change detection.
- Semantic segmentation.

Figure – Raw image.

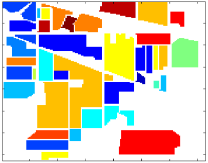

Figure – Segmented image. One color = one class (grass, woods, ...).

Example of change detection.

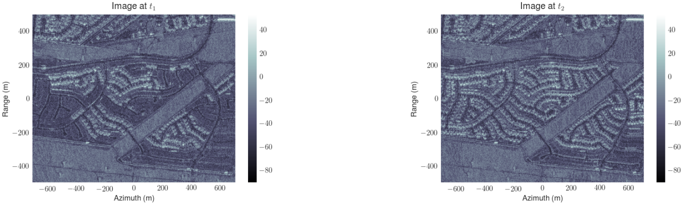

Figure – Time series

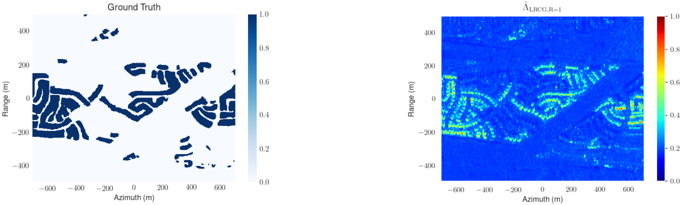

Figure – Ground truth vs prediction from (Mian et al., 2020).

Clustering pipeline.

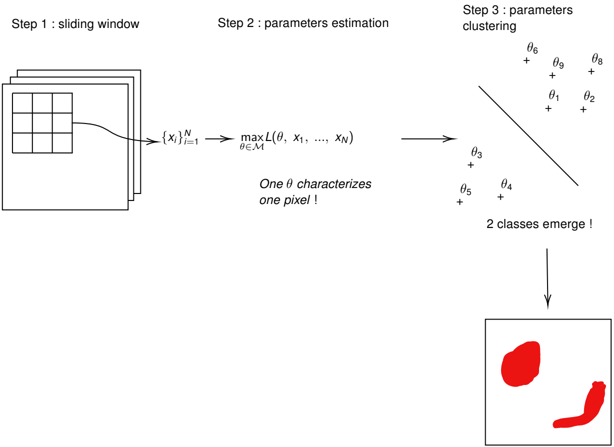

Figure – Clustering pipeline.

Objectives for parameter estimation.

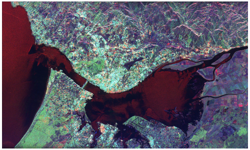

Figure – Example of a SAR image (from `nasa.gov`).

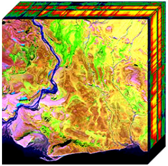

Figure – Example of a hyperspectral image (from `nasa.gov`).

Remark

For the parameter estimation step, we have to develop :

- robust estimators, i.e estimators that handle strong noise of SAR images
- "low-rank" estimators, i.e estimators that handle high dimension of hyperspectral images

Riemannian geometry.

A tool of interest in parameters estimation is the Riemannian geometry.
Briefly, a Riemannian manifold is a couple $(\mathcal{M}, g)$ where

- $\mathcal{M}$ is a *smooth manifold* (i.e locally Euclidean).
- g is a dot product on the tangent spaces called the *Riemannian metric*.

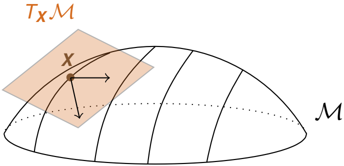

Figure – A manifold $\mathcal{M}$ with its tangent space $T_x\mathcal{M}$.

Remark

A tangent space at $x \in \mathcal{M}$, denoted $T_x\mathcal{M}$, is a linear space approximating $\mathcal{M}$ around x.

Optimization on manifolds is detailed in (Absil, Mahony, & Sepulchre, 2008).

Introduction to optimization on matrix manifolds.

Let f be a real-valued function to minimize over its parameter space :

$$\min_{\boldsymbol{X} \in \mathcal{M}} f(\boldsymbol{X})$$

where $\mathcal{M}$ is a Riemannian manifold representing the constraints of our problem.

Examples of Riemannian manifolds $\mathcal{M}$:

- linear space (no constraints) : $\mathbb{C}^{p \times p}$
- orthogonality constraints : $\mathrm{St}_{p,k} = \{\boldsymbol{X} \in \mathbb{C}^{p \times k} : \boldsymbol{X}^H \boldsymbol{X} = \boldsymbol{I}_k\}$
- positivity constraints : $\mathcal{H}_p^{++} = \{\boldsymbol{\Sigma} \in \mathcal{H}_p : \forall x \neq 0 \in \mathbb{C}^p,\ x^H \boldsymbol{\Sigma} x > 0\}$
- rank constraints : $\mathcal{H}_{p,k}^{+} = \{\boldsymbol{\Sigma} \in \mathcal{H}_p : \mathrm{rank}(\boldsymbol{\Sigma}) = k\}$
- norm constraints : $S^{p^2-1} = \{\boldsymbol{X} \in \mathbb{C}^{p \times p} : \|\boldsymbol{X}\|_F = 1\}$
- ...

Remark

In all these examples, $\mathcal{M}$ is a sub-manifold of an Hermitian space.

Intuition of optimization on a matrix manifold.

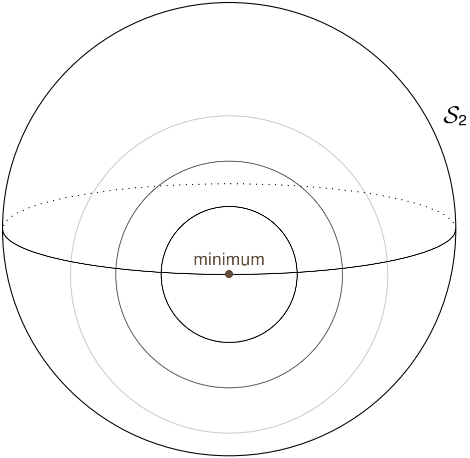

Figure – Example of the minimisation of a function on the sphere $\mathcal{S}_2$.

Intuition of optimization on a matrix manifold.

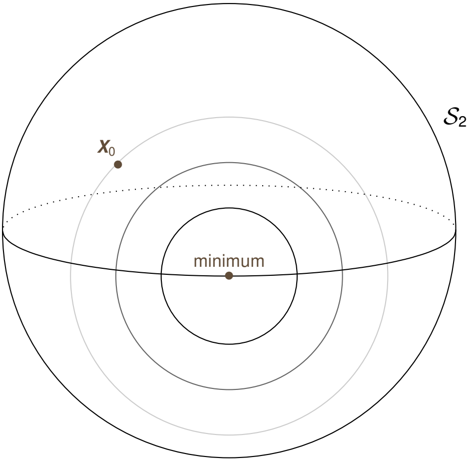

Figure – Example of the minimisation of a function on the sphere $\mathcal{S}_2$.

Intuition of optimization on a matrix manifold.

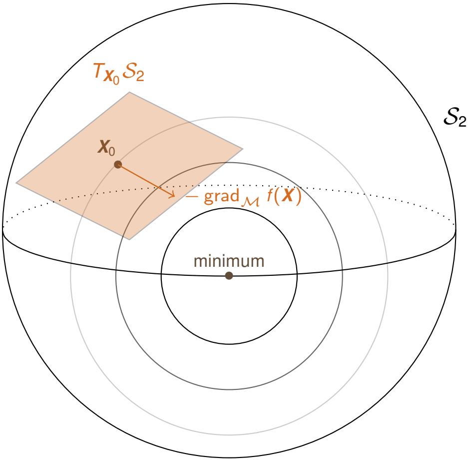

Figure – Example of the minimisation of a function on the sphere $\mathcal{S}_2$.

Intuition of optimization on a matrix manifold.

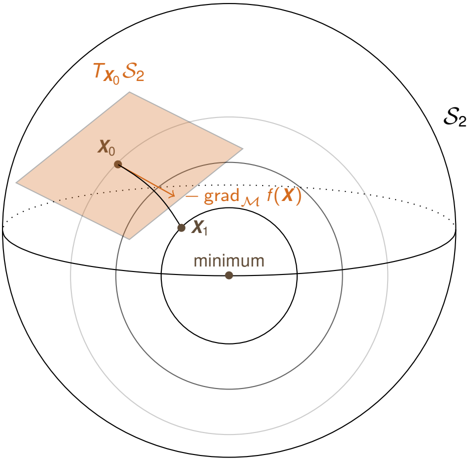

Figure – Example of the minimisation of a function on the sphere $\mathcal{S}_2$.

Study of a low rank model.

In the case of p-dimensional data, we can assume that data belongs to a subspace (lower k-dimensional vector space)(Tipping & Bishop, 1999).

In particular we studied the model where the signal follows a non Gaussian distribution in the subspace span($\boldsymbol{U}$) (**Collas**, Bouchard, Breloy, Ginolhac, et al., 2020) :

$$\underbrace{\boldsymbol{x_i}}_{\in\mathbb{C}^p} |\tau_i \overset{d}{=} \underbrace{\sqrt{\tau_i}\,\boldsymbol{U}\,\boldsymbol{g}_i}_{\text{signal}\in\text{span}(\boldsymbol{U})} + \underbrace{\boldsymbol{n_i}}_{\text{noise}\in\mathbb{C}^p} \tag{1}$$

where $\boldsymbol{g}_i \sim \mathbb{C}\mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_k)$ and $\boldsymbol{n_i} \sim \mathbb{C}\mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_p)$ are independent ; $\boldsymbol{\tau} \in (\mathbb{R}_*^+)^n$ contains the unknown deterministic textures τ_i ; and $\boldsymbol{U} \in \mathbb{C}^{p\times k}$ is an orthogonal basis of the subspace.

Study of a low rank model.

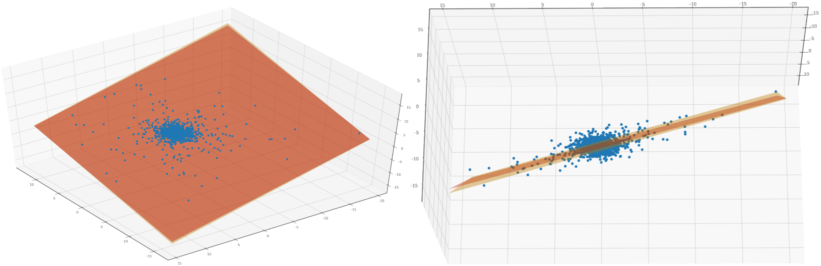

Figure – Scatter plot of samples $\{\boldsymbol{x_i}\}_{i=1}^{1000}$ with real and estimated subspaces respectively in orange and red in the case $\mathbb{E}[\tau_i] = 10$.

Remark

Both subspaces are really close !

MLE and intrinsic Cramèr-Rao bound.

Maximization of the likelihood while respecting the constraints :

- $\boldsymbol{U}$: orthogonal basis of the subspace (and thus invariant by rotation !)
- $\boldsymbol{\tau} \in (\mathbb{R}_*^+)^n$ (positivity constraint)

One iteration of the Riemannian steepest descent :

$$\boldsymbol{U}_{i+1}, \boldsymbol{\tau}_{i+1} = \underbrace{\exp^{\mathcal{M}}_{\boldsymbol{U}_i, \boldsymbol{\tau}_i}}_{\text{geodesic function}} (-\alpha_i \underbrace{\mathrm{grad}_{\mathcal{M}} L(\boldsymbol{U}_i, \boldsymbol{\tau}_i)}_{\text{Riemannian gradient}}) \tag{2}$$

Study of the performance through intrinsic Cramèr-Rao bounds (Smith, 2005) :

$$\overbrace{\mathbb{E}[d^2_{\mathrm{Gr}_{p,k}}(\pi(\hat{\boldsymbol{U}}), \pi(\boldsymbol{U}))]}^{\text{subspace estimation error}} \geq \frac{(p-k)k}{nc_{\boldsymbol{\tau}}} \approx \frac{(p-k)k}{n\mathbb{E}[\tau]} \tag{3}$$

$$\underbrace{\mathbb{E}[d^2_{(\mathbb{R}_*^+)^n}(\hat{\boldsymbol{\tau}}, \boldsymbol{\tau})]}_{\text{texture estimation error}} \geq \frac{1}{k} \sum_{i=1}^{n} \frac{(1+\tau_i)^2}{\tau_i^2} \tag{4}$$

Numerical experiment.

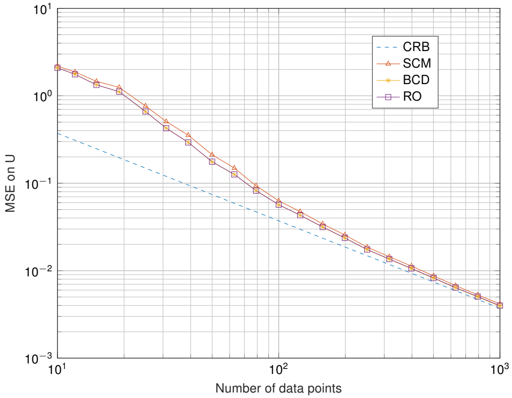

Figure – Subspace estimation error over $N = 500$ simulated sets $\{\boldsymbol{x}_i\}$ ($p = 15$, $k = 3$) with respect to the number of samples n for the three estimators (RO=Riemannian optimization).

A Tyler-type estimator using Riemannian optimization.

We also studied the joint estimation of location $\boldsymbol{\mu}$ and scatter matrix $\boldsymbol{\Sigma}$ in a robust manner (**Collas**, Bouchard, Breloy, Ren, et al., 2020) :

$$\boldsymbol{x_i} \sim \mathbb{C}\mathcal{N}(\boldsymbol{\mu}, \tau_i \boldsymbol{\Sigma}) \tag{5}$$

with $\boldsymbol{\mu} \in \mathbb{C}^p$, $\boldsymbol{\tau} \in (\mathbb{R}_*^+)^n$ and $\boldsymbol{\Sigma} \in \mathcal{H}_p^{++} (\boldsymbol{\Sigma} \succ 0)$.

- $\boldsymbol{\mu} = 0$: Tyler's estimator converges to the MLE (Tyler, 1987).
- $\boldsymbol{\mu}$ unknown : no estimator realizing the MLE exists ...

Idea

Optimizing the likelihood using a Riemannian gradient descent in order to respect the constraints $\tau_i > 0$ and $\boldsymbol{\Sigma} \succ 0$.

A Tyler-type estimator using Riemannian optimization.

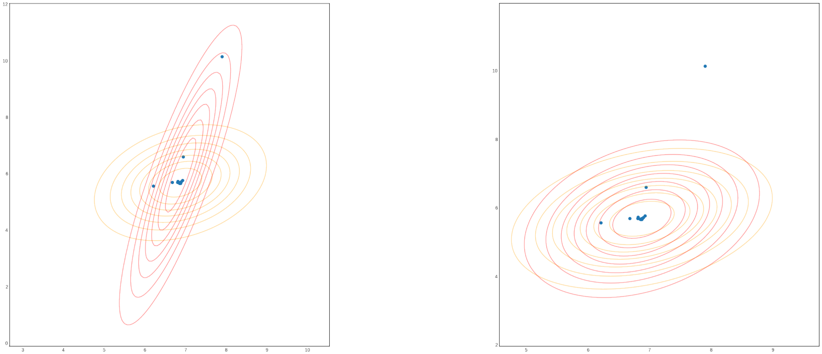

Figure – Scatter plot of samples $\{\boldsymbol{x_i}\}_{i=1}^{10}$ with real and estimated p.d.f respectively in orange and red. Left are the Gaussian estimators. Right are our estimators.

Numerical experiment.

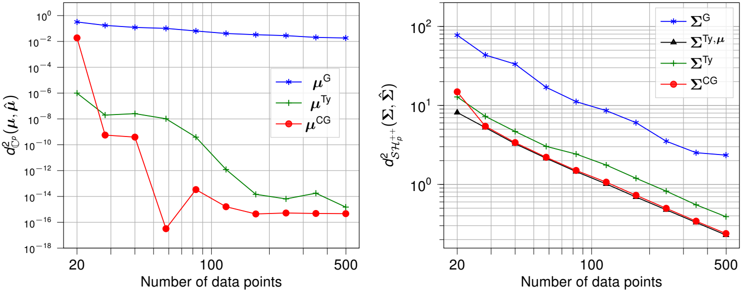

Figure – Subspace estimation error over $N = 200$ simulated sets $\{\boldsymbol{x_i}\}$ ($p = 10$) with respect to the number of samples n for different estimators. CG is the Riemannian Conjugate Gradient and "Ty,$\boldsymbol{\mu}$" is the Tyler's estimator with $\boldsymbol{\mu}$ known.

K-means on manifold.

Next step : cluster data which lie on a Riemannian manifold. For example, $(\boldsymbol{\tau}, \boldsymbol{U})$ parameters from the low rank model and $(\boldsymbol{\mu}, \boldsymbol{\tau}, \boldsymbol{\Sigma})$ parameters from the Compound Gaussian distribution.
To do K-means on a Riemannian manifold :

- estimation of parameters
- compute Riemannian distances between parameters and barycenters
- compute barycenters on the Riemannian manifold

References

Absil, P.-A., Mahony, R., & Sepulchre, R. (2008). *Optimization algorithms on matrix manifolds*. Princeton, NJ, USA : Princeton University Press.

Mian, A., **A. Collas**, Breloy, A., Ginolhac, G., & Ovarlez, J. P. (2020). Robust low-rank change detection for multivariate sar image time series. *IEEE Journal of Selected Topics in Applied arth Observations and Remote Sensing*, *13*, 3545-3556. doi: 10.1109/JSTARS.2020.2999615

Smith, S. (2005, 06). Covariance, subspace, and intrinsic Cramér-Rao bounds. *Signal Processing, IEEE Transactions on*, *53*, 1610 - 1630. doi: 10.1109/TSP.2005.845428

Collas, A., Bouchard, F., Breloy, A., Ginolhac, G., Ren, C., & Ovarlez, J.-P. (2020). *A riemannian geometry for probabilistic pca with compound gaussian signals.* (Preprint)

Collas, A., Bouchard, F., Breloy, A., Ren, C., Ginolhac, G., & Ovarlez, J.-P. (2020). *A tyler-type estimator of location and scatter leveraging riemannian optimization.* (Submitted in ICASSP 2021)

Tipping, M. E., & Bishop, C. M. (1999). Probabilistic principal component analysis. *Journal of the Royal Statistical Society : Series B*