MOM-based robust nonlinear anomaly detection for multispectral and hyperspectral data

Gaspar Massiot
joint work with Sidonie Lefebvre, Matthieu Lerasle, Zoltán Szabó, Guillaume Lecué and Eric Moulines.

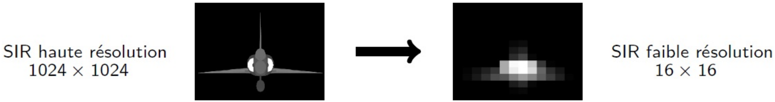
50èmes Journées de Statistique de la SFdS
Paris Saclay

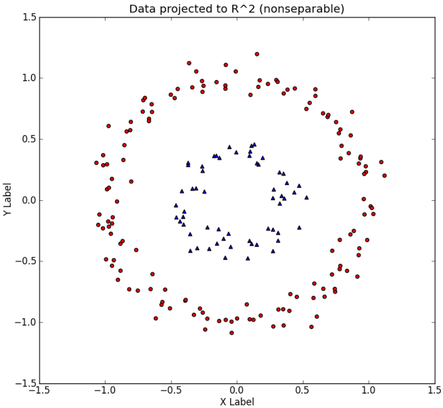
ONERA
THE FRENCH AEROSPACE LAB

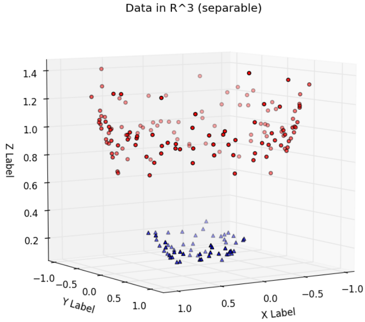
ONERA Plane dataset - Multispectral dataset

- Simulated multispectral **aircraft signatures** by ONERA

- Distance target/sensor ⇒ low resolution







source : reddit.com/r/MachineLearning/

- $\mathcal{X}$ ambient space
- $\mathcal{H}$ is a reproducing kernel Hilbert space (RKHS)
- $\exists! K : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ such that $K(\cdot, x) \in \mathcal{H}$ for all $x \in \mathcal{H}$ and $f(x) = \langle K(\cdot, x), f \rangle_{\mathcal{H}}$ for all $f \in \mathcal{H}$, $x \in \mathcal{X}$.
- K is called the reproducing kernel of $\mathcal{H}$ as

$$K(x, y) = \langle K(\cdot, x), K(\cdot, y) \rangle_{\mathcal{H}}$$

  and is symmetric and positive definite.
- $\mathcal{H} = \overline{\text{span}\{K(\cdot, x) : x \in \mathcal{X}\}}$ (linear span of kernel functions)

- **Mean embedding*** of prob. meas. $\mathbb{P}$

$$\mu_{\mathbb{P}} = \int_{\mathcal{X}} K(\cdot, x) \mathrm{d}\mathbb{P}(x) \in \mathcal{H}_K,$$

- $\mu_{\mathbb{P}}$ is well defined
  - ▷ when $\int_{\mathcal{X}} \sqrt{K(x,x)} \mathrm{d}\mathbb{P}(x) < \infty$,
  - ▷ specifically if K is bounded, example: Gaussian kernel.

$\Rightarrow$ representation of a prob. dist. on $\mathcal{X}$ as a point in $\mathcal{H}_K$.

- **Mean reproducing property:**

$$\mathbb{E}_{x \sim \mathbb{P}} f(x) = \langle f, \mu_{\mathbb{P}} \rangle_{\mathcal{H}_K} \quad (\forall f \in \mathcal{H}_K)$$

*Berlinet & Thomas-Agnan, 2004 ; Smola et al., 2007

- **Maximum mean discrepancy** (MMD)*

$$\text{MMD}_K(\mathbb{P}, \mathbb{Q}) = \|\mu_{\mathbb{P}} - \mu_{\mathbb{Q}}\|_{\mathcal{H}_K} .$$

$\Rightarrow$ (semi-)metric on distributions on $\mathcal{X}$.

- **Gretton et al. (2012):** Unbiased estimation

$$\text{MMD}_u(\mathbb{P}, \mathbb{Q}) = \frac{1}{N(N-1)} \sum_{i \neq j = 1}^{N} h(z_i, z_j),$$

where $z_i = (x_i, y_i)$ and
$h(z_i, z_j) = K(x_i, x_j) + K(y_i, y_j) - K(x_i, y_j) - K(x_j, y_i)$.

*Smola et al., 2007 ; Gretton et al., 2012

- Intuitively, MONs* replace the linear operation of expectation with the median of averages taken over non-overlapping blocks of the data, in order to get a robust estimate thanks to the median step.

$$x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad x_6 \quad \cdots \quad x_{N-2} \quad x_{N-1} \quad x_N$$

$$\mathbb{P}_{B_1} := \frac{1}{|B_1|}(\delta_{x_1} + \delta_{x_2} + \delta_{x_3}) \qquad \mathbb{P}_{B_2} \qquad \mathbb{P}_{B_Q}$$

- Median Of meaN (MON)

$$\mathrm{MON}_Q[f] := \underset{q \in [Q]}{\mathrm{med}} \left\{ \mathbb{P}_{B_q} f \right\} = \underset{q \in [Q]}{\mathrm{med}} \left\{ \left\langle f, \mu_{B_q} \right\rangle_{\mathcal{H}_K} \right\}.$$

*Jerrum et al., 1986 ; Lugosi & Mendelson, 2017

- Robust estimation of the mean embedding based on MON: $\hat{\mu}_{\mathbb{P},Q}$

$$\hat{\mu}_{\mathbb{P},Q} = \hat{\mu}_{\mathbb{P},Q}(x_{1:N}) \in \operatorname*{argmin}_{f \in \mathcal{H}_K} \sup_{g \in \mathcal{H}_K} J(f,g),$$

where $J(f,g) = \mathrm{MON}_Q\left[\|f - K(\cdot,x)\|_K^2 - \|g - K(\cdot,x)\|_K^2\right]$.

- Robust estimation of the MMD based on MON: $\widehat{\mathrm{MMD}}_Q(\mathbb{P},\mathbb{Q})$:

$$\widehat{\mathrm{MMD}}_Q(\mathbb{P},\mathbb{Q}) = \sup_{f \in \mathcal{B}_K} \operatorname*{med}_{q \in [Q]} \left\{ \left\langle f, \mu_{B_q,\mathbb{P}} - \mu_{B_q,\mathbb{Q}} \right\rangle_K \right\}.$$

where $\mathcal{B}_K = \{f \in \mathcal{H}_K : \|f\|_K \leq 1\}$.

- **Mean embedding:**
  - N_c elements $(x_{n_j})_{j=1}^{N_c}$ of the sample are arbitrarily corrupted.
- **MMD:**
  - $\{(x_{n_j}, y_{n_j})\}_{j=1}^{N_c}$ can be contaminated.
- The number of corrupted samples can be (almost) half of the number of blocks: $\exists \delta \in (0, 1/2]$ such that $N_c \leq Q(1/2 - \delta)$.
- $\delta = \frac{1}{2} \Rightarrow$ no contamination.

Theorem (Consistency and outlier robustness)

Assume $\Sigma_{\mathbb{P}}, \Sigma_{\mathbb{Q}} \in \mathcal{L}_1(\mathcal{H}_K)$. *Let* $c_1 = 2\left(1+\sqrt{2}\right)$. *Then with prob. at least* $1 - e^{-\frac{Q\delta^2}{18}}$

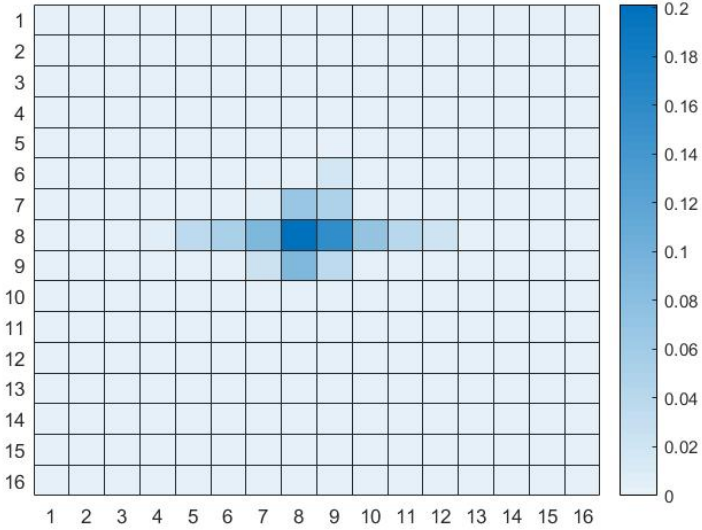
$$\left\|\hat{\mu}_{\mathbb{P},Q} - \mu_{\mathbb{P}}\right\|_K \leq c_1 \max\left(\sqrt{\frac{3\left\|\Sigma_{\mathbb{P}}\right\| Q}{\delta N}}, \frac{12}{\delta}\sqrt{\frac{\mathrm{Tr}\left(\Sigma_{\mathbb{P}}\right)}{N}}\right)$$

and

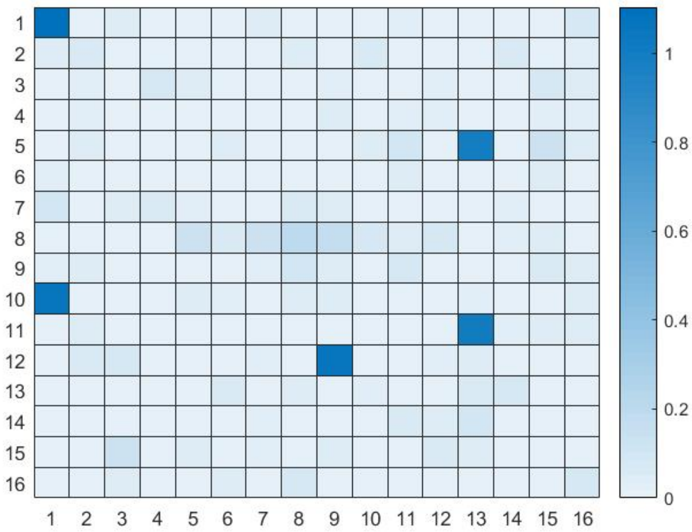
$$\left|\widehat{MMD}_Q(\mathbb{P},\mathbb{Q}) - MMD_Q(\mathbb{P},\mathbb{Q})\right| \leq$$

$$\leq 2\max\left(\sqrt{\frac{3\left(\left\|\Sigma_{\mathbb{P}}\right\| + \left\|\Sigma_{\mathbb{Q}}\right\|\right) Q}{\delta N}} \frac{12}{\delta}\sqrt{\frac{\mathrm{Tr}\left(\Sigma_{\mathbb{P}}\right) + \mathrm{Tr}\left(\Sigma_{\mathbb{Q}}\right)}{N}}\right).$$

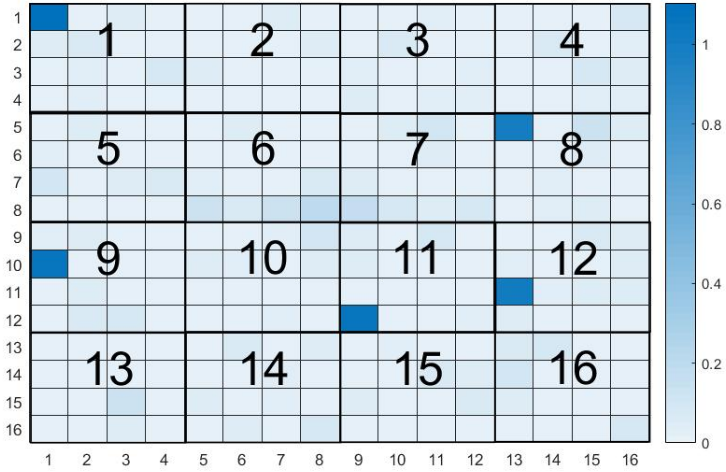
Pixel patches, plane in the middle

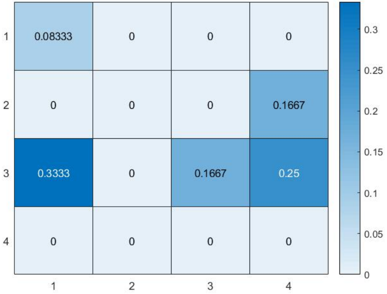


Plane + noise + contamination

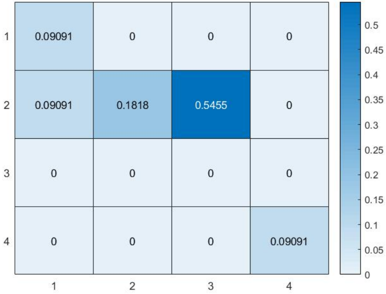


Approximation of the null distribution





Anomaly detection using MMD_u



Anomaly detection using MONK estimator

Conclusion

- New outlier-robust mean embedding and MMD estimators.
- Obtained estimators
  - obey optimal sub-Gaussian deviation bounds
  - are robust to contamination

Lerasle, M., Szabó, Z., Lecué, G., Massiot, G., & Moulines, E. (2018). MONK – Outlier-Robust Mean Embedding Estimation by Median-of-Means. *arXiv preprint arXiv:1802.04784.*

References I

- Kernel methods

Schölkopf, B., & Smola, A. J. (2001). *Learning with kernels: support vector machines, regularization, optimization, and beyond.* MIT press.

Steinwart, I., & Christmann, A. (2008). *Support vector machines*. Springer.

- Mean embedding

Muandet, K., Fukumizu, K., Sriperumbudur, B., & Schölkopf, B. (2017). Kernel mean embedding of distributions: A review and beyond. *Foundations and Trends in Machine Learning, 10*(1-2), 1-141.

- MOM / MON

Lugosi, G., & Mendelson, S. (2017). Sub-Gaussian estimators of the mean of a random vector. *To appear in Ann. Statist. arXiv preprint arXiv:1702.00482.*