Supervised classification of multidimensional and irregularly sampled signals
Alexandre Constantin, Mathieu Fauvel, Stéphane Girard, Serge Iovleff

To cite this version:
Alexandre Constantin, Mathieu Fauvel, Stéphane Girard, Serge Iovleff. Supervised classification of multidimensional and irregularly sampled signals. Statlearn 2019 - Workshop on Challenging problems in Statistical Learning, Apr 2019, Grenoble, France. pp.1-1. hal-02092347
Supervised classification of multidimensional and irregularly sampled signals.

Alexandre Constantin¹, Mathieu Fauvel², Stéphane Girard¹ and Serge Iovleff³

¹ Université Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, Grenoble, France
² CESBIO, Université de Toulouse, CNES/CNRS/IRD/UPS/INRA, Toulouse, France
³ Laboratoire Paul Painlevé - Université Lille 1, CNRS, Inria, France

Introduction

Background:
Recent space missions, such as Copernicus Sentinel-2a, provide high resolution Satellite Image Time Series (SITS) to study continental surfaces, with a very short revisit period (5 days for sentinel-2). In order to process such statistical models are regularly used [1, 2], which usually require a regular temporal sampling. However, for SITS, clouds and shadows (eg. figure from [3]), as well as the satellite orbit, an irregular temporal sampling is common.

Contribution:
A new statistical approach using Gaussian processes is proposed to classify irregularly sampled signals without temporal rescaling. Moreover, the model offers a theoretical framework to impute missing values such as cloudy pixels.

Gaussian Processes (GP) model:
Let $S = \{(y_i, z_i)\}_{i=1}^n$ be a set of multidimensional and irregularly sampled signals. A signal $Y$ is modeled as a vector of $p$ independent random processes $T \rightarrow \mathbb{R}^p$, with $T = [0,T]$. The associated label is modeled by a discrete random variable $Z$ taking its values in $\{1, \ldots, C\}$. The model introduced here is based on two assumptions: 1) The coordinate processes $Y_{i,b}$, $b \in \{1, \ldots, p\}$ of $Y$ are independent, 2) Each process $Y_i$ is, conditionally to $Z = c$, a Gaussian process. Then

$$Y_i(t)|Z = c \sim \mathcal{GP}(m_{b,c}(t), K_{b,c}(t,s)),$$

where $m_{b,c}: T \rightarrow \mathbb{R}^p$ is a mean function, and $K_{b,c}$ a covariance kernel with hyperparameters $\theta_{b,c}$. For example $\theta_{b,c} = \{\gamma^2_{b,c}, h_{b,c}, \sigma^2_{b,c}\}$ with

$$K_{b,c}(t,s) = \gamma^2_{b,c} k(t,s) h_{b,c} + \sigma^2_{b,c} \delta_{t,s}.$$

An irregularly sampled noisy signal $y_i$ is observed on $T_i$ time stamps $\{t_1^i, \ldots, t_{n_i}^i\} \in T$ and its $b$th coordinate is represented by a vector in $\mathbb{R}^{T_i}$. We write $y_{i,b} = [y_1^{i,b}, \ldots, y_{n_i}^{i,b}]^\top$, with

$$y_{i,b} | Z_i = c \sim \mathcal{N}_{T_i}(\mu_{i,b,c}, \Sigma_{b,c}).$$

There $\mu_{i,b,c} = B_i \alpha_{b,c}$ is the sample mean projected on a finite-dimensional space ($B_i$ is the fixed design matrix, $\alpha_{b,c}$ is the unknown vector of coordinates). $\Sigma_{b,c}$ is the kernel $K_{b,c}$ evaluations at $\{t_1^i, \ldots, t_{n_i}^i\}$.

Estimation:
• $\alpha_{b,c}$ and $\theta_{b,c}$ are estimated by maximizing the log-likelihood,

$$-\frac{1}{2} \sum_{i|Z_i = c} \log \left| \Sigma'_{b,c} \right| + (y_{i,b} - B_i \alpha_{b,c})^\top \Sigma'_{b,c}^{-1} (y_{i,b} - B_i \alpha_{b,c}).$$

• $\alpha_{b,c}$ is given by an explicit formula, while $\theta_{b,c}$ is computed thanks to a gradient technique.

Classification and Imputation of missing values

The assigned class is given by the MAP rule from the posterior probability

$$P(Z = c | y_i) = \frac{\hat{\pi}_c \prod_{b=1}^p f_{t_b}(y_j, B_i \alpha_{b,c}, \Sigma_{b,c}(\theta_{b,c}))}{\sum_{c'} \hat{\pi}_{c'} \prod_{b=1}^p f_{t_b}(y_j, B_i \alpha_{b,c}, \Sigma_{b,c}(\theta_{b,c}))}.$$

When the class is known to be $c$, the missing value at $t^*$ is estimated through the computation of conditional expectation.

$$\hat{Y}_{i,c}(t^*) = B_i(t^*) \alpha_{b,c} + \Sigma'_{b,c}(\theta_{b,c})^{-1} (y_{i,b} - B_i \alpha_{b,c})$$

$$\text{var}(\hat{Y}_{i,c}(t^*)) = K_{b,c}(t^*, t^*) - \Sigma'_{b,c}(\theta_{b,c})^{-1} K_{b,c}(t_1, t^*) \Sigma'_{b,c}(\theta_{b,c})^{-1}.$$

We also generalized this imputation when the class is unknown.

Future work

We are now implementing the model for massive real data (Sentinel-2). We are also working on a new model when the bands are correlated.

“https://www.esa.int/Dur_Activities/Observing_the_Earth/Copernicus/Sentinel-2”

Reference: