



HAL
open science

Interpretable sparse sliced inverse regression for functional data

Victor Picheny, Rémi Servien, Nathalie N. Villa-Vialaneix

► **To cite this version:**

Victor Picheny, Rémi Servien, Nathalie N. Villa-Vialaneix. Interpretable sparse sliced inverse regression for functional data. Workshop Learning with Functional Data, Oct 2016, Lille, France. hal-02799481

HAL Id: hal-02799481

<https://hal.inrae.fr/hal-02799481>

Submitted on 5 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Interpretable space sliced inverse regression for functional data

V. Picheny, R. Servien (speaker), N. Villa-Vialaneix

The focus is on the functional regression model in which a real random variable has to be predicted from functional predictors. We use the semiparametric framework of Sliced Inverse Regression (SIR). SIR is an effective method for dimension reduction of high-dimensional data which computes a linear projection of the predictors in a low-dimensional space, without loss on regression information. We address the issue of variable selection in functional SIR in order to improve the interpretability of the components. We extend the approaches of variable selection developed for multidimensional SIR to select intervals rather than separated evaluation points in the definition domain of functional predictors. SIR is formulated in different and equivalent ways which are alternatively used to produce ridge estimates of the components which are shrunk afterwards with a group-LASSO like penalty. An iterative procedure which automatically defines the relevant intervals is also introduced to provide a priori information to the sparse model. The approach is proved efficient on simulated data.