Voxel-wise nonlinear analysis toolbox for neurodegenerative diseases and aging

Authors
Signal Theory and Communications Department
Universitat Politècnica de Catalunya, BarcelonaTech
Barcelona, Spain
email

Abstract

The abstract paragraph should be indented ½ inch (3 picas) on both the left- and right-hand margins. Use 10 point type, with a vertical spacing (leading) of 11 points. The word **Abstract** must be centered, bold, and in point size 12. Two line spaces precede the abstract. The abstract must be limited to one paragraph.

1 Introduction

2 The toolbox

The toolbox comprises an independent *fitting library*, made up of different *model fitting* and *fit evaluation* methods, a *processing* module that interacts with the aforementioned *fitting library* providing the formatted data obtained from the *file system*, several *visualization* tools and a *CLI interface* that allows the interaction between the user and the *processing* module, supported by a *configuration file*.

2.1 Model fitting techniques

A model fitting consists on finding a parametric or a nonparametric function of some explanatory variables (**predictors**) and possibly some confound variables (**correctors**) that best fits the observations of the target variable in terms of a given quality metric or, conversely, that minimizes the loss between the prediction of the model and the actual observations.

- **General Linear Model (GLM)**
  The General Linear model is a generalization of multiple linear regression to the case of more than one dependent variable. As in the case of multiple linear regression, the most common lost function is the Residual Sum of Squares, and the optimization procedure used is Ordinary Least Squares, which yields the well-known normal equation $X^T X\beta = X^T y$, and from that we solve for the β parameters to obtain the final solution: $\beta = (X^T X)^{-1} X^T y$.

  A possible approach to model nonliniearities with this model is a polynomial basis expansion of degree d, that is, the input space $\mathfrak{X}$ is mapped into another feature space $\mathfrak{F}$ that also includes the polynomial terms of the variables: ($\Phi : \mathfrak{X} \rightarrow \mathfrak{F}$).
- **Generalized Additive Model (GAM)**
  A Generalized Additive Model is a Generalized Linear Model in which the observations of the target variable depend linearly on unknown smooth functions of some predictor variables: $f(X) = \alpha + \sum_{i=1}^{k} f_i(X_i)$. Here $f_1, f_2, ..., f_k$ are nonparametric smooth functions that are simultaneously estimated using scatterplot smoothers by means of the **backfitting**

30th Conference on Neural Information Processing Systems (NIPS 2016), Barcelona, Spain.

algorithm. Several fitting methods can be accomodated in this framework by using different smoother operators, such as cubic splines, polynomial or Gaussian smoothers.

- **Support Vector Regression (SVR)**

  The regression counterpart of the well-known Support Vector Machines, Support Vector Regression, is based on the following idea: the goal is to find a function that has at most ϵ deviation from the observations and, at the same time, is as flat as possible. However, the ϵ deviation contraint is not feasible sometimes, and a hyperparameter that controls the degree up to which deviations larger than ϵ are tolerated is introduced, C. The linear function for SVR is $f(x) = \langle w, x \rangle + b$, and then the optimization problem is formulated as follows:

  $$\text{minimize } \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(\zeta_i + \zeta_i^*) \text{ subject to } \begin{cases} y_i - \langle w, x_i \rangle - b \leq \epsilon + \zeta_i \\ \langle w, x_i \rangle + b - y_i \leq \epsilon + \zeta_i^* \\ \zeta_i, \zeta_i^* \geq 0 \end{cases} \tag{1}$$

  whose solution is $f(x) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)\langle x_i, x \rangle + b$, where $(\alpha_i - \alpha_i^*)$ are the coefficients of the dual problem that are found in the optimization process.

  In context of SVR the nonlinearities are introduced with the "kernel trick", that is, a kernel function $k(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle$ is introduced that implicitly maps the inputs from their original space into another high-dimensional space without requiring to know the explicit mapping $\Phi(\cdot)$. The solution using this kernel function is then $f(x) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)k(x_i, x) + b$.

  The kernel function used in this toolbox is the Radial Basis Function or Gaussian kernel, which is defined as $k(x_i, x_j) = exp(-\gamma\|x_i - x_j\|^2)$.

2.2 Hyperparameters search algorithm

2.3 Fit evaluation methods

- **F-test**
- **PRSS[1], Variance-Normalized PRSS**

2.3.1 Error functions

1. MSE
2. Cp statistic
3. ANOVA-based error

2.4 Interactive visualization tools

3 Implementation details

4 Experiments

4.1 Dataset

5 Conclusions

Acknowledgments

References

[1] Alexander, J.A. & Mozer, M.C. (1995) Template-based algorithms for connectionist rule extraction. In G. Tesauro, D.S. Touretzky and T.K. Leen (eds.), *Advances in Neural Information Processing Systems 7*, pp. 609–616. Cambridge, MA: MIT Press.

[1] Penalized Residual Sum of Squares