Near-perfect classification rules for second-order stochastic processes

Alberto Suárez*, José Luis Torrecilla, Carlos Ramos-Carreño, Adrián Muñoz-Perera
Computer Science Department
Universidad Autónoma de Madrid, Spain
{alberto.suarez, joseluis.torrecilla, carlos.ramos}@uam.es, pereradrian@gmail.com

January 26, 2024

Abstract

Functional data classification is an active field of research that has multiple applications in different areas, such as medicine, weather modelling and forecasting, and speech recognition, among others. In these types of classification problems, the instances available for induction are characterized by functions of a continuous parameter, such as trajectories in time or curves in space [11, 9]. Functional classification problems exhibit significant qualitative differences with their multivariate counterparts [3, 13]. These differences hinge on the infinite-dimensional nature of the data, the existence of a natural ordering within the functional observations, their underlying smoothness, and the fact that probability density functions generally does not exist for random functions [5]. For equivalent Gaussian processes, even if the individual class-conditional probability densities do not exist, an optimal classification rule can be formulated in terms of the Radon-Nikodym derivative between the corresponding measures, which plays the role of the likelihood ratio in these types of infinite-dimensional problems [1]. In this work we focus on cases in which this derivative is ill-defined and near-perfect classification (zero Bayes error in the population limit) is obtained [6, 12]. Specifically, we derive explicit expressions of optimal prediction rules for binary classification problems in which the data instances are characterized by trajectories X, sampled from second order stochastic processes defined on the interval $[0, T]$ in the real line. The stochastic processes from which the instances are sampled are different for each of the two classes. This problem has been analyzed earlier in the literature in both the homo- and heteroscedastic settings [6, 7, 4, 2, 12]. The conditions for near perfect classification when the stochastic processes have different means were first derived in [6]. The current paper builds on that work by considering cases in which the singularities are associated to the covariance structure of the processes. The main novel contribution is to derive classification rules when the means of the processes are equal [12]. These rules are expressed in terms of limits of approximations to the inner products in the reproducing kernel Hilbert spaces associated to the covariance functions of the processes. In the general case, these limits are singular and, as a result, near-perfect classification is obtained. Carrying out a detailed analysis of these rules and their singularities, we provide novel derivations of some known results and gain insight into the mechanisms by which near perfect classification occurs. A further novel result is the derivation of explicit tests to determine whether two Gaussian processes are equivalent or mutually singular [10, 8]. These tests are derived from the condition that for the two processes to be equivalent, the singularities that appear in the corresponding classification rule must cancel out.

*Part of this work was carried out while in residence at the Sorbonne Center for Artificial Intelligence (SCAI, Paris)

References

[1] Amparo Baíllo, Antonio Cuevas, and Juan Antonio Cuesta-Albertos. Supervised Classification for a Family of Gaussian Functional Models. *Scandinavian Journal of Statistics*, 38(3):480–498, 2011.

[2] José R. Berrendero, Antonio Cuevas, and José L. Torrecilla. On the use of reproducing kernel Hilbert spaces in functional classification. *Journal of the American Statistical Association*, 113(523):1210–1218, 2018.

[3] Antonio Cuevas. A partial overview of the theory of statistics with functional data. *Journal of Statistical Planning and Inference*, 147:1 – 23, 2014.

[4] Xiongtao Dai, Hans-Georg Müller, and Fang Yao. Optimal Bayes classifiers for functional data and density ratios. *Biometrika*, 104(3):545–560, 2017.

[5] Aurore Delaigle and Peter Hall. Defining probability density for a distribution of random functions. *Ann. Statist.*, 38(2):1171–1193, 04 2010.

[6] Aurore Delaigle and Peter Hall. Achieving near perfect classification for functional data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 74(2):267–286, 2012.

[7] Aurore Delaigle and Peter Hall. Classification using censored functional data. *Journal of the American Statistical Association*, 108(504):1269–1283, 2013.

[8] Jacob Feldman. Equivalence and perpendicularity of Gaussian processes. *Pacific J. Math.*, 8(4):699–708, 1958.

[9] Frédéric Ferraty and Philippe Vieu. *Nonparametric Functional Data Analysis: Theory and Practice*. Springer Series in Statistics. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.

[10] Jaroslav Hájek. A property of J-divergences of marginal probability distributions. *Czechoslovak Mathematical Journal*, 08(3):460–463, 1958.

[11] James O. Ramsay and Bernard W. Silverman. *Functional Data Analysis*. Springer Series in Statistics. Springer, 2nd edition, 2005.

[12] José L. Torrecilla, Carlos Ramos-Carreño, Manuel Sánchez-Montañés, and Alberto Suárez. Optimal classification of gaussian processes in homo- and heteroscedastic settings. *Statistics and Computing*, 30(4):1091–1111, jul 2020.

[13] Jane-Ling Wang, Jeng-Min Chiou, and Hans-Georg Müller. Functional data analysis. *Annual Review of Statistics and Its Application*, 3(1):257–295, 2016.