VARIABLE SELECTION IN LINEAR REGRESSION: SEVERAL APPROACHES BASED ON NORMALIZED MAXIMUM LIKELIHOOD

Ciprian Doru Giurcăneanu, Seyed Alireza Razavi, Antti Liski

Department of Signal Processing
Tampere University of Technology, Finland
ciprian.giurcaneanu@tut.fi, alireza.razavi@tut.fi, antti.liski@tut.fi

ABSTRACT

In this talk, we discuss the application of the normalized maximum likelihood (NML) for model selection in Gaussian linear regression. All the results which will be presented have been recently published in [1].

1. INTRODUCTION

The use of the NML for model selection in Gaussian linear regression poses troubles because the normalization coefficient is not finite. The most elegant solution has been proposed by Rissanen and consists in applying a particular constrain for the data space [2]. The resulting criterion is independent of arbitrarily selected hyper-parameters. Surprisingly, for about one decade, it was totally ignored the important fact that the closed-form expression of the criterion depends on the particular constraint which has been involved in its derivation. Only recently, it was shown in [3] that two other criteria can be obtained by employing constraints which are different of the one used in [2]. Hence, novel NML-based criteria can be devised by enforcing various constraints.

2. MAIN RESULTS

Our talk is focused on the following aspects:

- We demonstrate that the methodology introduced by Rissanen can be applied in a more general framework, and not only for the ellipsoidal constraints which have been considered in [2, 3]. We also discuss the particular case of rhomboidal constraint.
- We analyze the relationship between Rissanen criterion and the two criteria that have been introduced in [3].
- We compare experimentally the capabilities of the NML-based selection rules against BIC [4], $\mathrm{AIC}_{\mathrm{c3}}$ [5], CME [6], $\mathrm{MML}_{\mathrm{g}}$ and $\mathrm{MML}_{\mathrm{u}}$ [7].
- We also provide some guidance on the use of various criteria in model selection.

3. REFERENCES

[1] C.D. Giurcăneanu, S.A. Razavi, and A. Liski, "Variable selection in linear regression: Several approaches based on normalized maximum likelihood," *Signal Processing*, vol. 91, no. 8, pp. 1671–1692, 2011.

[2] J. Rissanen, "MDL denoising," *IEEE Transactions on Information Theory*, vol. 46, no. 7, pp. 2537–2543, 2000.

[3] E. Liski and A. Liski, "Minimum Description Length model selection in Gaussian regression under data constraints," in *Statistical Inference, Econometric Analysis and Matrix Algebra, Festschrift in Honour of Götz Trenkler*, B. Schipp and W. Krämer, Eds., pp. 201–208. Springer, `http://www.springerlink.com/content/n32t4671713w87g0/fulltext.pdf`, 2009.

[4] G. Schwarz, "Estimating the dimension of a model," *The Annals of Statistics*, vol. 6, no. 2, pp. 461–464, Mar. 1978.

[5] A.-K. Seghouane, "Asymptotic bootstrap corrections of AIC for linear regression models," *Signal Processing*, vol. 90, no. 1, pp. 217 – 224, 2010.

[6] S. Kay, "Conditional model order estimation," *IEEE Transactions on Signal Processing*, vol. 49, no. 9, pp. 1910–1917, Sep. 2001.

[7] D.F. Schmidt and E. Makalic, "MML invariant linear regression," in *AI 2009: Advances in Artificial Intelligence, 22nd Australasian Joint Conference, Melbourne, Australia, December 1-4, 2009. Proceedings*, A.E. Nicholson and X. Li, Eds. 2009, vol. 5866 of *Lecture Notes in Computer Science*, pp. 312–321, Springer.