Simultaneous estimation of the mean and the variance in heteroscedastic Gaussian regression

Xavier Gendre, *Laboratoire J.-A. Dieudonné*

Abstract

Let Y be a Gaussian vector of R^{n} of mean s and diagonal covariance matrix $Gamma$. Our aim is to estimate both s and the entries $sigma_{i}=Gamma_{i,i}$, for $i=1,dots,n$, on the basis of the observation of two independent copies of Y. Our approach is free of any prior assumption on s but requires that we know some upper bound $gamma$ on the ratio $max_{i}sigma_{i}/min_{i} sigma_{i}$. For example, the choice $gamma=1$ corresponds to the homoscedastic case where the components of Y are assumed to have common (unknown) variance. In the opposite, the choice $gamma>1$ corresponds to the heteroscedastic case where the variances of the components of Y are allowed to vary within some range. Our estimation strategy is based on model selection. We consider a family ${S_{m} times Sigma_{m}$, $min mathcal{M}}$ of parameter sets where S_{m} and $Sigma_{m}$ are linear spaces. To each $m in mathcal{M}$, we associate a pair of estimators $(hat{s}_{m}, hat{sigma}_{m})$ of $(s,sigma)$ with values in $S_{m}timesSigma_{m}$. Then we design a model selection procedure in view of selecting some $hat{m}$ among $mathcal{M}$ in such a way that the Kullback risk of $(hat{s}_{hat{m}}, hat{sigma}_{hat{m}})$ is as close as possible to the minimum of the Kullback risks among the family of estimators ${(hat{s}_{m},hat{sigma}_{m})$, $minmathcal{M}}$. Then we derive uniform rates of convergence for the estimator $(hat{s}_{hat{m}},hat{sigma}_{hat{m}})$ over Hölderian balls. Finally, we carry out a simulation study in order to illustrate the performances of our estimators in practice.

AMS 2000 subject classifications: 62G08.

Keywords: Gaussian regression, heteroscedasticity, model selection, Kullback risk, convergence rate.

Full Text: PDF

Gendre, Xavier, Simultaneous estimation of the mean and the variance in heteroscedastic Gaussian regression, Electronic Journal of Statistics, 2, (2008), 1345-1372 (electronic). DOI: 10.1214/08-EJS267.

References

[1] Akaike, H. (1969). Statistical predictor identification. Annals Inst. Statist. Math.. MR0286233

[2] Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In Proceedings 2nd International Symposium on Information Theory, P. Petrov and F. Csaki, Eds. Akademia Kiado, Budapest, 267–281. MR0483125

[3] Akaike, H. (1974). A new look at the statistical model identification. IEEE Trans. on Automatic Control. MR0423716

[4] Arlot, S. (2007). Rééchantillonnage et sélection de modèles. Ph.D. thesis, Université Paris 11.

[5] Baraud, Y. (2000). Model selection for regression on a fixed design. Probab. Theory

Related Fields 117, 467–493. MR1777129

[6] Baraud, Y., Giraud, C., and Huet, S. (2006). Gaussian model selection with an unknown variance. To appear in Annals of Statistics.

[7] Barron, A., Birgé, L., and Massart, P. (1999). Risk bounds for model selection via penalization. Probab. Theory Related Fields 113, 301–413. MR1679028

[8] Birgé, L. and Massart, P. (1997). From model selection to adaptive estimation. Festschrift for Lucien Lecam: Research Papers in Probability and Statistics. MR1462939

[9] Birgé, L. and Massart, P. (2001a). Gaussian model selection. Journal of the European Mathematical Society 3, 3, 203–268. MR1848946

[10] Birgé, L. and Massart, P. (2001b). A generalized c_p criterion for gaussian model selection. Prépublication 647, Universités de Paris 6 and Paris 7.

[11] Boucheron, S., Bousquet, O., Lugosi, G., and Massart, P. (2005). Moment inequalities for functions of independent random variables. Annals of Probability 33, 2, 514–560. MR2123200

[12] Comte, F. and Rozenholc, Y. (2002). Adaptive estimation of mean and volatility functions in (auto-)regressive models. Stochastic Processes and their Applications 97, 1, 111–145. MR1870963

[13] DeVore, R. and Lorentz, G. (1993). Constructive approximation. Vol. 303. Springer-Verlag. MR1261635

[14] Galtchouk, L. and Pergamenshchikov, S. (2005). Efficient adaptive nonparametric estimation in heteroscedastic regression models. Université Louis Pasteur, IRMA, Preprint. MR2187340

[15] Laurent, B. and Massart, P. (2000). Adaptive estimation of a quadratic functional by model selection. Annals of Statistics 28, 5, 1302–1338. MR1805785

[16] Mallows, C. (1973). Some comments on c_p. Technometrics 15, 661–675.

[17] McQuarrie, A. and Tsai, C. (1998). Regression and times series model selection. World Scientific Publishing Co, Inc. MR1641582

[18] R Development Core Team. (2007). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, http://www.R-project.org.

[19] Wang, L. and al. (2008). Effect of mean on variance function estimation in nonparametric regression. The Annals of Statistics 36, 2, 646–664. MR2396810

Electronic Journal of Statistics. ISSN: 1935-7524