

Accepted Manuscript

On asymptotic normality in nonlinear regression

Harry Haupt, Walter Oberhofer

PII: S0167-7152(08)00514-2
DOI: 10.1016/j.spl.2008.10.037
Reference: STAPRO 5266

To appear in: *Statistics and Probability Letters*

Received date: 17 September 2008

Revised date: 30 October 2008

Accepted date: 30 October 2008

Please cite this article as: Haupt, H., Oberhofer, W., On asymptotic normality in nonlinear regression. *Statistics and Probability Letters* (2008), doi:10.1016/j.spl.2008.10.037

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



On asymptotic normality in nonlinear regression ¹

Harry Haupt ^{a,*} Walter Oberhofer ^b

^a*Bielefeld University, Germany*

^b*University of Regensburg, Germany*

Abstract

In this paper we reconsider the results in Wang (1995, 1996), who studies the conditions for asymptotic normality of L_1 - and (constrained) L_2 -norm regression estimators in the nonlinear regression model. Unfortunately, Wang fails to state necessary global conditions.

Key words: Asymptotic normality, nonlinear regression.

1 Nonlinear regression

We consider the nonlinear regression model

$$y_i = f(x_i, \vartheta) + e_i, \quad i = 1, \dots, N, \quad (1)$$

where $\vartheta \in R^m$ is a vector of unknown parameters and $\{e_i\}$ is an error process. An extensive literature investigates the asymptotic properties of M-estimators $\hat{\vartheta}_N$ of ϑ , minimizing the loss function

$$M_N(\vartheta) = \frac{1}{N} \sum_{i=1}^N \rho(y_i - f(x_i, \vartheta)), \quad (2)$$

for different choices of ρ -functions $\rho_p = |z|^p$, $p \geq 1$, (e.g., Berinet, Liese and Vajda, 2000). Leading examples are L_1 ($p = 1$; e.g. Wang, 1995) and L_2 ($p = 2$; e.g. Wang, 1996) estimators.

* Corresponding author. Postal address: Bielefeld University, Centre for Statistics, Department of Business Administration and Economics, PO Box 100131, 33501 Bielefeld, Germany. Phone: +49 521 106 4873, fax: +49 521 106 6425.

Email address: hhaupt@wiwi.uni-bielefeld.de (Harry Haupt).

¹ We wish to thank the Editor Hira Koul for helpful suggestions.

For independent and identically distributed (iid) processes, consistency and asymptotic normality conditions of model (1) have been studied for L_2 regression by Wu (1980). For the same model Oberhofer (1982) establishes consistency for L_1 regression and Jureckova and Prochazka (1994) derive conditions for consistency and asymptotic normality for the asymmetrically weighted L_1 (quantile) regression. The results of Wu and Oberhofer have been extended to dependent and/or heterogeneous processes in subsequent research (e.g., Prakasa Rao, 1984). A comprehensive treatment and annotated bibliography for M-estimation of model (1) is given by Liese and Vajda (2003, 2004).

In his main result stated in Theorem 2, Wang (1995) studies the asymptotic normality of the L_1 estimator of model (1). His main theorem is based on three assumptions: A1 concerns the errors e_i , in particular that e_1, \dots, e_N are iid, A2 and A3 concern the smoothness and boundedness of $\nabla f(x_i, \vartheta)$. The requirements of the latter two assumptions, however, are postulated in A2 only for ϑ in a neighborhood of the true value ϑ_0 , and in A3 for $\vartheta = \vartheta_0$.

The solely local nature of these assumptions is rather astounding, as even identifiability requires an assertion global in ϑ . An example for a suitable identification assumption would be: for every $\epsilon > 0$ there exists a positive δ_ϵ such that

$$\liminf_N \inf_{\|\vartheta - \vartheta_0\| \geq \epsilon} \frac{1}{N} \sum_{i=1}^N |f(x_i, \vartheta) - f(x_i, \vartheta_0)| > \delta_\epsilon. \quad (3)$$

Usually such an assertion is made when discussing consistency, e.g., Oberhofer (1982) for L_1 estimators and Wu (1980), who discusses the consistent directions of L_2 estimators. Then asymptotic normality is proved given consistency and rate of convergence of the estimator $\hat{\vartheta}_N$ minimizing $M_N(\vartheta)$ (e.g., Liese and Vajda, 2003, p. 458).

As a consequence, asymptotic normality cannot be established globally without consistency, where the latter is based on an assertion global in ϑ . It lies near at hand to ask, whether and where these considerations are reflected in the argumentation of Wang?

2 Asymptotic normality

Wang (1995) refers to Prakasa Rao (1987) and argues that the latter uses an analogous argumentation for the least squares (L_2) estimation of ϑ in (1). Unfortunately, Wang's adaption is incorrect, though this may be difficult to see at first sight. Indeed, Theorem 6.2.1 (p. 373) in Prakasa Rao (1987) holds only locally, i.e. for ϑ such that

$$\vartheta = \vartheta_0 + N^{-1/2}\phi, \quad (4)$$

where ϕ is contained in a neighborhood V of ϑ_0 and the local parameter $\hat{\phi}_N = N^{1/2}(\hat{\vartheta}_N - \vartheta)$ minimizes the loss function

$$M_N(\phi) = M_N(\vartheta_0 + N^{-1/2}\phi) - M_N(\vartheta_0). \quad (5)$$

But this is derestricted by Prakasa Rao (1987) in the remark after Theorem 6.2.5 (p. 377), and requires a (global) identifiability assumption such as (3) or assumption (C_1) in Prakasa Rao (1987, p. 372). The latter step, which is based on the work of Ivanov (1976), establishes a rate of convergence result and is missing in Wang's argumentation. In the last paragraph of p. 237, Wang (1995) merely mentions that V can be arbitrarily large. As V depends on N , however, this argument is ill-founded. V can be arbitrarily large, but independently from N . Hence, a correct proof requires the consistency of the estimator of ϑ and in addition its sufficiently fast convergence to ϑ_0 (e.g., Prakasa Rao, 1987, equations (6.2.28) and (6.2.29)).

Note that the same reasoning applies analogously to a more recent paper of Wang (1996), which considers the asymptotic normality of the L_2 estimator of the parameter vector in (1), subject to a set of nonlinear equality and inequality constraints. In Theorem 6, Wang (1996) shows that $N^{1/2}(\hat{\vartheta} - \vartheta_0)$ converges to a normal distribution if $N^{1/2}(\vartheta - \vartheta_0)$ is contained in a bounded set. Again, however, all assumptions employed are only local.

References

- [1] Berline, A., Liese, F., Vajda, I., 2000. Necessary and sufficient conditions for consistency of M-estimates in regression models with general errors. *J. Statist. Plann. Inference* 89, 243-267.
- [2] Ivanov, A. V., 1976. An asymptotic expansion for the distribution of the least squares estimator of the non-linear regression parameter. *Theory Probab. Appl.* 21, 557-570.
- [3] Jureckova, J., Prochazka, B., 1994. Regression quantiles and trimmed least squares estimators in nonlinear regression models. *J. Nonpar. Statist.* 3, 201-222.
- [4] Liese, F., Vajda, I., 2003. A general asymptotic theory of M-estimators. I. *Math. Meth. Statist.* 12, 454-477.
- [5] Liese, F., Vajda, I., 2004. A general asymptotic theory of M-estimators. II. *Math. Meth. Statist.* 13, 82-95.
- [6] Oberhofer, W., 1982. The consistency of nonlinear regression minimizing the L_1 norm. *Ann. Statist.* 10, 316-319.
- [7] Prakasa Rao, B. L. S., 1984. The rate of convergence of the least squares estimator in a nonlinear regression model with dependent errors. *J. Multivariate Anal.* 14, 315-322.
- [8] Prakasa Rao, B. L. S., 1987. *Asymptotic theory of statistical inference*. Wiley Series in Probability and Mathematical Statistics.
- [9] Wang, J., 1995. Asymptotic normality of L_1 -estimators in nonlinear regression. *J. Multivariate Anal.* 54, 227-238.
- [10] Wang, J., 1996. Asymptotics of least-squares estimators for constrained nonlinear regression. *Ann. Statist.* 4, 1316-1326.
- [11] Wu, C.F., 1980. Characterizing the consistent directions of least squares estimates. *Ann. Statist.* 8, 789-801.
- [12] Wu, C.F., 1981. Asymptotic theory of nonlinear least squares estimation. *Ann. Statist.* 9, 501-513.