IWQW

Institut für Wirtschaftspolitik und Quantitative Wirtschaftsforschung

Diskussionspapier
Discussion Papers

No. 8/2010

A note Hadamard differentiability and differentiability in quadratic mean

Fabian Tinkl
University of Erlangen-Nuremberg

ISSN 1867-6707

Friedrich-Alexander-Universität
IWQW
Institut für Wirtschaftspolitik und Quantitative Wirtschaftsforschung

A note on Hadamard differentiability and differentiability in quadratic mean

Working Paper, 21. Oktober 2010

Fabian Tinkl[1]
University of Erlangen-Nuremberg
Fabian.Tinkl@wiso.uni-erlangen.de

Keywords and phrases: Hadamard differential, Differentiability in quadratic mean

Abstract

We proof that Hadamard differentiability in addition with usual assumptions on the loss function for M estimates implies differentiability in quadratic mean. Thus both concepts are exchangeable.

[1]Correspondence author: Fabian Tinkl, Department of Statistics and Econometrics, University of Erlangen-Nuremberg, Lange Gasse 20, D-90403 Nuremberg, E-Mail: Fabian.Tinkl@wiso.uni-erlangen.de

1 Introduction

As it was pointed out by [Pollard, 1997] there is a connection between differentiability in an L^2 sense, differentiability in quadratic mean and local asymptotic normality. The idea of local asymptotic normality (LAN) goes back to [LeCam, 1970], who showed that differentiability in quadratic mean implies an quadratic approximation for the log-likelihood function, without using the 2nd derivative. We will show, how a different concept of differentiability, namely the Hadamard differentials, are under mild regularity conditions equivalent to the definition of differentials in quadratic mean, and therefore the results of [Pollard, 1997], [LeCam, 1970] and [Fernholz, 1983] will coincide for M estimates. Thus it can be seen, that it is also sufficient for M estimates to be Hadamard differentiable to enjoy the LAN property.

2 The equivalence Theorem

Throughout this article we regard an σ-finite probability space and a sampling X_i (not necessarily iid.) from a measurable random variable $X : \Omega \to \mathbb{R}$, where X has a distribution function, which is absolutely continuous w.r.t. some σ-finite measure. Define a M estimate as the solution of the following minimization problem

$$\hat{\theta}_n = \arg\min_{\theta\in\Theta} \frac{1}{n}\sum_{i=1}^{n} \rho(X_i, \theta), \tag{1}$$

where θ is some unknown parameter, where $\theta \subset \Theta$, and Θ is some metric space with the uniform metric. Under standard conditions, that $\theta \to \rho(\theta)$ is continuous and Θ is compact as well as ρ is uniformly integrable w.r.t. θ, we get that $\hat{\theta}_n \xrightarrow{P} \theta_0$, where

$$\theta_0 = \arg\min_{\theta\in\Theta} E[\rho(X, \theta)]. \tag{2}$$

We suppress the dependence of ρ on $\omega \in \Omega$ from now on, and treat ρ as a function of the parameter(-vector) θ. If ρ is differentiable w.r.t. to θ, one may take a first-order Taylor series expansion around $\hat{\theta}_n$ in neighborhood of θ_0, where (2) attains its minimum:

$$\rho(\hat{\theta}_n) = \rho(\theta_0) + (\hat{\theta}_n - \theta_0)^T \rho'_\theta(\theta_0) + Rem(\theta)(\hat{\theta}_n - \theta_0). \tag{3}$$

Suppose that $Rem \to 0$ in some proper sense. If $Rem \xrightarrow{L^2} 0$, i.e. $E[|Rem(\hat{\theta}_n - \theta_0)|^2] = 0$, we call ρ differentiable in quadratic mean (DQM). If ρ is Hadamard differentiable, then $Rem \xrightarrow{P} 0$, see [Fernholz, 1983]. Under usual regularity conditions on ρ and its derivative we can conclude, that $\sqrt{n}(\hat{\theta}_n - \theta_0) \to N(0, \sigma^2)$, with $\sigma^2 = var(\phi(X))$, and ϕ is the influence function. Define Hadamard differentials as follows (see [van der Vaart, 1998]):

Definition 1 *A function $\rho : D_\theta \to W$ is Hadamard differentiable (or compact differentiable) at $\theta \in D_\theta$ if there exists $\rho'_\theta \in L_1(D_\theta, W)$ such that for any compact set Γ of D_θ,*

$$\lim_{t\to 0} \left\| \frac{\rho(\theta + th_t) - \rho(\theta)}{t} - \rho'_\theta(h) \right\|_W \to 0, \tag{4}$$

for every $h_t \to h$, s.t. $\theta + th_t \in D_\theta$ for $t > 0$ and t small enough.

One may write instead of (4):

$$Rem_1(\hat{\theta_n} - \theta_0) = \rho(\hat{\theta}_n) - \rho(\theta_0) - (\hat{\theta}_n - \theta_0)^T \rho'_\theta(\theta_0) \tag{5}$$

as the remainder term of (3). Note that ρ'_θ indicates the Hadamard derivative. It is shown by [Fernholz, 1983], that if

$$\lim_{\theta\to\theta_0} \frac{Rem_1(\theta - \theta_0)}{|\theta - \theta_0|} = 0,$$

then: $\sqrt{n} Rem_1(\theta - \theta_0) = o_P(1)$. DQM can be defined using (3):

Definition 2 *$\rho(\theta)$ is said to be DQM if there exists a function $D\rho_\theta \in L^2(P)$, s.t.*

$$\lim_{\theta\to\theta_0} \frac{\|\rho(\theta) - \rho(\theta_0) - (\theta - \theta_0)^T D\rho_\theta(\theta_0)\|_2}{\|\theta - \theta_0\|} = 0 \tag{6}$$

Once again we write for (3):

$$Rem_2(\hat{\theta_n} - \theta_0) = \rho(\hat{\theta}_n) - \rho(\theta_0) - (\hat{\theta}_n - \theta_0)^T D\rho_\theta(0) \tag{7}$$

We now state the result:

Theorem 1 *Suppose $\rho(\theta)$ is Hadamard differentiable w.r.t. θ in a neighborhood N_θ of θ_0, $\sup E[|\rho(\theta)|] < \infty$ and $\sup E[\rho^2(\theta)] < \infty$ for $\theta \in N_\theta$ and $\hat{\theta}_n$ is a consistent estimate for θ_0, then $\rho(\theta)$ is also differentiable in quadratic mean, and both derivative coincide on N_θ.*

Proof Write $t = \hat{\theta}_n - \theta_0$, $\theta = \theta_0$ and $h_t \to 0$ in (4), so we can rewrite (4):

$$\lim_{\hat{\theta}_n \to \theta_0} \left\| \frac{\rho(\hat{\theta}_n) - \rho(\theta_0) - (\hat{\theta}_n - \theta_0)^T \rho'_\theta(0)}{\hat{\theta} - \theta_0} \right\|_W = 0 \tag{8}$$

Because $\hat{\theta}_n \xrightarrow{P} \theta_0$ is a Cauchy sequence we have for $\hat{\theta}_n \in N_\theta$, that $|\hat{\theta}_n - \theta_0| \to 0$ and we get expression (6). We now have to show, that both remainder terms converge in a L^2 sense, as well as in probability to zero. Then from (3) we can deduce that $Rem_1 = Rem_2$ a.s. and therefore both derivatives must coincide a.s.. From the assumptions imposed on ρ we can conclude that ρ is uniformly integrable (see corollary 6.22 in [Klenke, 2006]). If $\hat{\theta}_n \xrightarrow{P} \theta_0$ we have that $\sqrt{n} Rem_1(\hat{\theta}_n - \theta_0) = o_P(1)$ and $\sqrt{n} Rem_2(\hat{\theta}_n - \theta_0) \xrightarrow{L^2} 0$. As $\rho(\theta) \in L^2(P)$ it follows with continuity of $\rho(\theta)$ w.r.t. θ that $h = |\rho'_\theta| \in L^2$, where ρ'_θ is the Hadamard derivative. Therefore we have $Rem_1(\hat{\theta}_n - \theta_0) \in L^2$ for n large enough. We now have: Rem_1 is uniformly integrable, and $Rem_1 \xrightarrow{P} 0$ and thus

$$|\sqrt{n} Rem_1(\hat{\theta}_n - \theta_0)|^2 \overset{n \to \infty}{\to} |0|^2$$

Write $g = |\sqrt{n} Rem_1(\hat{\theta}_n - \theta_0) - 0|^2$, then $g_n \xrightarrow{P} 0$ and g_n is uniformly integrable, as

$$g_n = |\sqrt{n} Rem_1(\hat{\theta}_n - \theta_0)|^2 \leq 4 \cdot |o_P(1)|^2.$$

Therefore $g_n \in L^2(P)$ and thus $D\rho_\theta = \rho'_\theta$ a.s. in N_θ.

□

3 Conclusion

We proved the equivalence of Hadamard differentiability and differentiability in quadratic mean for M estimates usually considered in robust statistics. This result can be used to deduce alternative assumptions imposed on M estimators to be consistent and asymptotically normally distributed.

References

[Fernholz, 1983] Fernholz, L. (1983). *Von Mises calculus for statistical functionals.* New York: Springer-Verlag.

[Klenke, 2006] Klenke, A. (2006). *Wahrscheinlichkeitstheorie*. Berlin: Springer-Verlag.

[LeCam, 1970] LeCam, L. (1970). On the assumptions used to prove asymptotic normality of maximum likelihood estimates. *The Annals of Statistics*, 41, 802–828.

[Pollard, 1997] Pollard, D. (1997). Another look at differentiability in quadratic mean. In *Festschrift for Lucien Le Cam: Research papers in probability and statistics* (pp. 305–314). Springer-Verlag.

[van der Vaart, 1998] van der Vaart, A. (1998). *Asymptotic statistics*. Cambridge: Cambridge University Press.

Diskussionspapiere 2010
Discussion Papers 2010

01/2010 **Mosthaf, Alexander, Schnabel, Claus and Stephani, Jens:** Low-wage careers: Are there dead-end firms and dead-end jobs?

02/2010 **Schlüter, Stephan and Matt Davison:** Pricing an European Gas Storage Facility using a Continuous-Time Spot Price Model with GARCH Diffusion

03/2010 **Fischer, Matthias, Gao, Yang and Herrmann, Klaus:** Volatility Models with Innovations from New Maximum Entropy Densities at Work

04/2010 **Schlüter, Stephan, Deuschle, Carola:** Using Wavelets for Time Series Forecasting – Does it Pay Off?

05/2010 **Feicht, Robert, Stummer, Wolfgang:** Complete closed-form solution to a stochastic growth model and corresponding speed of economic recovery.

06/2010 **Hirsch, Boris, Schnabel, Claus:** Women Move Differently: Job Separations and Gender.

07/2010 **Gartner, Hermann, Schank, Thorsten, Schnabel, Claus:** Wage cyclicality under different regimes of industrial relations.

Diskussionspapiere 2009
Discussion Papers 2009

01/2009 **Addison, John T. and Claus Schnabel:** Worker Directors: A German Product that Didn't Export?

02/2009 **Uhde, André and Ulrich Heimeshoff:** Consolidation in banking and financial stability in Europe: Empirical evidence

03/2009 **Gu, Yiquan and Tobias Wenzel:** Product Variety, Price Elasticity of Demand and Fixed Cost in Spatial Models

04/2009 **Schlüter, Stephan:** A Two-Factor Model for Electricity Prices with Dynamic Volatility

05/2009 **Schlüter, Stephan and Fischer, Matthias**: A Tail Quantile Approximation Formula for the Student t and the Symmetric Generalized Hyperbolic Distribution

06/2009 **Ardelean, Vlad:** The impacts of outliers on different estimators for GARCH processes: an empirical study

07/2009 **Herrmann, Klaus:** Non-Extensitivity versus Informative Moments for Financial Models - A Unifying Framework and Empirical Results

08/2009 **Herr, Annika:** Product differentiation and welfare in a mixed duopoly with regulated prices: The case of a public and a private hospital

09/2009 **Dewenter**, Ralf, **Haucap**, Justus and **Wenzel**, Tobias: Indirect Network Effects with Two Salop Circles: The Example of the Music Industry

10/2009 **Stuehmeier**, Torben and **Wenzel**, Tobias: Getting Beer During Commercials: Adverse Effects of Ad-Avoidance

11/2009 **Klein**, Ingo, **Köck**, Christian and **Tinkl**, Fabian: Spatial-serial dependency in multivariate GARCH models and dynamic copulas: A simulation study

12/2009 **Schlüter**, Stephan: Constructing a Quasilinear Moving Average Using the Scaling Function

13/2009 **Blien,** Uwe, **Dauth,** Wolfgang**, Schank,** Thorsten and **Schnabel,** Claus**:** The institutional context of an “empirical law”: The wage curve under different regimes of collective bargaining

14/2009 **Mosthaf**, Alexander, **Schank**, Thorsten and **Schnabel**, Claus: Low-wage employment versus unemployment: Which one provides better prospects for women?

Diskussionspapiere 2008
Discussion Papers 2008

01/2008 **Grimm, Veronika and Gregor Zoettl**: Strategic Capacity Choice under Uncertainty: The Impact of Market Structure on Investment and Welfare

02/2008 **Grimm, Veronika and Gregor Zoettl**: Production under Uncertainty: A Characterization of Welfare Enhancing and Optimal Price Caps

03/2008 **Engelmann, Dirk and Veronika Grimm**: Mechanisms for Efficient Voting with Private Information about Preferences

04/2008 **Schnabel, Claus and Joachim Wagner**: The Aging of the Unions in West Germany, 1980-2006

05/2008 **Wenzel, Tobias**: On the Incentives to Form Strategic Coalitions in ATM Markets

06/2008 **Herrmann, Klaus**: Models for Time-varying Moments Using Maximum Entropy Applied to a Generalized Measure of Volatility

07/2008 **Klein, Ingo and Michael Grottke:** On J.M. Keynes' “The Principal Averages and the Laws of Error which Lead to Them” - Refinement and Generalisation