Bootstrapping of M-smoothers in Homoscedastic and Heteroscedastic Models

Matúš Maciak♯

Charles University in Prague, Faculty of Mathematics and Physics

Department of Probability and Mathematical Statistics, Sokolovská 83, Prague, Czech Republic

Models & Motivation

- observations $\{(X_i, Y_i);\ i = 1, \ldots, N\}$ for some $N \in \mathbb{N}$ form a random sample from $F(x, y)$;
- conditional dependence of Y given X modelled by $m(x)$ under nonparametric assumptions;
- under the given variance $\Rightarrow$ two possibilities:

$$Y_i = m(X_i) + \sigma\varepsilon_i,\ i = 1, \ldots, N;$$

Homoscedasticity

- for some unknown variance parameter $\sigma^2 > 0$;

$$Y_i = m(X_i) + \sigma(X_i)\varepsilon_i,\ i = 1, \ldots, N;$$

Heteroscedasticity

- for some common variance function $\sigma(\cdot) > 0$;
- for ε_i's to be an i.i.d. sequence of random variables with a distribution function $G(\cdot)$;

Change-point problem:
We assume a model with possible discontinuities (change-points) up the the order $p \in \mathbb{N}$, where p is an order of local polynomial approximation!

Local Polynomial M-smoothers with Change-points problem.

Local polynomial estimation methods with robust approach...
while taking into account all order discontinuities up to the order of approximation.

Local Asymptotics

- asymptotic normality derived under some common assumptions related to nonparametric regression and M-estimates theory;
- asymptotically unbiased M-smoother estimate of the unknown regression function $m(x)$, for some $x \in [0, 1]$, where we assume the domain of interest for X to be the interval $[0, 1]$;
- asymptotic normality derived also for all order derivatives $m^{(\nu)}(x)$ of the unknown regression function $m(\cdot)$, for some $x \in [0, 1]$ and $\nu \in \{1, 2, \ldots, p\}$;

$$\sqrt{Nh_N} \cdot \left|\hat{m}^{(\nu)}(x) - m^{(\nu)}(x)\right| \xrightarrow{\mathsf{D}} \mathsf{N}\left(0, \frac{c^2 \mathsf{E}\psi^2(\epsilon)}{f(x)} \cdot \mathbf{e}_{\nu+1}^{\top} \mathbf{h}_N{}^{\top} \mathbf{S}_1^{-1}\mathbf{S}_2\mathbf{S}_1^{-1}\mathbf{h}_N \mathbf{e}_{\nu+1}\right)$$

Homoscedasticity

$$\sqrt{Nh_N} \cdot \left|\hat{m}^{(\nu)}(x) - m^{(\nu)}(x)\right| \xrightarrow{\mathsf{D}} \mathsf{N}\left(0, \frac{c_x^2 \mathsf{E}\psi^2(\epsilon \cdot \sigma(x))}{f(x)} \cdot \vec{e}_{\nu+1}^{\top} \vec{h}_N{}^{\top} \mathbf{S}_1^{-1}\mathbf{S}_2\mathbf{S}_1^{-1}\vec{h}_N \vec{e}_{\nu+1}\right)$$

Heteroscedasticity

Under the notation:

- where $c^{-1} = \lambda'_G(0)$ is a common term coming naturally form an M-estimates theory and c_x^{-1} respectively is its analogy for heteroskedastic case given by $c_x^{-1} =_x \lambda'_G(0) = -\left[\frac{\partial}{\partial t}\int_{-\infty}^{\infty} \psi(e \cdot \sigma(x) - t)\mathrm{d}G(e)\right]_{\{t=0\}}$.
- moreover, $\mathbf{S}_1$ and $\mathbf{S}_2$ being $(p+1) \times (p+1)$ matrices with elements $s_{i,j} = \int_{-1}^{1} u^{i+j}K(u)\mathrm{d}u$ in case of $\mathbf{S}_1$ and $s^*_{i,j} = \int_{-1}^{1} u^{i+j}K^2(u)\mathrm{d}u$ in case of $\mathbf{S}_2$, for $i, j = 0, 1, \ldots, p$, where $p \in \mathbb{N}$ is a degree of local polynomial approximation;
- vectors $\vec{e}_\nu$, $\vec{h}_N$ respectively, stands for a vector with zeros everywhere and a one on its ν^{th} place and $(1, h_N^{-1}, \ldots, h^{-p})$;

Application to a hypothesis testing problem ...
about the occurrence of a change-point at some $x_0 \in [0, 1]$.

Bootstrapping of M-smoothers

- for the test about the change-point occurrence at some $x_0 \in [0, 1]$ the test statistics $T_N(x_0) = \sqrt{Nh_N} \cdot \left|\hat{m}_+^{(\nu)}(x_0) - \hat{m}_-^{(\nu)}(x_0)\right|$ follows the same distribution (up to the additive constant) as the one derived above;
- however, this distribution heavily depends on some unknown quantities $\Rightarrow$ bootstrap approximation for the unknown distribution;

Bootstrapping algorithms:

a) compute (estimate) residuals $\hat{e}_i = Y_i - \hat{m}(X_i)$, for $i = 1, \ldots, N$;
b) centering of residuals: $\widetilde{e}_i = \hat{e}_i - \frac{1}{N}\sum_i \hat{e}_i$;
c) resample new residuals e_i^* from $\{\widetilde{e}_i, \ldots, \widetilde{e}_N\}$ with replacement, for $i = 1, \ldots, N$.
d) apply a bias-correction procedure where $e_i^* \leftarrow e_i^* + a_N \cdot Z_i$, such that Z_i's are independent centered random variables and a_N is a bias-corrected bandwidth;
e) define new data points (X_i, Y_i^*), where $Y_i^* = \hat{m}(X_i) + e_i^*$;
f) use new data to compute the test statistics $T_N(x_0)$ → repeat for a sufficient number of times to mimic the unknown distribution;

Homoscedasticity

a) compute residuals $\hat{e}_i = Y_i - \hat{m}(X_i)$, for $i = 1, \ldots, N$;
b) compute standardized residuals as $\widetilde{e}_i = \frac{\hat{e}_i - \frac{1}{N}\sum_j \hat{e}_j}{\sqrt{\frac{1}{N}\sum_j(\hat{e}_j - \frac{1}{N}\sum_l \hat{e}_l)^2}}$;
c) similarly to the previous case, let $e_i^* = \widetilde{e}_i + a_N \cdot Z_i$;
d) form a new data set (X_i, Y_i^*), where $Y_i^* = \hat{m}(X_i) + \hat{\sigma}(X_i) \cdot e_i^*$, where $\hat{\sigma}(\cdot)$ is a nonparametric estimate of the variance function under the considered model;
e) use the new data set to compute the test statistic → repeat to mimic the distribution under the interest;

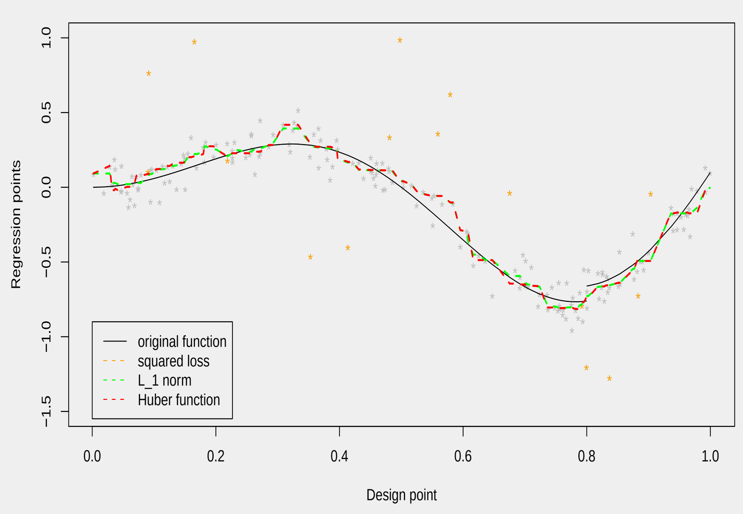
Heteroscedasticity

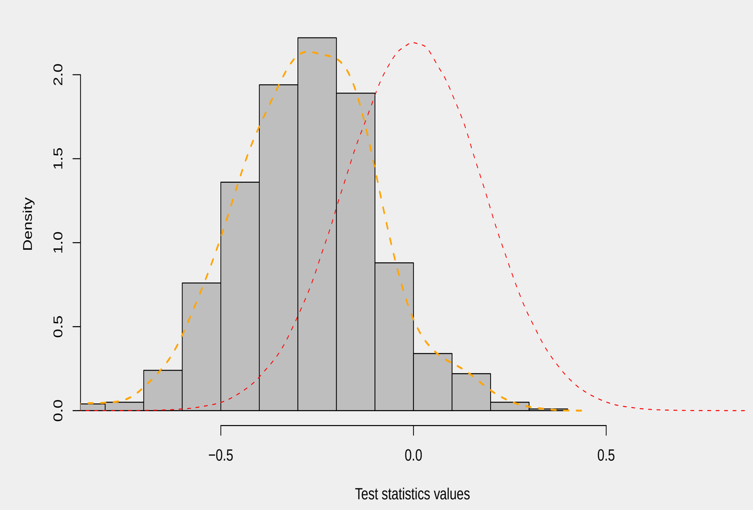
Local testing procedure for the regression function itself.

An extensive simulation study proposed ...
in order to see how the algorithms behave.

Simulations

Remarks, Acknowledgement & Bibliography





RESULTS from the simulations	Loss function used for minimization	$T_N(x_0 = 0.3)$ → slope: 0.338	$T_N(x_0 = 0.5)$ → slope: -3.24	$T_N(x_0 = 0.8)$ → change-point
Homoscedasticity	*Squared loss*	0.340 *(0.364)*	−0.495 *(0.397)*	0.373 *(0.307)*
	L_1 norm	0.183 *(0.232)*	−0.196 *(0.299)*	0.262 *(0.153)*
	Huber function	0.182 *(0.228)*	−0.191 *(0.279)*	0.275 *(0.182)*
Heteroscedasticity	*Squared loss*	0.387 *(0.369)*	−0.395 *(0.380)*	0.403 *(0.335)*
	L_1 norm	0.194 *(0.240)*	−0.236 *(0.319)*	0.312 *(0.188)*
	Huber function	0.190 *(0.235)*	−0.211 *(0.322)*	0.305 *(0.190)*

Considered for the test statistic $T_N(x_0) = \sqrt{Nh_N} \cdot (\hat{m}_+(x_0)) - \hat{m}_-(x_0))$ for a data set with $10\ \%$ of outlying observations.

Results

Bibliography:

[1] Bickel P.J. and Freedman D.A., Some Asymptotic Theory for the Bootstrap; *The Annals of Statistics*, (1981), Vol.9, No.6, 1196-1217.
[2] Freedman D.A., Bootstrapping Regression Models; *The Annals of Statistics*, (1981), Vol.9, No.6, 1218-1228.
[3] Neumeyer N., Bootstrap Procedures for Empirical Processes of Nonparametric Residuals; *Habilitationsschrift*, Ruhr-Universität (2006).

Acknowledgement:
The author would like to express many thanks for support provided to him from the grant GAČR 201/09/0755.

♯ http://www.mmatthew.matfyz.cz maciak@karlin.mff.cuni.cz