Erratum to: Weighted Halfspace Depths and Their Properties (dissertation thesis)

Lukáš Kotík

March 22, 2015

Page 64, lines 1 - 19: There exists a set of elementary events $A_{r+K} \in \mathcal{A}$ such that $\mathsf{P}(A_{r+K}) = 1$ and such that the following properties hold for any $\omega \in A_{r+K}$ (subscript emphasises the dependence on r, K from previous paragraph). Additional parenthesis with ω symbol in the text that follows denotes the dependence on ω.

1. There exists $n_1(\omega) \in \mathbb{N}$ such that $\sup_{\boldsymbol{x} \in C_K} |\mathrm{D}_{w,n}(\boldsymbol{x})(\omega) - \mathrm{D}_w(\boldsymbol{x})| < \varepsilon$ for all $n \geq n_1(\omega)$.
2. There exists $n_2(\omega) \in \mathbb{N}$ such that $P_{\mathrm{n}}(C^c)(\omega) < 2\varepsilon$ for all $n \geq n_2(\omega)$ since $P(C^c) < \varepsilon$. This follows from the *Glivenko-Cantelli* theorem applied to the random sample $\|\boldsymbol{X}_1\|, \|\boldsymbol{X}_2\|, \ldots, \|\boldsymbol{X}_n\|, \ldots$.

Further as the weight function is considered bounded, say $w(\boldsymbol{x}, \boldsymbol{u}) \leq b$, it follows that for all $\boldsymbol{x} \in C_K^c$ it holds that

$$\mathrm{D}_{w,n}(\boldsymbol{x})(\omega) \leq b\, P_{\mathrm{n}}(C^c)(\omega) + \varepsilon,$$
$$\mathrm{D}_w(\boldsymbol{x}) \leq b\, P(C^c) + \varepsilon.$$

Therefore there exists some constant B which does not depend on ε such that it holds that for all $n \geq n_2(\omega)$

$$\sup_{x \in C_K^c} \big|\mathrm{D}_{w,n}(\boldsymbol{x})(\omega) - \mathrm{D}_w(\boldsymbol{x})\big| \leq b\Big(P_{\mathrm{n}}(C^c)(\omega) + P(C^c)\Big) + 2\varepsilon < B\varepsilon. \tag{3.13}$$

So eventually, it holds that for all $n \geq n_0(\omega) = \max\{n_1(\omega), n_2(\omega)\}$

$$\begin{aligned}\sup_{\boldsymbol{x} \in \mathbb{R}^p} \big|\mathrm{D}_{w,n}(\boldsymbol{x})(\omega) - \mathrm{D}_w(\boldsymbol{x})\big| \\ &\leq \max\left\{\sup_{\boldsymbol{x} \in C_K} \big|\mathrm{D}_{w,n}(\boldsymbol{x})(\omega) - \mathrm{D}_w(\boldsymbol{x})\big|, \sup_{\boldsymbol{x} \in C_K^c} \big|\mathrm{D}_{w,n}(\boldsymbol{x})(\omega) - \mathrm{D}_w(\boldsymbol{x})\big|\right\} \\ &< B\varepsilon.\end{aligned}$$

The set A_{r+K} depends on choice of the sets C (radius r) and C_K (radius $r + K$) and hence on the value of ε. For any $\varepsilon > 0$ there exist $r(\varepsilon), K(\varepsilon) \in \mathbb{N}$ which satisfy the preceding lines of the proof. Denote $l(\varepsilon) = \min\{r(\xi) + K(\xi) : \ \xi \geq \varepsilon\}$. Let us define a set of elementary events

$$A = \bigcap_{\varepsilon > 0} A_{l(\varepsilon)}.$$

It is clear that there exists a series of natural numbers $l_1, l_2, l_3, \ldots$ such that

$$A = \bigcap_{i \in \mathbb{N}} A_{l_i}.$$

It follows $\mathsf{P}(A) = 1$ and also that this set does not depend on a choice of C and C_K.

We have proved so far: there exists a set A, $\mathsf{P}(A) = 1$, such that for any $\omega \in A$ it holds that $\forall \varepsilon > 0 \ \exists n_0(\omega), \ n > n_0(\omega) \Rightarrow \sup_{\boldsymbol{x} \in \mathbb{R}^p} \big|\mathrm{D}_{w,n}(\boldsymbol{x})(\omega) - \mathrm{D}_w(\boldsymbol{x})\big| < \varepsilon$. These lines finishes the proof of uniform almost sure convergence over $\mathbb{R}^p$, (3.4).

Page 65, lines 24 - 25:

$$\mathcal{H}_2 = \{\boldsymbol{x} : \ \exists \boldsymbol{u}_{\boldsymbol{x}} \in \mathcal{S}^p, \ \mathsf{E}\, w(\boldsymbol{X} - \boldsymbol{x}, \boldsymbol{u}_{\boldsymbol{x}}) = 0 \text{ and } \mathsf{E}\, w(\boldsymbol{X} - \boldsymbol{x}, -\boldsymbol{u}_{\boldsymbol{x}}) \geq \delta\},$$

where $\delta > 0$ is a constant.

Page 67, lines 11 - 17: Since w is positive then from the definition of the set $\mathcal{H}_2$ it follows that for any $\boldsymbol{x} \in \mathcal{H}_2$ it holds that

$$\frac{1}{n} \sum_{i=1}^{n} w(\boldsymbol{X}_i - \boldsymbol{x}, \boldsymbol{u}_{\boldsymbol{x}}) = 0 \text{ a.s.}$$

Further from Theorem 17 and Lemma 29 it follows that

$$\begin{aligned} &\left| \inf_{\boldsymbol{x} \in \mathcal{H}_2} \frac{1}{n} \sum_{i=1}^{n} w(\boldsymbol{X}_i - \boldsymbol{x}, -\boldsymbol{u}_{\boldsymbol{x}}) - \inf_{\boldsymbol{x} \in \mathcal{H}_2} \mathsf{E}\, w(\boldsymbol{X} - \boldsymbol{x}, -\boldsymbol{u}_{\boldsymbol{x}}) \right| \\ &\leq \sup_{\boldsymbol{x} \in \mathcal{H}_2} \left| \frac{1}{n} \sum_{i=1}^{n} w(\boldsymbol{X}_i - \boldsymbol{x}, -\boldsymbol{u}_{\boldsymbol{x}}) - \mathsf{E}\, w(\boldsymbol{X} - \boldsymbol{x}, -\boldsymbol{u}_x) \right| \\ &\leq \sup_{\boldsymbol{x} \in \mathbb{R}^p} |\mathrm{D}_{w,n}(\boldsymbol{x}) - \mathrm{D}_w(\boldsymbol{x})| \xrightarrow{n \to \infty} 0 \quad \text{a.s.} \end{aligned}$$

Since $\mathsf{E}\, w(\boldsymbol{X} - \boldsymbol{x}, -\boldsymbol{u}_x) \geq \delta, \ \forall \boldsymbol{x} \in \mathcal{H}_2$, then it follows

$$\sup_{\boldsymbol{x} \in \mathcal{H}_2} \mathrm{RD}_n(\boldsymbol{x}) \leq \sup_{\boldsymbol{x} \in \mathcal{H}_2} \widehat{\mathrm{RD}}_n(\boldsymbol{x}, \boldsymbol{u}_{\boldsymbol{x}}) \xrightarrow{n \to \infty} 0 \text{ a.s.}$$

Page 73, last 3 lines: Since

$$\mathsf{E}\, \sqrt{n} R_n = \sqrt{n}\big(\mathsf{E}\, \mathrm{D}_n(\boldsymbol{x}) - \mathrm{D}_w(\boldsymbol{x})\big)$$

one has (by using Markov's inequality) that (3.21) holds if

$$\sqrt{n}\big(\mathsf{E}\, \mathrm{D}_n(\boldsymbol{x}) - \mathrm{D}_w(\boldsymbol{x})\big) \xrightarrow{n \to \infty} 0. \tag{3.22}$$

Page 75, lines 13 - 17: For any $k \in \mathbb{N}$ and $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k \in \mathcal{S}^p$ it holds that

$$\begin{pmatrix} Y^n_{\boldsymbol{u}_1} \\ \vdots \\ Y^n_{\boldsymbol{u}_k} \end{pmatrix} \xrightarrow[\text{in Law}]{n\to\infty} \begin{pmatrix} Y_{\boldsymbol{u}_1} \\ \vdots \\ Y_{\boldsymbol{u}_k} \end{pmatrix} \sim \mathcal{N}(\mathbf{0}, \mathbf{R}),$$

where $R_{ij} = R(\boldsymbol{u}_i, \boldsymbol{u}_j)$. It follows directly from the *Multivariate Central Limit Theorem.* If w is not continuous on $\mathbb{R}^p \times \mathcal{S}^p$ the trajectories $\boldsymbol{u} \mapsto Y^n_{\boldsymbol{u}}(\omega)$ can be modified to be continuous on $\mathcal{S}^p$. To show

$$Y^n \xrightarrow[\text{in Law}]{n\to\infty} Y.$$

we need to check that for each $\eta, \varepsilon > 0$ there exists $\delta > 0$ such that

$$\limsup_{n\to+\infty} \mathsf{P}(\sup_{\measuredangle(\boldsymbol{u},\boldsymbol{v})<\delta} |Y^n_{\boldsymbol{u}} - Y^n_{\boldsymbol{v}}| > \eta) < \varepsilon$$

(Stochastic equicontinuity - see e.g. Theorem 10.2 in *Pollard, D. (1990). EMPIRICAL PROCESSES: THEORY AND APPLICATIONS. IMS.*) Using Chebyshev's inequality stochastic equicontinuity holds if

$$\lim_{\delta\to 0+} \mathsf{var}\left(\sup_{\measuredangle(\boldsymbol{u},\boldsymbol{v})<\delta} \frac{1}{n} \sum_{i=1}^{n} \Big(w(\boldsymbol{X}_i - \boldsymbol{x}, \boldsymbol{u}) - w(\boldsymbol{X}_i - \boldsymbol{x}, \boldsymbol{v}) \Big) \right) = 0.$$

This condition usually holds for absolutely continuous distributions if a reasonable weight function is chosen. For hypothetical use of the proposed asymptotic distribution is more of interest the covariance structure $R(\boldsymbol{u}, \boldsymbol{v})$ of the limit process than the technical apparatus behind the proof of tightness. Hence the proof of the stochastic equicontinuity is not shown in this note.

I would like to thank Ing. Marek Omelka, Ph.D. for comments that lead to improvement of this erratum.