ON TRIMMED POLYA ALGORITHM

Jüri Lember

Abstract. Let $(E, \|\cdot\|)$ be a separable Banach space, $L_\infty(\Omega, E)$ the corresponding Lebesgue-Bochner space of essentially bounded random elements. Let $X \in L_p(\Omega, E)$. We consider the trimmed p-predictions - the best L_p-approximations of X by the constants, obtained by the trimmed procedure. The behavior of the trimmed p-predictions as $p \to \infty$ is investigated. This is the setup of Polya algorithm. We show that the (trimmed) L_p-distance from X to E does not always converge to the (trimmed) L_∞-distance. The topological conditions in terms of E for such a convergence are presented. We prove that any sequence of trimmed p-predictions converges to the set of best L_∞-approximations, if E has uniform ν-Kadec-Klee property with respect to a suitable topology ν.

Key words: Kadec-Klee property, Polya algorithm, trimming.

1 Introduction

Generally, the Polya algorithm is the construction of a best L_∞-approximation as the limit of best L_p-approximations, as $p \to \infty$. This is a quite general approach and several approximation problems fall within this framework. Often the L_∞-approximation of real-valued functions has been considered: the papers (Darst, 81) and (Al-Rashed, 83) deal with the approximation of essentially bounded random variables by the functions that are measurable with respect to a sub σ-algebra; in (Darst, *et al.* 83; Cuesta, Matran, 89b) the sub σ-algebra is generalized by the sub σ-lattice; the Polya algorithm in the space of continuous bounded function is studied in (Legg, Towswend, 89; Li, 95). The abovementioned papers consider the best approximation of the elements of certain function-space by the elements of a subspace. The approximation (and corresponding Polya algorithm) by the elements of a convex set that is not necessary a subspace is carried out in (Egger, Huotari, 90; Huotari, Li, 94).

In the present paper we consider the Polya algorithm in the following settings. Let $(E, \|\cdot\|)$ be a separable real Banach space, $(\Omega, \mathcal{F}, \mathbf{P})$ a probability space and $L_\infty(\Omega, E)$ the corresponding Lebesgue-Bochner space. We investigate the approximations of the functions in $L_\infty(\Omega, E)$ by the elements of E. So, in a sense, the approximation setup is the simplest - the subspace of possible approximations consists of constant-valued functions. The complexities connected with such kind of approximation are caused by the general nature of E and P. The Polya algorithm corresponding to such setup was studied in (Darst *et al.*, 83). They considered an uniformly convex Banach space E and showed the convergence of Polya algorithm. However, as pointed out in (Cuesta, Matran, 89b), the proof of it was not satisfactory. The complete proof of the convergence of Polya algorithm for uniformly convex E was given in (Cuesta, Matran, 87). Moreover, in this paper an useful connection between the approximations of $X \in L_\infty(\Omega, E)$ by means of constants and the approximation of X by the elements of $L_\infty(\Omega, \mathcal{G}, \mathbf{P}, E)$ with $\mathcal{G}$ being a sub σ-algebra of $\mathcal{F}$ was established - a best L_p-approximation of regular conditional probability measure of X given $\mathcal{G}$ by constants is a version of a best L_p-approximation of X by the elements of $L_\infty(\Omega, \mathcal{G}, \mathbf{P}, E)$. This relation shows that in many cases the approximation by the constants is not very restrictive - the corresponding results for the approximation from the subspace $L_\infty(\Omega, \mathcal{G}, \mathbf{P}, E)$ can be easily deduced.

For now on, let $X \in L_\infty(\Omega, E)$, $X \sim P$, $V_p(a) := \|X - a\|_p$ and

$$V_p := \inf_{a \in E} V_p(a), \quad p \leq \infty. \tag{1.1}$$

The elements minimizing (1.1) are called p-predictions, the quantity V_p will be referred to as p-variance. The ∞-predictions are the Chebyshev centers of the support of P and ∞-variance is usually called the Chebyshev radius (of the support of P). The set of p-predictions ($p \leq \infty$) will be denoted by $\mathcal{P}_p$.
In this paper we investigate the following questions:

1) When

$$V_p \to V_{p_o} \quad \text{as} \quad p \to p_o? \tag{1.2}$$

2) What is the limit $\lim_{p \to p_o} V_p$, if the convergence in 1) fails?

3) When $a_n \to \mathcal{P}_{p_o}$, if $p_n \to p_o$, $\epsilon_n \to 0$, $V_{p_n}(a_n) \leq V_{p_n} + \epsilon_n$?

As mentioned previously, the case $p_o = \infty$ is of our special interest. However, we briefly consider also the case $p_o < \infty$.
At the first sight, (1.2) might seem obvious. However, it turns out that (1.2) does not always hold and this fact justifies the question 2).
Note that convergence of the sequence $\{a_n\}$ to the set $\mathcal{P}_{p_o}$ can occur in the following ways:

a) $a_n \to a \in \mathcal{P}_{p_o}$.
When this holds and $p_o = \infty$, then the Polya algorithm converges. If E is convex, then the p-predictions (when they exists) are unique. This needs not hold for the Chebyshev centers (unless E is uniformly convex) and, therefore, the limit of Polya algorithm is a distinguished element of $\mathcal{P}_{p_o}$. This particular element is often called the best best approximation.

b) the sequence $\{a_n\}$ is relatively compact with all cluster points in $\mathcal{P}_{p_o}$.
If the limit set is unique, then this means the convergence of $\{a_n\}$. The all convergence results in this paper are of this type.

c) $d(a_n, \mathcal{P}_{p_o}) \to 0$, i.e. the distance of a_n from to $\mathcal{P}_{p_o}$ goes to 0.
Since the limit set $\mathcal{P}_{p_o}$ is bounded, for a finite-dimensional E this type of convergence coincides with the type b). However, as pointed out in a remark after Corollary 5.2, for a infinite dimensional E, this type of convergence might be "too general".

We do not impose any restrictions on P, and we investigate the stated questions in terms of topological and geometrical properties of E. Moreover, we aim to consider all questions in a more general framework of convergent metrics (section 2) as well as in terms of certain type of loss functions (sub-sections 4.3, 4.4). It turns out that the sequence $\{a_n\}$ ($p_o \leq \infty$) minimizes the loss function of type (4.5), and the convergence results for such (minimizing) sequences now apply (see e.g. Lember, 2000).

The (impartial) trimming procedure is a known method to robustify location parameter like p-prediction (see e.g. Gordaliza, 91b; Cuesta, Matran, 97). In order to carry out the trimming procedure we use the trimming functions (for P) introduced in (Gordaliza, 91a). Let $\delta \in [0,1)$. A trimming function for P at level δ is a measurable map $\tau : E \mapsto [0,1]$ satisfying $\int \tau(x)P(dx) \geq 1-\delta$. The trimmed p-prediction at level δ is any solution of the minimization problem

$$\inf_{\tau} \inf_{a\in E} \frac{1}{\int \tau(x)P(dx)} \int \tau(x)\|x-a\|^p P(dx),$$

where the minimization with respect to τ is over all possible set of trimming functions at level δ. The trimming at level 0 yields the usual sense of p-predictions. The trimmed Chebyshev centers are defined analogously. Having the trimmed p-predictions and trimmed p-variances, we investigate the questions 1), 2) and 3) also for $\delta > 0$.

The paper is organized as follows. In section 2 we consider some very basic best approximation principles in connection with convergent metrics and norms. We observe that if p_o, ∞ then (1.2) follows from corresponding results of convergent metrics (Lemma 2.1).
In Section 3 we introduce some preliminaries about trimmings. We also generalize Lemma 2.1 for case $\delta > 0$ and show its relation with the generalities in Section 2.
In Section 4 we give the answer to the question 2) (Lemma 4.1) and we study the topological conditions to guarantee the convergence 1). It turns out that the latter holds if the loss function (4.5) attains its minimum on E. This allows to consider the question 2) in terms of (a wide class of) loss functions. It is known that (4.5) attains its minimum on E if E admits a vector space topology ν such that every $\|\cdot\|-$closed ball is sequentially ν-compact (condition **A** in the sequel). When this holds, the (non-trimmed) p-predictions always exists; the question of the existence of trimmed p-predictions under **A** is still open.
In Section 5 we study the question 3) under the assumption that E satisfies **A**. We show that for $p_o < \infty$ the convergence in 3) holds, if E has ν-Kadec-Klee property, i.e. the ν-convergence in the unit sphere of E yields strong convergence (Theorem 5.2). If $p_o = \infty$, then 3) holds under a slightly stronger condition: E has uniform ν-Kadec-Klee property (Theorem 5.1). Since every uniformly convex space has uniform weak Kadec-Klee property, our result generalizes the (non-trimmed) convergence result in (Cuesta, Matran, 87).

2 Some preliminaries about convergent metrics

In this section we represent some general principles about convergent metrics and norms.

2.1 Metrics

Let S be a set, $G \subset S$. Suppose d_n, d are the metrics on S satisfying $d_n \to d$ (i.e. $d_n(x,y) \to d(x,y)\ \forall x, y \in S$).

We consider an arbitrary $x \in S,\ x \notin G$ and an arbitrary sequence of ϵ_n-optimal best d_n-approximations $\{g_n\}$, where $g_n \in \mathcal{P}_n^{\epsilon_n} := \{g \in G : d_n(x,g) \leq d_n(x,G) + \epsilon_n\}$ and $\epsilon_n \to 0$. We are interested in the following convergences:

$$d_n(x,G) \to d(x,G); \tag{2.1}$$

$$d(g_n, x) \to d(x,G). \tag{2.2}$$

For L_p-distances, (2.1) and (2.2) means $V_{p_n} \to V_p$ and $\|X - a_n\|_p$, as $p_n \to p$, respectively. When (2.2) holds then $\{g_n\}$ is called to be d-minimizing. This obviously implies that $\{g_n\}$ is bounded with respect to the d-metric and any cluster point of $\{g_n\}$ belongs to $\mathcal{P} := \{g \in G : d(x,g) = d(x,G)\}$.

At first note that

$$\limsup_n d_n(x,G) \leq d(x,G)$$

and, hence, the sequence $\{d_n(x,g_n)\}$ is bounded. When the sequence of metrics d_n tends to d from above, i.e. $d_n \searrow d$, then (2.1) and (2.2) holds. In terms of L_p-metrics, $p_n \searrow p$ yields (1.2). However, in the following we mostly concentrate on the convergence $p_n \nearrow p \leq \infty$.

The following proposition turns out to be useful.

Proposition 2.1. *When*

$$\lim_n \sup_G \{|\, d(x,g) - d_n(x,g)\,|\} = 0 \tag{2.3}$$

then (2.1) and (2.2) holds.

Proof. The relation (2.3) means that for every sequence of ϵ_n-optimal best approximations $\{g_n\}$, $|\, d(x,g_n) - d_n(x,g_n)\,| \to 0$. Thus $\limsup_n d(x,g_n) = \limsup_n d_n(x,g_n)$ and the inequalities

$$d(x,G) \leq \limsup_n d(x,g_n) = \limsup_n d_n(x,g_n) = \limsup d_n(x,G) \leq d(x,G)$$

yield (2.2). Now (2.1) immediately follows:

$$d(x,G) = \lim_n d(x,g_n) = \lim_n d_n(x,g_n) = \lim_n d_n(x,G). \qquad \triangle$$

Suppose now that S is a linear space, $G \subset S$ a subspace, $\|\cdot\|_n \to \|\cdot\|$ convergent norms in S. As previously, we consider a fixed $x \notin G$.

Since G is a subspace, the relation (2.3) generally fails. However, when the sequence of best approximations $\{g_n\}$ is bounded, then, as the following proposition shows,the supreme in (2.3) can be replaced by the supreme over the unit ball of G.

Proposition 2.2. *Let G be a subspace. Assume that $\forall y \in S$*

$$\lim_n \sup\Big\{\big|\|y-g\|_n - \|y-g\|\big| : \|g\| \le 1\Big\} = 0. \tag{2.4}$$

Then we have (2.1) and (2.2).

Proof. Take $y = 0$. Then $\lim_n \sup\Big\{\big|\|g\|_n - \|g\|\big| : \|g\| \le 1\Big\} = 0$. This implies that

$$\sup_{g\in G} \Big| \frac{\|g\|_n}{\|g\|} - 1 \Big| \to 1$$

or $\|g_n\| \sim \|g_n\|_n$. Since $\|x - g_n\|_n$ is bounded, the same holds for $\|g_n\|_n$ and, hence, $\{g_n\}$ is bounded in $\|\cdot\|$-metrics. Let this bound be m.
Since (2.4) holds for each $y \in S$, it is not hard to see that in (2.4) the constant 1 can be replaced by m. Now use Proposition 2.1 with $G_m = \{g : \|g\| \le m\}$ in place of G. △

Corollary 2.1. *Suppose G is finite dimensional and $\|\cdot\|_n \nearrow \|\cdot\|$. Then (2.1) and (2.2) hold. Moreover, $g_n \to \mathcal{P}$ in the sense of b).*

Proof. Since $\|\cdot\|_n \nearrow \|\cdot\|$, the norms $\|\cdot\|_n$ are continuous on $(E, \|\cdot\|)$. By finite dimensionality of G, the uniform convergence in (2.4) clearly holds. Proposition 2.2 now establishes (2.1) and (2.2). Since $\{g_n\}$ us bounded in G, it is relatively compact with all cluster points in $\mathcal{P}$ (by (2.2)). △

Remark. Because of the finite-dimensionality of G, the continuity of $\|\cdot\|_n$ holds even if the convergence of norms is not monotone. Thus the monotone convergence is not necessary in Corollary 2.1. This is Kripke's theorem (e.g. Holmes, p.119).

2.3 L_p-norms

Consider now Lebesgue-Bochner spaces. The next proposition proves (2.4) provided $p_n \to p < \infty$.

Proposition 2.3. *Let $p_o < \infty$, $Y \in L_{p_o}(\Omega, \Re)$. Then $\forall X \in L_{p_o}(\Omega, E)$ the convergence $p_n \to p$, where $p_n \in [1, p_o]$ yields*

$$\sup_{Z\in\mathcal{Y}} \big|\|X - Z\|_p - \|X - Z\|_{p_n}\big| \to 0, \quad \text{where} \quad \mathcal{Y} := \{Z \in L_p(\Omega, E) : \|Z(\omega)\| \le Y(\omega)\}.$$

Proof. W.l.o.g. we assume $Y(\omega) \geq 1$.
Let us estimate

$$\sup_{Z\in\mathcal{Y}} |\|X-Z\|_p^p - \|X-Z\|_{p_n}^{p_n}| \leq \sup_{Z\in\mathcal{Y}} \int |\|X(\omega)-Z(\omega)\|^p - \|X(\omega)-Z(\omega)\|^{p_n}| d\mathbf{P} \leq$$
$$\sup_{Z\in\mathcal{Y}} \Big\{ \int_{\{\|X(\omega)-Z(\omega)\|\leq 1\}} |\|X(\omega)-Z(\omega)\|^p - \|X(\omega)-Z(\omega)\|^{p_n}| d\mathbf{P} +$$
$$\int_{\{\|X(\omega)-Z(\omega)\|>1\}} |\|X(\omega)-Z(\omega)\|^p - \|X(\omega)-Z(\omega)\|^{p_n}| d\mathbf{P} \Big\} \leq$$
$$\sup_{Z\in\mathcal{Y}} \int_{\{\|X(\omega)-Z(\omega)\|\leq 1\}} |\|X(\omega)-Z(\omega)\|^p - \|X(\omega)-Z(\omega)\|^{p_n}| d\mathbf{P} +$$
$$+ \sup_{Z\in\mathcal{Y}} \int_{\{\|X(\omega)-Z(\omega)\|>1\}} |\|X(\omega)-Z(\omega)\|^p - \|X(\omega)-Z(\omega)\|^{p_n}| d\mathbf{P}.$$

When $x \in [0,1]$ and $p > q$, then $|x^p - x^q| = x^q - x^p \leq \left(\frac{q}{p}\right)^{\frac{q}{p-q}} - \left(\frac{q}{p}\right)^{\frac{p}{p-q}} := t(p,q) \to 0$, provided $q \to p$ or $p \to q$.
When $x \in [1,k]$ and $p > q$, then $|x^p - x^q| = x^p - x^q \leq k^p - k^q \to 0$, provided $q \to p$.
Hence

$$\sup_{Z} \int_{\{\|X(\omega)-Z(\omega)\|\leq 1\}} |\|X(\omega)-Z(\omega)\|^p - \|X(\omega)-Z(\omega)\|^{p_n}| d\mathbf{P} \leq t(p,p_n) \to 0.$$

For each $Z \in \mathcal{Y}$, denoting $u_n = \max\{p, p_n\}$ and $v_n = \min\{p, p_n\}$ we now have

$$\big|\|X(\omega)-Z(\omega)\|^p - \|X(\omega)-Z(\omega)\|^{p_n}\big| \leq \big(\|X(\omega)\| + Y(\omega)\big)^{u_n} - \big(\|X(\omega)\| + Y(\omega)\big)^{v_n},$$

provided $\|X(\omega) - Z(\omega)\| > 1$.
Therefore,

$$\int_{\{\|X(\omega)-Z(\omega)\|>1\}} |\|X(\omega)-Z(\omega)\|^p - \|X(\omega)-Z(\omega)\|^{p_n}| d\mathbf{P} \leq$$
$$\int [(\|X(\omega)\| + Y(\omega))^{u_n} - (\|X(\omega)\| + Y(\omega))^{v_n}] d\mathbf{P}.$$

For last inequality recall the assumption $Y(\omega) \geq 1 \ \forall \omega$.
Hence

$$\sup_{Z\in\mathcal{Y}} \int_{\{\|X(\omega)-Z(\omega)\|>1\}} |\|X(\omega)-Z(\omega)\|^p - \|X(\omega)-Z(\omega)\|^{p_n}| d\mathbf{P} \leq \int g_n(\omega) d\mathbf{P},$$

where $g_n(\omega) := (\|X(\omega)\| + Y(\omega)^{u_n} - (\|X(\omega)\| + Y(\omega))^{v_n} \to 0$. Clearly $g_n(\omega) \to 0, \forall \omega \in \Omega$. Since $g_n(\omega) \leq (\|X(\omega)\| + Y(\omega))^{p_o}$ and, by assumption, $\|X\|, Y \in L_p(\Omega, \Re)$, by dominated convergence we have $\int g_n d\mathbf{P} \to 0$.
Thus

$$\sup_{Z\in\mathcal{Y}} |\|X-Z\|_p^p - \|X-Z\|_{p_n}^{p_n}| \to 0.$$

The claim now follows easily. △

Corollary 2.2. *Let* $p_o < \infty$, $X \in L_{p_o}(\Omega, E)$, $p_n \to p$, $p_n \in [1, p_o]$. *Then,* $\forall m < \infty$

$$\sup\Big\{\big|\|X-a\|_p - \|X-a\|_{p_n}\big| : a \in E, \|a\| \le m\Big\} \to 0.$$

Moreover, from the proof we get

$$\sup_{\|a\|\le m} \int \big|\, \|x-a\|^p - \|x-a\|^{p_n} \,\big| P(dx) \to 0. \tag{2.5}$$

We can now immediately deduce the convergence of p-variances and the minimizing property (2.2) in bounded interval. Let $\mathcal{P}^{\epsilon}_{p_o}$ stand for the set of ϵ-optimal p-predictions.

Lemma 2.1. *Let* $p_o < \infty$, $\epsilon_n \to 0$. *Then the following statements hold*
i) the mapping $p \mapsto V_p$ *is continuous on the set* $[1, p_o]$;
ii) if $p_n \to p$, $p_n \in [1, p_o]$, *then for each sequence* $\{a_n\}$, *satisfying* $a_n \in \mathcal{P}^{\epsilon_n}_{p_n}$, *we have*

$$\|X - a_n\|_p \to V_p. \tag{2.6}$$

iii) $\exists K < \infty : \|a\| < K, \quad \forall a \in \mathcal{P}^{\epsilon}_p, \quad \forall p \in [1, p_o]$.

Proof. i) ja ii) follow from Proposition 2.2.
iii) holds, because any minimizing sequence is bounded. The constant K can be estimated: for all $a \in \mathcal{P}^{\epsilon}_p$, $p \in [1, p_o]$: $\|a\| \le \|X\|_{p_o} + V_{p_o} + \epsilon < \infty$.

Note that Lemma 2.1 holds for general E. For uniformly convex E, the statements of Lemma 2.1 were presented in (Cuesta, Matran, 89a). Their proof does not apply for general E. We aimed to show how the claims of Lemma 2.1 represent some general principles of convergent metrics.

Remark. Suppose $X \in L_p(\Omega, E)$, $p \in [1, p_o)$, $p_o < \infty$. Then an obvious generalization of ii) holds: $V_p \nearrow \infty$ as $p \nearrow p_o$

3 Trimming

We now represent some basics about (impartial) trimming. This concept was introduced in (Gordaliza, 91a). For notation and terminology we follow (Gordaliza, 91a; Cuesta, Matran, 97). In the end of this section we prove the trimmed version of Corollary 2.2 as well as that of Lemma 2.1. We also give a shorter proof for the existence of trimmed p-predictions for finite-dimensional E.

Let $\delta \in [0, 1)$. If δ is 0, it will be skipped in the notation.
At first, let us define the set of trimming functions $\tau^{\delta-}$.

$$\tau^{\delta} := \{\tau : E \to [0,1], \text{measurable}, \int \tau(x) dP = 1 - \delta, \}, \quad \tau^{\delta-} := \cup_{\beta \le \delta} \tau^{\beta}.$$

Let $\tau \in \tau^\delta$, $1 \le p < \infty$ and define the trimmed p-loss-functions

$$V_p^\tau(a) := (\frac{1}{1-\delta}\int \tau(x)\|x-a\|^p P(dx))^{\frac{1}{p}}, \quad V_\infty^\tau(a) := \|\tau(X)(X-a)\|_\infty.$$

For all $1 \le p \le \infty$ and $\tau \in \tau^{\delta-}$, let

$$V_p^\tau := \inf_a V_p^\tau(a), \quad V_p^\delta(a) := \inf_{\tau\in\tau^{\delta-}} V_p^\tau(a),$$

$$V_p^\delta := \inf_{\tau\in\tau^{\delta-}} V_p^\tau = \inf_a V_p^\delta(a). \tag{3.1}$$

If τ^* and a^* are solutions of (3.1), they are called L_p-best trimming function at level δ and (impartial) trimmed p-prediction at level δ, respectively.
The set of trimmed p-predictions at level δ will be denoted by $\mathcal{P}_p^\delta$, the set of ϵ-optimal p-predictions will be denoted by $\mathcal{P}_p^{\epsilon\delta}$.
When $\delta > 0$, $V_p^\delta < \infty$. In the following we do not assume that P has bounded support. Therefore, if not specified, we also allow the case $V_\infty = \infty$.

Define the trimmed essential supreme (trimmed radius) of $\|X-a\|$:

$$r^\delta(a) =: \|X-a\|_\infty^\delta := \inf\{r : P(\|x-a\| > r) \le \delta\}.$$

Note that $r^\delta(a)$ could be represented as

$$r^\delta(a) = \inf\{\|I_{\overline{B}(a,r)}(X)(X-a)\|_\infty : I_{\overline{B}(a,r)} \in \tau^{\delta-}\}. \tag{3.2}$$

The following proposition shows the obvious relation between trimmed ∞-loss function and trimmed radius.

Proposition 3.1. $V_\infty^\delta(a) = r^\delta(a)$.

Proof. By (3.2), we have

$$V_\infty^\delta(a) = \inf_{\tau\in\tau^{\delta-}} \|\tau(X)(X-a)\|_\infty \le \inf\{\|I_{\overline{B}(a,r)}(X)(X-a)\|_\infty : I_{\overline{B}(a,r)} \in \tau^{\delta-}\} = r^\delta(a).$$

On the other hand, for any $\tau \in \tau^{\delta-}$, we obviously have

$$r := \|\tau(X)(X-a)\|_\infty = \|I_{\overline{B}(a,r)}(X)(X-a)\|_\infty \tag{3.3}$$

where $I_{\overline{B}(a,r)} \in \tau^{\delta-}$. Hence,

$$\inf_{\tau\in\tau^{\delta-}} \|\tau(X)(X-a)\|_\infty = \inf\{\|I_{\overline{B}(a,r)}(X)(X-a)\|_\infty : I_{\overline{B}(a,r)} \in \tau^{\delta-}\} = r^\delta(a)$$

and

$$V_\infty^\delta(a) = r^\delta(a) = \|X-a\|_\infty^\delta = \|I_{\overline{B}(a,r^\delta(a))}(X)(X-a)\|_\infty. \qquad \triangle \tag{3.4}$$

We now describe the trimming functions minimizing $V_p^\delta(a)$ over $\tau^{\delta-}$ (a is fixed). Let

$$\tau^\delta(a) := \{\tau \in \tau^\delta : I_{B(a,r^\delta(a))} \le \tau \le I_{\overline{B}(a,r^\delta(a))} \quad P-a.s.\}.$$

Thus, $\tau^\delta(a)$ consists of trimming functions that have mass $1-\delta$ and may differ of the indicator of $\overline{B}(a, r^\delta(a))$ on the boundary of $\overline{B}(a, r^\delta(a))$, only. Note that $\tau^\delta(a)$ does not depend on p. The following proposition, proved in (Gordaliza, 91a), shows that for a fixed a and finite p, each element of $\tau^\delta(a)$ minimizes the criterion $V_p^\delta(a)$ over all $\tau \in \tau^{\delta-}$. Moreover, among all the functions in τ^δ, only the functions form $\tau^\delta(a)$ can achieve $V_p^\delta(a)$.

Proposition 3.2. *Let* $p \in [1, \infty)$ *and* $\tau \in \tau^\delta$. *Then* $V_p^\tau(a) = V_p^\delta(a)$ *iff* $\tau \in \tau^\delta(a)$.

When $p = \infty$, then the set of trimming functions, that minimize $V_\infty^\tau(a)$ over τ^δ can be bigger than $\tau^\delta(a)$. The following proposition gives a description of this set.
Define

$$\tau_\infty^\delta(a) := \{\tau \in \tau^\delta : \tau \leq I_{\overline{B}(a, r^\delta(a))} \quad P - a.s.\}.$$

Proposition 3.3. *Let* $\tau \in \tau^\delta$. *Then* $V_\infty^\tau(a) = V_\infty^\delta(a)$ *iff* $\tau \in \tau_\infty^\delta(a)$.

Proof. The claim of the proposition is already mentioned in (Gordaliza, 91a; Cuesta, Matran, 98).
Let $\tau \in V_\infty^\tau(a)$. Since $\tau \leq I_{\overline{B}(a, r^\delta(a))}$, from (3.4) we get

$$\|\tau(X)(X - a)\|_\infty \leq \|I_{\overline{B}(a, r^\delta(a))}(X)(X - a)\|_\infty = V_\infty^\delta(a). \tag{3.5}$$

By Proposition 1, the inequality in (3.5) is, in fact, equality. This proves the 'if' part.
Suppose $\tau \in \tau^\delta$ is such that $V_\infty^\tau(a) = V^\delta(a) = r^\delta(a)$. By (3.3), this means $\tau \leq I_{B(a, r^\delta(a))}$, i.e. $\tau \in V_\infty^\tau(a)$. △

Suppose $p < \infty$. Proposition 3.2 does not specify, whether there exist trimming functions that minimize $V_p^\tau(a)$ but do not belong to $\tau^\delta(a)$. Suppose τ is such a function. By 'only if' part of Proposition 3.2, the total mass of τ must be strictly bigger than $1 - \delta$. Let us denote it via $1 - \beta > 1 - \delta$. By 'if' part of Proposition 3.2, $\tau \in \tau^\beta(a)$ and $V_p^\beta(a) = V_p^\delta(a)$. As showed in (Gordaliza, 91a), this equality holds iff $r^\delta(a) = r^\beta(a)$, and $P(B(a, r^\delta(a))) = 0$.
Suppose a^* is a trimmed p-prediction at level δ and $\tau^* \in \tau^\delta(a^*)$ is one of the best trimming functions for a^* (such functions depend on p via a^*, only). Since $V^{\tau^*}(a^*) = V_p^\delta$, the trimmed p-prediction a obviously minimizes $V_p^{\tau^*}(\cdot)$ over E. Hence, there is a dual relationship between best trimming functions and trimmed p-predictions - each trimmed p-prediction is a center of a ball that essentially defines the corresponding (the best) trimming function: $B(a^*, r^\delta(a^*))$, and in the same time the trimmed p-prediction minimizes the p-loss function over this ball. See also (Gordaliza, 91a; Cuesta, Matran, 98).
There might exist other points having the described duality property. For example, if P is a probability measure in $\Re$ consisting on three atoms: $P(0) = 0.5$, $P(-1) = P(1) = 0.25$, then the best trimmed 2-predictions at level 0.25 are, obviously, -0.5 and 0.5 with corresponding radiuses $r^{0.25}(-0.5) = r^{0.25}(0.5) = 0.5$. On the other hand, the point 0 also has the above mentioned duality property ($r^{0.25}(0) = r(0) = 1$), although 0 is not a 2-prediction at level 0.25.

Suppose τ minimizes V_∞^τ over $\tau^{\delta-}$, and the total mass of τ is bigger than $1 - \delta$, i.e. $\tau \in \tau^\beta$ for some $\beta < \delta$. Then, by Proposition 3.3, $\tau \in \tau_\infty^\beta$, and $r^\beta = r^\delta$. Hence, the set of trimming functions minimizing $V_\infty^\tau(a)$ over $\tau^{\delta-}$ is

$$\tau_\infty^{\delta-}(a) := \{\tau \in \tau^{\delta-} : \tau \leq I_{\overline{B}(a, r^\delta(a))} \quad P - a.s.\}$$

For example, $I_{\overline{B}(a, r^\delta(a))}$ always belongs to $\tau_\infty^{\delta-}(a)$ even if it does not belong to $\tau_\infty^\delta(a)$.

We now generalize Corollary 2.2 for trimming.

Corollary 3.1. *Assume the hypotheses of Corollary 2.2. Then, for each* $\delta \in [0,1)$ *we have*

$$\sup_{\|a\|\le m} \left|V_{p_n}^{\delta}(a) - V_p^{\delta}(a)\right| \to 0. \tag{3.6}$$

Proof. Let $a \in E$. Then, by Proposition 3.2,

$$|(V_p(a))^p - (V_{p_n}(a))^{p_n}| = |\int (\|x-a\|^p - \|x-a\|^{p_n})\tau(x)P(dx)| \le$$
$$\int |\|x-a\|^p - \|x-a\|^{p_n}|\tau(x)P(dx) \le \int |\|x-a\|^p - \|x-a\|^{p_n}|P(dx),$$

where $\tau \in \tau^{\delta}(a)$. Hence, by (2.5)

$$\sup_{\|a\|\le m} |(V_p^{\delta}(a))^p - (V_{p_n}^{\delta}(a))^{p_n}| \le \sup_{\|a\|\le m} \int |\|x-a\|^p - \|x-a\|^{p_n}|P(dx) \to 0$$

and (3.6) follows. △

Although, for positive trimming ($\delta > 0$) we do not have a corresponding metric for random elements, the claims of Lemma 2.1 still hold.

Lemma 3.1. *Let* $p_o < \infty$, $\delta \in (0,1)$, $\epsilon_n \to 0$ *. Then the following statements hold*
i) the mapping $p \mapsto V_p^{\delta}$ *is non-decreasing and continuous on the set* $[1, p_o]$*;*
ii) if $p_n \to p$, $p_n \in [1, p_o]$*, then for each sequence* $\{a_n\}$*, satisfying* $a_n \in \mathcal{P}_{p_n}^{\delta\epsilon_n}$*, we have*

$$V_p^{\delta}(a_n) \to V_p^{\delta}. \tag{3.7}$$

iii) $\exists K < \infty : \|a\| < K, \quad \forall a \in \mathcal{P}_p^{\delta\epsilon}, \quad \forall p \in [1, p_o].$

Proof. The monotonicity - $V_p^{\delta}(a) \le V_q^{\delta}(a)$, if $p \le q$ - follows from the monotonicity of L_p norm together with Proposition 3.2. Now, the same for the mapping $p \mapsto V_p^{\delta}$ follows.
iii) Note that, for each $a \in E$ and $\tau \in \tau^{\delta}(a)$, it holds

$$V_p^{\tau}(a) \le \frac{V_p(a)}{(1-\delta)^{\frac{1}{p}}}.$$

Also note that $\forall \tau \in \tau^{\delta-}$, $\forall a, b \in E$ we have the analogies to the triangle inequalities

$$V^{\tau}(a) \le V^{\tau}(b) + \|a-b\| \quad \text{and} \quad \|a\| \le V_p^{\tau}(a) + V_p^{\tau}(0) \tag{3.8}$$

Now, for an arbitrary $a \in \mathcal{P}_p^{\delta\epsilon}$ with corresponding $\tau \in \tau^{\delta}(a)$ we get

$$\|a\| \le V_p^{\tau}(a) + V_p^{\tau}(0) \le V_p^{\delta} + \epsilon + \frac{\|X\|_p}{(1-\delta)^{\frac{1}{p}}} \le V_{p_o}^{\delta} + \epsilon + \frac{\|X\|_{p_o}}{(1-\delta)^{\frac{1}{p}}} =: K.$$

The proofs of i) and ii) go along the same line as in Proposition 2.1. Let $\{a_n\}$ be an arbitrary sequence of ϵ_n-optimal trimmed p_n-predictions at level δ. We know that $\{a_n\}$ is bounded by K. Hence, by (3.6) $\lim_n |V^\delta_{p_n}(a_n) - V^\delta_p(a_n)| \to 0$. Now proceed as in the proof of Proposition 2.1 to obtain i) and ii). $\triangle$

The sequences satisfying (3.7) are called minimizing (for V^δ_p). When such a sequence converges, the limit belongs to $\mathcal{P}^\delta_p$. Indeed, suppose $\{a_n\}$ satisfies (3.7), $a_n \to a$. Let $\tau_n \in \tau^\delta(a_n)$. Now, $V^\delta_p(a) \le V^{\tau_n}_p(a) \le V^{\tau_n}_p(a_n) + \|a_n - a\| = V^\delta_p(a_n) + \|a_n - a\| \to V^\delta_p$. The first inequality follows from the definition of $V^\delta_p(a)$ and the second inequality follows from (3.8).
This observation gives a short proof of the existence of trimmed p-prediction for finite-dimensional E. Originally the existence of trimmed p-predictions for Euclidean space was proved (in Gordaliza, 91a). Our proof here uses triangle inequalities (3.8) and is much shorter.

4 The function t^δ and limits

In this section we study the limit $\lim_{p\to\infty} V^\delta_p$. We show, that generally the limit is not $V^\delta_\infty = r^\delta$, and we study the conditions that guarantee $\lim_{p\to\infty} V^\delta_p = r^\delta$

4.1 Limit $\lim_{p\to\infty} V^\delta_p$

By Proposition 3.1,

$$V^\delta_\infty = \inf_a r^\delta(a) =: r^\delta.$$

The quantity r^δ can be interpreted as the trimmed Chebyshev radius. We now introduce another quantity, which in many cases is equal to r^δ. Let

$$\delta(t) := \inf_{a\in E} P\{x : \|x - a\| > t\}, \quad t^\delta := \inf\{t : \delta(t) \le \delta\}.$$

At first note that

$$r^\delta \ge t^\delta, \quad \forall\delta. \tag{4.1}$$

We allow the case $t^\delta = \infty$. Observe that $t^\delta < \infty$ yields $r^\delta \le \infty$.
Also observe that for $\delta > 0$, (4.1) would be an equality, if in the definition of t^δ the "$\le \delta$" requirement were replaced by "$< \delta$". We shall show that t^δ is the limit of V^δ_p as $p \nearrow \infty$.
For each $a \in E$, $\delta \in [0, 1)$, it holds $V^\delta_p(a) \nearrow V^\delta_\infty(a)$.
Indeed, take an arbitrary $\tau \in \tau^\delta(a)$. By (3.4) we know that $\|\tau(X)(X - a)\|_\infty = r^\delta(a) = V^\delta_\infty(a)$ and, therefore

$$V^\delta_p(a) = V^\tau_p(a) = \Big(\frac{1}{1-\delta}\int \tau(X)\|X - a\|^p d\mathbf{P}\Big)^{\frac{1}{p}} \nearrow \|\tau(X)(X - a)\|_\infty = V^\delta_\infty(a).$$

Although, $V^\delta_p(a) \to r^\delta(a)$, $\forall a$, the following lemma shows that the same needs not hold for the minimums.

Lemma 4.1. *For each $\delta \in [0,1)$, it holds*

$$V_p^\delta \to t^\delta \quad \text{if} \quad p \to \infty. \tag{4.2}$$

Proof. Fix an arbitrary $\delta \le 1$, and $\epsilon > 0$. By definition of t^δ, $\inf_a P\{x : \|x-a\| > t^\delta - \epsilon\} \ge \delta + \gamma$, for some $\gamma > 0$. Fix now $a \in E$ and $\tau \in \tau^\delta(a)$. Then, for $p < \infty$.

$$(1-\delta)\big(V_p^\delta(a)\big)^p = \int \tau(x)\|x-a\|^p P(dx) \ge (t^\delta - \epsilon)^p \int_{\{\|x-a\|>t^\delta-\epsilon\}} \tau(x) P(dx) \ge (t^\delta-\epsilon)^p \gamma,$$

because

$$P\{x : \|x-a\| \le t^\delta - \epsilon\} + \int_{\{\|x-a\|>t^\delta-\epsilon\}} \tau(x)P(dx) = 1-\delta \quad \text{and}$$

$$P\{x : \|x-a\| \le t^\delta - \epsilon\} \le 1 - \gamma - \delta.$$

Let now $p_n \to \infty$, $\epsilon_n \to 0$, $a_n \in \mathcal{P}_{p_n}^{\delta\epsilon_n}$, $\tau_n \in \tau^\delta(a_n)$. Then,

$$V_{p_n}^\delta + \epsilon_n \ge V_{p_n}^{\tau_n}(a_n) \ge \Big(\frac{\gamma}{1-\delta}\Big)^{\frac{1}{p_n}} (t^\delta - \epsilon) \to t^\delta - \epsilon. \tag{4.3}$$

The other side. Since $\delta(t^\delta + \epsilon) = 0$, there exists a sequence $\{b_n\}$ such that $\lim_n P(A_n) \to \delta$, where $A_n = \{x : \|x-b_n\| > t^\delta + \epsilon\}$. Then $P(B_n) \to 0$, where

$$B_n := \{x : t^\delta + \epsilon < \|x-b_n\| \le r_n(b_n)\}.$$

Let $\tau_n \in \tau^\delta(b_n)$. Recall that by definition of $\tau^\delta(b_n)$,

$$\tau_n \le I_{\overline{B}(b_n, r^\delta(b_n))} \quad P-\text{a.s.}.$$

It is easy to see that $\{b_n\}$ is bounded. Then, obviously, $r^\delta(b_n)$ is bounded as well. When $\delta > 0$, then the boundedness holds even if P does not posses a bounded support. Now, for each p,

$$\begin{aligned}(1-\delta)(V_p^\delta(b_n))^p &= \int_{A_n} \tau_n(x)\|x-b_n\|^p P(dx) + \int_{A_n^c} \tau_n(x)\|x-b_n\|^p P(dx) \\ &\le \int_{A_n} \tau_n(x)\|x-b_n\|^p P(dx) + (t^\delta+\epsilon)^p P(A_n^c) \\ &\le \int_{B_n} \|x-b_n\|^p P(dx) + (t^\delta+\epsilon)^p P(A_n^c) \\ &\le r^\delta(b_n)P(B_n) + (t^\delta+\epsilon)^p P(A_n^c) \to (t^\delta+\epsilon)^p(1-\delta).\end{aligned}$$

Consequently, $V_p^\delta(b_n) \to t^\delta + \epsilon$, as $n \to \infty$, implying that $V_p^\delta \le t^\delta + \epsilon$. Thus, if $p_n \to \infty$, then for each $\epsilon > 0$, $t^\delta - \epsilon \le \liminf_n V_{p_n}^\delta \le \limsup_n V_{p_n}^\delta \le t^\delta + \epsilon$. △

Corollary 4.1. *Let P have a bounded support. Then $V_p \nearrow t$, where $t = \inf\{r : \delta(r) = 0\}$.*

4.2 The mappings $\delta \mapsto r^\delta$ and $\delta \mapsto t^\delta$

Let us investigate the mapping $\delta \mapsto r^\delta(a)$, $a \in E$. Clearly this mapping is non-increasing and lower semi-continuous. Hence, for each $a \in E$, the convergence $\delta \searrow \delta_o$ yields $r^\delta(a) \nearrow r^{\delta_o}(a)$. Again, the same property needs not hold when considering the minimal values. The following property shows, that the limits are, in fact, t^{δ_o}.

Proposition 4.1. *Let $\delta_o \in [0,1)$. Then*

$$\lim_{\delta \searrow \delta_o} r^\delta = t^{\delta_o}. \tag{4.4}$$

Proof. Let $t > t^{\delta_o}$. Then $\inf_a P(\|x-a\| > t) \le \delta_o$. Hence, for each $\delta > \delta_o$, we can find a $a \in E$ such that $P(\|x-a\| > t) \le \delta$. By definition, $r^\delta \le t$. Hence, $\lim_{\delta \searrow \delta_o} r^\delta \le t^{\delta_o}$.
On the other hand, let $t < t^{\delta_o}$. Then $\inf_a P(\|x-a\| > t) > \delta_o$. Hence, for each $\delta < \delta_o$, $a \in E$, $P(\|x-a\| > t) > \delta$, implying that $r^\delta(a) > t$ and $r^\delta \ge t$. Hence, $\lim_{\delta \searrow \delta_o} r^\delta \ge t$, implying that $\lim_{\delta \searrow \delta_o} r^\delta \ge t^{\delta_o}$. △

Corollary 4.2. *The functions $\delta \mapsto t^\delta$ and $\delta \mapsto r^\delta$ differ at most countable set of points in $(0,1]$.*

Proof. If $\delta \mapsto r^\delta$ is right-continuous at δ_o, then, by Proposition 4.2, $r^{\delta_o} = t^{\delta_o}$. A non-decreasing function has at most countable set of discontinuity points. △

Corollary 4.3. *Let P have the bounded support. Then $r^\delta \to t$, if $\delta \searrow 0$.*

If we refer to the limit $\lim_{\delta \nearrow 0} V_\infty^\delta$ as "zero-trimming", then Propositions 3.1 and 4.1 state that the zero-trimming can be much better (in the sense of minimizing the loss-function V_∞) than no trimmings, i.e. $t < r$.

Example. Let $E = c_o$, $P = \sum_n \delta_{e_n} p_n$, $p_n > 0 \ \forall n$, $X \sim P$. Here, e_n is the sequence that has 0-s everywhere but 1 in n-th place.
Let us calculate the functions $\delta \mapsto r^\delta$ and $\delta \mapsto t^\delta$.
We start with the function $\delta(t)$. At first, convince that $\delta(0.5) = 0$. Consider the sequence $a_m = (a_i^m)$, $m = 1, 2, \ldots$, where $a_i^m = \frac{1}{2}$, if $i \le m$ and $a_i^m = 0$ else. Now

$$\|e_n - a_m\| = \begin{cases} \frac{1}{2}, & \text{if } n \le m \\ 1 & \text{if } n > m. \end{cases}$$

Hence

$$P\{x \| \|x - a_m\| > \frac{1}{2}\} = P\{x \| \|x - a_m\| = 1\} = \sum_{n > m} p_n \to 0, \quad \text{as} \quad m \to \infty$$

i.e. $\delta(0.5) = 0$.
Let $t = 0.5 - \epsilon$. It is easy to see that if there exists e_n such that $\|e_n - a\| < 0.5 - \epsilon$, then $\|e_k - a\| > 0.5 + \epsilon$ for each $k \ne n$. Therefore

$$\delta(0.5 - \epsilon) = \inf_{a \in E} P\{e_n : \|e_n - a\| > 0.5 - \epsilon\} = 1 - \max_n p_n.$$

Hence

$$\delta(t) = \begin{cases} 1 - \max_n p_n & \text{if } t \in [0, \frac{1}{2}) \\ 0, & \text{if } t \in [\frac{1}{2}, 1]. \end{cases}$$

Thus, the function $\delta \mapsto t^\delta$ is the following

$$t^\delta = \begin{cases} \frac{1}{2}, & \text{if } t \in [0, 1 - \max_n p_n) \\ 0, & \text{if } t \in [1 - \max_n p_n, 1]. \end{cases}$$

Let us now investigate the trimmed Chebyshev radius r^δ. At first we see that $r = 1$. Indeed, let $\epsilon > 0$ and consider $a = (a_1, a_2, \ldots) \in c_o$. Since $a_n \to 0$ there $\exists$ N: $|a_n| < \epsilon$, when $n > N$. This means $\|e_n - a\| > 1 - \epsilon$, if $n > N$. Consequently, $\|X - a\|_\infty > 1 - \epsilon$. Since a was arbitrary, we get $\inf_{a \in E} \|X - a\|_p \geq 1 - \epsilon$ or $r = 1$.
Finally, let us convince that $r^\delta = t^\delta$, when $\delta > 0$. Let $0 < \delta < 1 - \max_n p_n$. Then there exists n_o such that $\sum_{n > n_o} p_n \leq \delta$. Take $a = (\frac{1}{2}, \ldots, \frac{1}{2}, 0, \ldots)$, where first n_o components are $\frac{1}{2}$. Then, obviously, $r^\delta(a) = \frac{1}{2}$ and the same argument as before shows that $r^\delta = r^\delta(a) = \frac{1}{2}$.
When $\delta \geq 1 - \max_n p_n$, then, obviously $r^\delta(e_{n^*}) = 0$, where $p_{n^*} = \max_n p_n$.

Consequently, by Corollary 4.1, $V_p \nearrow 0.5$, as $p \to \infty$, but $V_\infty = 1$.
In fact, it is easy to see that, $\limsup_n V_{p_n} \leq 0.5$. Let $p < \infty$ be arbitrary and estimate V_p. Taking a_m as previously, we get

$$\|X - a_m\|_p^p = \sum_{n=1}^{\infty} \|e_n - a_m\|^p p_n = \left(\frac{1}{2}\right)^p \sum_{n=1}^{n} p_n + \sum_{n > m} p_n \to \left(\frac{1}{2}\right)^p.$$

This means, that $V_p \leq 0.5$.
Similarly, $V_\infty^\delta \to 0.5$, when $\delta \searrow 0$. △

Remark. This simple example also shows the inconsistency of empirical Chebyshev radiuses. Indeed, for any finite sample from P, the empirical Chebyshev radius r_n satisfies $r_n \leq 0.5$ (here n reflects the sample size) with an inequality iff all elements of the sample are the same. Hence, $r_n \to t = 0.5$ a.s., while $r = 1$. Of course, for any positive amount of trimming, $\delta > 0$, we get consistency of empirical trimmed Chebyshev radiuses, i.e. $r_n^\delta \to r^\delta$ a.s., if $\delta > 0$.
In comparison with the case of the finite p we remark that the convergence $V_p^n \to V_p$ a.s. holds (Lember, 99). Here V_p^n stands for the empirical analogue of V_p.

4.3 Connection with loss-functions

In this section we define a class of loss-functions on E. The loss functions used for non-trimmed p-predictions belongs to this class. We also introduce some generalizations for trimmed case. The (semi)-continuity properties of this type of loss-functions turn out to be important for the existence and for the convergence of p-predictions as well as for the equality $t^\delta = r^\delta$. These questions with related topological assumptions will be studied in the next sections.

Let $\varphi : [0,\infty) \mapsto [0,\infty)$ be a lower semi-continuous non-decreasing discrepancy function. Define a loss function on E

$$W : E \mapsto [0,\infty), \quad W(a) = \int \varphi(\|x-a\|)P(dx) \tag{4.5}$$

(see also Lember, 99). Clearly the function $a \mapsto \big(V_p(a)\big)^p$ ($p < \infty$) is of type (4.5) with $\varphi(\cdot) = (\cdot)^p$. When $p = \infty$, then Chebyshev centers are related to the loss function (4.5) as follows. For each $s > 0$ define a continuous function

$$\phi_s(x) = \begin{cases} 0, & \text{if } x \leq s \\ \text{strictly increasing}, & \text{else.} \end{cases}$$

The loss function of (4.5) with ϕ_s as φ will be denoted by $W_s(a)$. We also denote

$$W_s := \inf_a W_s(a).$$

Now, a Chebyshev Center is a solution a minimizing problem $W_r(a) = W_r$ (see Cuesta, Matran, 88).

We now aim to describe the quantity t in terms of W_s.

Proposition 4.2. $t = \inf\{s : W_s = 0\}$, *where* $\inf \emptyset := \infty$.

Proof. Let $s < t$. Then, for each $a \in E$,

$$W_s(a) \geq \int_{\{\|x-a\|>s_1\}} \phi_s(\|x-a\|)P(dx) \geq \phi_s(s_1)\delta(s_1) > 0,$$

where $s < s_1 < t^\delta$.
So, if $t = \infty$, the Proposition is proved. Let us assume $t < \infty$ and consider $s > t$. Then we can find a sequence a_n such that $P(A_n) \to 0$, where $A_n = \{x : \|x - a_n\| > s\}$. It is not hard to see that the sequence $\{a_n\}$ is bounded. Thus there $\exists K > \infty$ such that $P\{x : \|x - a_n\| < K\} = 1$. This means

$$W_s(a_n) = \int_{A_n} \phi_s(\|x-a_n\|)P(dx) \leq \phi_s(K)P(A_n) \to 0. \qquad \triangle$$

We now consider a straightforward possibility to generalize (4.5) for $\delta > 0$. Define

$$W^\delta : E \to [0,\infty), \quad W^\delta(a) = \frac{1}{1-\delta}\int \tau(x)\varphi(\|x-a\|)P(dx), \quad \tau \in \tau^\delta(a) \tag{4.6}$$

(see, Cuesta, Matran, 98). Again $(\cdot)^p$ for φ yields trimmed p-predictions, any a^* satisfying $W_r^\delta(a^*) = 0$ is a trimmed Chebyshev center at level δ. Here, as previously, W_s^δ stands for a function (4.6) with ϕ_s as the discrepancy function. The minimum of $W_s^\delta(\cdot)$ over E will be denoted by W_s^δ.
Since τ in (4.6) depends on a, (4.6) is not a particular case of (4.5) any more. Still the analogue of Proposition 4.2 holds.

Proposition 4.3. $t^\delta = \inf\{s : W_s^\delta = 0\}$.

Proof. Let $s < t^\delta$ and $s < s_1 < t^\delta$. Then, for each $a \in E$, $\tau \in \tau^\delta(a)$, we get

$$(1-\delta)W_s^\delta(a) \geq \int_{\{\|x-a\|>s_1\}} \tau(x)\phi_s(\|x-a\|)P(dx) \geq \phi_s(s_1)(\delta(s_1)-\delta) > 0,$$

since

$$\int_{\{\|x-a\|>s_1\}} \tau(x)P(dx) = \int \tau(x)P(dx) - P(\|x-a\| \leq s_1) \geq 1-\delta-(1-\delta(s_1)) = \delta(s_1)-\delta(s) > 0.$$

Let $t^\delta < s$. Then there obviously exists a sequence $\{a_n\}$ such that

$$\limsup_n P(\|x-a_n\| > s) \leq \delta \quad \text{or} \quad \liminf_n P(\|x-a_n\| \leq s) \geq 1-\delta.$$

Now

$$\int_{A_n} \tau_n(x)P(dx) \to 0,$$

where $\tau_n \in \tau^\delta(a_n)$ and $A_n = \{x : \|x-a_n\| > s\}$. Indeed, if $s > r$, the statement is obvious. If $s \leq r$, then for each n, $\limsup_n \int_{A_n} \tau_n(x)P(dx) = (1-\delta) - \liminf_n P(\|x-a_n\| \leq s_1) = 0$. Since, again, a_n is bounded, as in the proof of Proposition 4.2 we obtain that for a $K < \infty$ large enough,

$$(1-\delta)W_s(a_n) = \int_{A_n} \tau_n(x)\phi_s(\|x-a_n\|)P(dx) \leq \phi_s(K)\int_{A_n} \tau_n(x)P(dx) \to 0. \qquad \triangle$$

4.4 The equality $t^\delta = r^\delta$ and the existence of p-prediction

Let us now investigate the conditions that guarantee the equality

$$t^\delta = r^\delta. \tag{4.7}$$

When this is the case, the limits (4.2) and (4.4) are, as one might expected, r^δ. We shall see, that this question is related to the existence of p-predictions.
The following proposition, which actually is a straightforward restatement of the definition of t^δ, gives necessary and sufficient condition for (4.7) in a more comfortable form (see also, Cuesta, Matran, 89) . In the following statement, when $r = \infty$, then $r - m$ must be interpreted as a (arbitrary large) finite number.

Proposition 4.4. *Let $\delta \in [0,1)$. Then $t^\delta = r^\delta$ if and only if, for each $m > 0$, there exists $\gamma(m) > 0$, such that*

$$P\{x : \|x-a\| > r^\delta - m\} > \gamma + \delta, \quad \forall a \in E. \tag{4.8}$$

Proof. If (4.8) holds, then $\delta(r_\delta - m) > \delta + \gamma$ for each $m > 0$. Hence, by definition of t^δ, $r^\delta \leq t^\delta \leq r^\delta$.
If, for some $m > 0$, (4.8) does not hold, then there would exists a sequence a_n such that $\limsup_n P\{x : \|x - a_n\| > r^\delta - m\} \leq \delta$. Hence, $\delta(r^\delta - m) \leq \delta$ and $t^\delta \leq r^\delta - m$. △

Let us investigate the conditions that guarantee (4.8). From the definition of t^δ, it is straightforward to see, that it holds, if, for each $\epsilon > 0$,

$$\exists a(\epsilon) \in E : \quad P(\|x - a\| > t^\delta + \epsilon) \leq \delta. \tag{4.9}$$

The latter holds, if for each s, the value $\delta(s)$ is attainable on E. Since $\delta(s)$ is in the form of (4.5) with $\varphi(x) = I_{(s,\infty)}(x)$, we can consider the question of (4.7) in a more general framework. So, the rest of this section deals with the question of the existence of $a^* \in E$ satisfying

$$W(a^*) = \inf_a W(a) \tag{4.10}$$

with W as in (4.5). As noted in the previous section, this question covers also the existence of p-predictions. See also (Herrendorf, 83; Lember 99).

Remark. One can see that (4.7) holds if, for each s, W_s^δ attains its minimum on E. In the case $\delta = 0$, W_s is also the type of (4.5).

The existence of the solution of (4.10) depends on the topological properties of E. Clearly $W(\cdot)$, being lower-semicontinuous, attains its minimum on a compact space. However, the compactness with respect to the metric-topology is too restrictive and, in many cases, it can be replaced by the (sequential) compactness with respect to another, usually weaker topology on E. Of course, the new topology cannot be arbitrary weak, it has to posses some lower-sem-icontinuity properties with respect to the norm $\|\cdot\|$ in E. We now describe the necessary topological conditions more closely.

Let ν be a topological vector space topology on E such that the norm $\|\cdot\|$ is ν-sequentially lower semi-continuous, i.e.

$$\liminf_{x_n \xrightarrow{\nu} x} \|x_n\| \geq \|x\|. \tag{4.11}$$

This holds iff $\|\cdot\|$-closed balls are sequentially ν-closed (see, e.g., Lember, 99). Note that, in general, ν needs not be comparable with the norm-topology in E.
Clearly the weak topology satisfies (4.11), if E is a dual, then $*$-weak topology satisfies (4.11). For other examples see (Khamsi, 96).

Suppose $a_n \xrightarrow{\nu} a$, where ν satisfies (4.11). Then, because of Fatou lemma $W(\cdot)$ is sequentially ν-lower semi-continuous ,i.e.

$$\liminf_{a_n \xrightarrow{\nu} a} W(a_n) \geq W(a). \tag{4.12}$$

Moreover, if, in addition, $\{a_n\}$ minimizes W, i.e. $W(a_n) \to \inf_a W(a)$, then the following result holds (Lember, 2000). From now on, let $S(P)$ denote the support of P.

Proposition 4.5. *Suppose* $a_n \xrightarrow{\nu} a$ *minimizes* W *with a continuous discrepancy function* φ*. Then*

$$\forall x \in S(P), \quad \varphi(\|x - a_n\|) \to \varphi(\|x - a\|) \tag{4.13}$$

If, in addition φ *is strictly increasing, we have*

$$\forall x \in S(P), \quad \|x - a_n\| \to \|x - a\| \tag{4.14}$$

The condition (4.11) holds, if (E, ν) satisfies the following condition

> **A:** every (norm)-closed ball is sequentially ν-compact.

Under **A** the function $W(\cdot)$ attains its minimum on E. Indeed, let $\{a_n\}$ be minimizing for W. Since $\{a_n\}$ is bounded (Lember 2000) and (E, ν) has **A**, $\{a_n\}$ has a ν-convergent subsequence: $a_{n'} \xrightarrow{\nu} a^*$. By (4.12), $W(a^*) = \inf_a W(a)$.

Proposition 4.6. *Suppose* E *admits a topological vector space topology* ν *satisfying* **A**. *Then there exists an element* $a^* \in E$ *satisfying (4.10).*

Corollary 4.4. *Assume* (E, ν) *satisfies* **A**. *Then, for each* $\delta \in (0, 1]$*, we have (4.7) and, for each* $p \leq \infty$*, we have* $\mathcal{P}_p \neq \emptyset$*.*

Note that the foregoing argument does not apply for a loss-function type (4.6), since for such a function, (4.12) has not proved. Therefore, we can not make statements about the existence of non-trimmed p-predictions. In the next section we show that **A** implies the existence of trimmed Chebyshev centres. The question of existence of trimmed p-predictions under **A** [or, equivalently, the question of (4.12) for a loss-function as in (4.6)] is still open.

Corollary 4.5. *If* E *is a dual, then for each* $\delta \in (0, 1]$ *we have (4.7) and* $\mathcal{P}_p \neq \emptyset$*.*

Proof. Since E is dual, by Alaoglu's theorem the unit ball of E is $*$-weak compact. Recall the assumption that E is separable. Then the $*$-weak topology in B_E is induced by a metric, and the unit ball is sequentially compact. Now Corollary 4.5 applies.

From Corollary 4.5 follows that in a reflexive Banach space or, obviously, in finite-dimensional space the equality (4.6) always holds. For reflexive E and $\delta = 0$, Corollary 4.5 was proved in (Cuesta, Matran, 89a).

5 The convergence of p-predictions

In this section we study the convergence of trimmed p_n-predictions as $p_n \to \infty$. Also the convergence of non-trimmed p_n-predictions as $p_n \to p < \infty$ is considered.

Let $p_n \to p_o$, $\epsilon_n \to 0$. We consider the sequence $\{a_n\}$, where

$$\|X - a_n\|_{p_n} \leq V^{\delta}_{p_n} + \epsilon_n. \tag{5.1}$$

Proposition 5.1. *Let $\{a_n\}$ be as in (5.1) with $p_o = \infty$. Then, for each $\epsilon > 0$,*

$$\limsup_n P(\|x - a_n\| > t^{\delta} + \epsilon) \leq \delta. \tag{5.2}$$

Proof. Suppose there exists $\epsilon > 0$ such that (5.2) fails. Then, there exist a $\gamma > 0$ such that along a subsequence, $\{a_{n'}\}$, we have

$$P(\|x - a_{n'}\| > t^{\delta} + \epsilon) > \delta + \gamma.$$

Then, by (4.3) we get

$$V^{\delta}_{p_{n'}} \geq \left(\frac{\gamma}{1-\delta}\right)^{\frac{1}{p_{n'}}} (t^{\delta} + \epsilon) \to t^{\delta} + \epsilon.$$

Consequently, $\limsup_n V^{\delta}_{p_n} \geq t^{\delta} + \epsilon$ - a contradiction with Lemma 4.1. △

Suppose $\{a_n\}$ is an arbitrary sequence that satisfies (5.2). Let $a_n \xrightarrow{\nu} a$, where ν satisfies (4.11). Then $\forall \epsilon > 0$ we get

$$P(t^{\delta} + \epsilon < \|x - a\|) \leq P(\liminf_n \|x - a_n\| > t^{\delta} + \epsilon) \leq$$
$$P(\|x - a_n\| > t^{\delta}, \text{eventually}) \leq \liminf_n P(\|x - a_n\| > t^{\delta} + \epsilon) \leq \delta.$$

Consequently, $P(\|x - a\| > t^{\delta}) \leq \delta$, i.e. $a \in \mathcal{P}^{\delta}_{\infty}$. Hence we have the result.

Proposition 5.2. *Let $\{a_n\}$ be as in (5.1) with $p_o = \infty$. Assume $a_n \xrightarrow{\nu} a$. Then $a \in \mathcal{P}^{\delta}_{\infty}$ and*

$$P(\|x - a\| \leq \liminf \|x - a_n\| \leq t^{\delta}) \geq 1 - \delta. \tag{5.3}$$

Remark. It can be shown that from (5.2) follows $W^{\delta}_{t^{\delta}}(a_n) \to 0$, where W^{δ}_{s} is defined as in previous sections. In particular, when $\delta = 0$, it means that $\{a_n\}$ minimizes a loss function as in (4.5). Then (4.13) applies and (5.3) can be strengthened as follows

$$P(\|x - a\| \leq \liminf_n \|x - a_n\| \leq \limsup_n \|x - a_n\| \leq t) = 0.$$

This approach does not directly apply in case $\delta > 0$.

From (5.4) we get $t^{\delta} = r^{\delta}$. Hence, (4.7) is a necessary condition for $\{a_n\}$ being ν-relatively compact. Clearly $\{a_n\}$ is bounded. If, in addition, every norm-closed ball is ν-sequentially compact, then $\{a_n\}$ is relatively compact. Let us state this observation as a corollary.

Corollary 5.2. *Assume* **A** *and let* $\{a_n\}$ *be as in (5.1) with* $p_o = \infty$. *Then any subsequence of* $\{a_n\}$ *contains a further subsequence converging in* ν *to a trimmed Chebyshev center at level* δ.

Note that Corollary 5.2 proves the existence of trimmed Chebyshev centers for a large class of Banach spaces. The existence of trimmed Chebyshev centers in $\Re^n$ is proved in (Gordaliza, 91a). The claim Corollary 5.2 can be interpreted as follows: $a_n \xrightarrow{\nu} \mathcal{P}^{\delta}_{\infty}$,
Since (4.7) is a necessary condition for a sequence satisfying (5.2) being ν-relatively compact and the metric-topology on E satisfies (4.11), we immediately obtain that a sequence of p_n-predictions ($p_n \to \infty$) is not relatively compact if $t^{\delta} < r^{\delta}$.
Recall Example. Let $p_n \to \infty$, $\epsilon_n := 0.5 + n^{-1} - V_{p_n}$. We know that $\epsilon_n \to 0$. We also know that for each p_n there exists an element $a_n := a_{m_n}$ from the sequence a_m such that $\|X - a_n\|_{p_n} \leq V_{p_n} + \epsilon_n$. Thus, the subsequence a_n is as in (5.1). This is not hard to see $m_n \to \infty$ as $n \to \infty$. Hence, there exists a further subsequence $a_{n'}$ that has no converging subsequence.
On the other hand, we cannot conclude that the sequence $\{a_n\}$ does not converge to the set of Chebyshev centers of P. In fact, any element of a_n is a Chebyshev center.

5.2 Strong convergence

We now turn to the question of the strong convergence of the sequence $\{a_n\}$ as in (5.1) with $p_o = \infty$. Corollary 5.2 states that having suitable topology ν is E, the sequence $\{a_n\}$ converges in ν to the set $\mathcal{P}^{\delta}_{\infty}$. Moreover, we know that $\{a_n\}$ is minimizing in the sense (5.2). When E is uniformly convex and $\delta = 0$, then the convergence of $\{a_n\}$ to the unique Chebyshev center was proved (in Cuesta, Matran, 87). Their proof uses (5.2) and cannot be directly used for $\delta > 0$ (the trimmed Chebyshev centers are generally not unique).
We generalize the mentioned result for a larger class of Banach spaces, allowing also $\delta > 0$. Let (E, ν), as previously, be a topological vector space that satisfies **A**.

Definition 1. (E, ν) satisfies ν-Kadec-Klee (ν-KK) property, if the ν-convergence on the unit sphere implies the strong convergence.

The definition can be restated as follows (See, e.g. Huff, 80).

Definition 1′ (E, ν) satisfies ν- Kadec-Klee property, if for each $\{x_n\}$ in the unit ball of E ν-convergent to x, we have

$$\|x\| < 1$$

provided that $\text{sep}\{x_n\} > 0$, where $\text{sep}\{x_n\} := \inf_{n \neq m} \|x_m - x_n\|$.

Indeed. Suppose (E, ν) is ν-KK and $\{x_n\}$ such a sequence that $\|x_n\| \leq 1$, $x_n \xrightarrow{\nu} x$ and $\text{sep}\{x_n\} > 0$. If $\|x\| = 1$, then $\|x_n\| \to \|x\|$, $x_n \xrightarrow{\nu} x$ and $x_n \not\to x$ a contradiction with ν-KK - property.

On the other hand, suppose (E,ν) has the property described in Definition 1'. At first note: since every norm-closed ball is sequentially ν-continuous, we have that for a sequence converging in ν and in the norm topology, the both limits are the same. [Let $x_n \xrightarrow{\nu} x$, $x_n \to y$, $x \neq y$. Then there exist $\epsilon > 0$ such that $x_n \in \overline{B}(y,\epsilon)$ eventually, but $x \notin \overline{B}(y,\epsilon)$. Since $x_n \xrightarrow{\nu} x$ the latter contradicts the assumption that every norm-closed ball is ν-sequentially closed.]
Suppose $x_n \xrightarrow{\nu} x$, $\|x_n\| \to \|x\|$. W.l.o.g. we may assume $x_n, x \in S_E$ (since (E,ν) is a topological vector space). If $x_n \not\to x$, then it has a subsequence, say $\{x_m\}$, that is not relatively compact [if not, then any subsequence would have a norm-convergent (sub)-subsequence with the limit - as just observed - x and, hence, $x_n \to x$.] Not being relatively compact in a complete metric space (recall the Banach space assumption) means the existence of $\epsilon > 0$ such that $\{x_m\}$ does not have finite ϵ-net. This means, $\{x_m\}$ has a subsequence, say $\{x_{m'}\}$ with $\text{sep}\{x'_m\} > 0$. By Definition $1'$ we get $\|x\| < 1$- a contradiction. Hence $\{x_n\}$ can not have a subsequence that is not relatively compact, implying $x_n \to x$.

Definition 2. (Huff) (E,ν) satisfies ν-uniform Kadec-Klee (ν-UKK) property, if, for each $\epsilon > 0$, there exists $v(\epsilon) > 0$ such that for each $\{x_n\}$ in the unit ball of E ν-convergent to x we have

$$\|x\| \leq 1 - v(\epsilon)$$

provided that $\text{sep}\{x_n\} > \epsilon$.

Usually (uniform) Kadec-Klee property is defined with respect to the weak topology. The ν-(uniform) Kadec-Klee property is a straightforward generalization, see also (Lennard, 90; Besbes, *et al.*, 94; Khamsi, 96). Most spaces encountered in practice that have ν-KK property, do actually have ν-UKK propoerty. But not all. For a counterexample see (van Dulst , Sims, 81). Reflexive space with (weak) Kadec-Klee property is sometimes called Jefimov - Stechkin space. Every uniformly convex space is Jefimov - Stechkin space with uniform Kadec-Klee property. Thus, for example $L_p(\Re)$, l_p $(1 < p < \infty)$ are weak-UKK. In particular, Hilbert space is weak-UKK.
The space l_1 is $*$-weak-UKK, sequential Orlicz spaces with the function having Δ_2-property are weak-UKK (van Dulst *et al.*, 83).
$L_p(\Re)-$spaces $(p > 0)$ with a.s.-topology as ν satisfy ν-UKK [of course, it does not have **A**, but still (4.11). This is useful in probabilistic applications; for example, with Skorohod embedding.] This property be generalized for Lebesgue-Bochner spaces as well (Dilworth *et al.*, 95) .
For more examples, see (Lennard, 90; Besbes *et al.* 94; Khamsi, 96).

When E is ν-UKK, then the ν-convergence of $\{x_n\}$ such that $\liminf_n \|x_n\| \leq l$ and $\text{sep}\{x_n\} > \epsilon$ implies the existence of $\gamma(\epsilon, l) > 0$ such that

$$\|x\| < l - \delta. \tag{5.4}$$

Indeed, for each $\alpha > 0$, there exists a subsequence $\{x_{n'}\}$ such that $\|x_{n'}\| \leq l+\alpha$, $\|x\| \leq l+\alpha$. Because of the ν-UKK

$$\|x\| \leq \Big(1 - v(\frac{\epsilon}{l+\alpha})\Big)(l+\alpha).$$

By choosing α small enough, we can see the existence of $\gamma > 0$ such that $\|x\| \le l - \gamma$. Now we show that Polya algorithm converges (in the sense of b)), provided that E admits a topology satisfying **A**.

Theorem 5.1. *Suppose* (E, ν) *satisfies* **A** *and is* ν*-UKK. Let* $\{a_n\}$ *be as in (5.1) with* $p_o = \infty$*. Then* $a_n \to \mathcal{P}_\infty^\delta$.

Proof. The proof is complete if we show that any subsequence of $\{a_n\}$ contains a converging subsequence. Consider an arbitrary subsequence of $\{a_n\}$, denoted as previously. By assumption **A**, it has a ν-convergent subsequence (denoted as previously by $\{a_n\}$). Suppose $\{a_n\}$ is not relatively compact. Then there $\exists \epsilon > 0$ and a subsequence of a_n, say $\{a_m\}$ such that $\text{sep}\{a_m\} > \epsilon$. Thus $\{a_m\}$ satisfies the assumptions of Proposition 5.2 , hence (5.3) holds. Let A be the corresponding set, $P(A) \le \delta$. For each $x \in A$ we get $x - a_m \xrightarrow{\nu} x - a$, $\text{sep}\{x - a_m\} > \epsilon$ and $\liminf_m (x - a_m) \le t^\delta$. Thus, by (5.4) we have the existence of $\gamma > 0$ such that $\|x - a\| < t^\delta - \gamma$. The latter holds for each $x \in A$, implying $P(\|x - a\| \le t^\delta - \gamma) \le \delta$. This contradicts the definition of t^δ. △

Remark. The foregoing proof relies on (5.3). When there exists a $x_o \in A$ $\|x_o - a\| = t^\delta$, then, obviously the uniform Kadec-Klee property in Theorem 1 can be weakened to the (nonumiform) Kadec-Klee property. However, it is easy to see that generally such an element x_o need not exists. However, for each $\epsilon > 0$ there exists a $x_\epsilon \in A$ such that $\|x_o - a\| < t^\delta + \epsilon$. This makes possible to use the uniform Kadec-Klee property. The question, whether Theorem 5.1 holds with (non-uniform) KK-property, is still open.

Also note that from Theorem 5.1 follows that the set $\mathcal{U}_\infty^\delta$ is compact. This generalizes Theorem 2 in (van Dulst, Sims 81).

We conclude with a convergence result for non-trimmed p_n predictions, as $p_n \to p_o < \infty$.

Theorem 5.2. *Assume* (E, ν) *satisfies* **A**. *Let* $\{a_n\}$ *be as in (5.1) with* $p_o < \infty$. *Then* $a_n \to \mathcal{P}_{p_o}$ *if one of the following conditions holds:*
i) (E, ν) *has* ν*-KK property;*
ii) $\mathcal{P}_{p_o} \subset S$.

Proof. It suffices to show that $\{a_n\}$ contains a strongly convergent subsequence. By (2.6), $\{a_n\}$ is minimizing for V_{p_o}. Thus, Proposition 4.5 applies.
By **A**, $\{a_n\}$ has a ν-convergent subsequence $a_m \xrightarrow{\nu} a$ with the limit in $\mathcal{P}_{p_o}$. From (4.13) we get that $\|x - a_m\| \to \|x - a\|$, when $x \in S(P)$.
If (E, ν) has ν-KK property, then the convergence $x - a_m \to x - a$ obviously follows. That proves i).
Since $a \in \mathcal{P}_{p_o}$, ii) is now obvious. △

The proof of Theorem 5.2 relies on (4.13) and cannot be used, if $\delta > 0$. Therefore, the question of the convergence of trimmed p_n-predictions (with $p_o < \infty$) under ν-KK-property is still open.

References

1. A. Al-Rashed (1983), L_∞-approximation in probability spaces. *J. Math. An. and Appl.* **91** 1-8.

2. M. Besbes, S. Dilworth, P. Dowling, C. Lennard (1994), New convexity and fixed point properties in Hardy and Lebesgue-Bochner spaces. *J. Funct. Anal.* **119**, 340-357.

3. J. Cuesta, C. Matran (1998), Trimmed best k-nets: A robustified version of a L_∞-based clustering method. *Stat. & Prob. Lett.* **36** 401-413.

4. J. Cuesta, C. Matran (1997), Trimmed k-means: An attempt to robustify quantizers. *The Annals of Statistics* **25** 553-576.

6. J. Cuesta, C. Matran (1989a), Uniform consistency of r-means. *Stat. & Prob. Lett.* **6**, 65-71.

7. J. Cuesta, C. Matran (1989b), Conditional Bounds and Best L_∞-approximation in probability spaces. *J. App. Theory* **56** 1-12.

8. J. Cuesta, C. Matran (1988), The strong law of large numbers for k-means and best possible nets. *Prob. Th. Rel. Fields* **78** 523-534.

9. J. Cuesta, C. Matran (1987), Asymptotic behavior of p-predictions for vector valued random variables. *Proc. Am. Math. Soc.* **100**, 1-5.

10. R.B. Darst (1981), Convergence of L_p-approximations as $p \to \infty$. *Proc. Amer. Math. Soc.* **3**, 433-436.

11. R.B. Darst, D.A. Legg, D.W. Dowsend (1983), The Polya algorithm in L_∞-approximation. *J. App. Theory* **38** 209-220.

12. S.J. Dilworth, M. Girardi, D. Kutrazova (1995), Banach spaces which admit a norm with the uniform Kadec-klee property. *Studia Math* **112**, 267-277.

13. D. van Dulst, B. Sims (1983), Fixed point of nonexpansive mappings and Chebyshev centers in Banach spaces with norms of type (KK). *Banach space theory and its applications* Lecture Notes in Math. 991, 35-43.

14. D. van Dulst, V de Valk (1986), (KK)-properties, normal structure and fixed points of nonexpansive mappings in Orlicz sequence spaces. *Can. J. Math.* **3** 728-750.

15. A. Egger, R. Huotari (1990), Rate of convergence of the discrete Polya algorithm. *J. App. Theory* **60** 24-30.

16. A. Gordaliza (1991a), Best approximations to random variables based on trimming procedures *J. App. Theory* **64**, 162-180.

17. A. Gordaliza (1991b), On the breakdown point of multivariate location estimators based on trimming procedures. *Stat. & Prob. Lett.* **11**, 387-394.

18. R. B. Holmes (1972), A Course of Optimization and Best Approximation. (Singer) 225p.

19. N. Herrendorf (1983), Approxiamtion of vector-valued random variables by constants. *J. App. Theory* **37**, 175-181.

20. R. Huff (1980), Banach spaces which are nearly uniformly convex. *Rocky Mount. J. Math.* **10**, 743-749.

21. R. Huotari, W. Li (1994), Continuity of metric projection, polya algorithm, strict best approximation, and tubularity of convex sets. *J. Math. An. and Appl.* **182** 836-856.

22. M. Khamsi (1996), On uniform Opial condition and uniform Kadec-Klee property in Banach spaces. *Nonlinear Anal.* **26**, 1733-1748.

23. D. A. Legg, D. A. Dowsend (1989), The Polya algorithm for convex approximation. *J. Math. An. and Appl.* **141**, 431-441.

24. J. Lember (1999), Consistency of empirical k-centres (Doctoral dissertation). *Diss. Math. Univ. Tartuensis* **20**, 146 p. (submitted)

25. J. Lember (2000), On minimizing sequences of k-centres. (submitted)

26. C. Lennard (1990), $\mathcal{C}_1$ is uniformly Kadec-Klee. *Proc. Amer. Math. Soc.* **109**, 71-77.

27. W. Li (1995), Convergence of Polya algorithm and continuous metric selections in space of continuous functions. *J. App. Theory* **80**, 164-179.