A Monte-Carlo algorithm for finding a near-optimal arrangement in the Steinitz functional

L.A.Chobanyan, *Institute TECHINFORMI, Georgian Technical University*

S.A.Chobanyan, Niko Muskhelishvili Institute of Computational Mathematics, Georgian Technical University

---

Supported in part by the Rustaveli Science Foundation grant FR/539/5-100/13.

1. **The main aim** of this talk is to apply our maximum inequality and transference theorem [1,2] - presented in the next section – to the following problem which is an important subtask of many problems of machine learning, scheduling theory and discrepancy theory [3-8].

Find or estimate the minimum in π the *Steinitz functional*

$$\Phi_x(\pi) = \max_{1 \le k \le n} \| \sum_1^k x_{\pi(i)} \|,$$

where $x = (x_1, \dots, x_n)$ is a fixed collection of elements of a normed space X, and $\pi : \{1, \dots, n\} \to \{1, \dots, n\}$ is a permutation.

The peculiarity of the related applied problems is that *n*, the dimension of X, is very large, so that the *brute force* idea as a rule does not work.

The problem was posed by E.Steinitz [9] who was solving the question on sum range of a conditionally convergent series in a finite dimensional space (the generalization of the famous Riemann problem). If time permits we'll come back to this and other related nice analytical problems (among them to the open so far Kolmogorov-Garsia problem on the system of almost everywhere convergence of a rearranged orthogonal sequence).

Literature: *[1] S.Chobanyan, G.Giorgobiani, Lecture Notes in Math.,1391, 1987, 33-46;[2] S.Chobanyan, Birkhauser, Prog. Probab. 35, 1994, 3-29; [3] S.Sevastyanov, Discrete Applied Math.; [4] J.Beck, V.T.Sos, Handbook in Combinatorics, v.2, Elsevier Sciernce B.V. and MIT Press, 1995; [5] N.Makai, Appl. Math. Comp., 150, 2004, 785-801; [6] L.Chobanyan, S.Chobanyan, Giorgobiani, Bull. Georgian National Acad. Sci., 5, 2011, 16-21; [7] L.Chobanyan, V.V.Kvaratskhelia, 9th International Conference CSIT-2013, Yerevan, Armenia, Proceedings: pp.58-60, 2013; [8]N.Harvey, S.Samadi, Workshop and Conference Proceedings 35, 1-18;2014; [9] E.Steinitz, J.Reine Angew. Math., 143 (1913), 128-175.*

2. The main maximum inequalities. In this talk we apply the following two maximum inequalities to problems related to the calculation or estimation of the Steinitz functional.

Theorem 1. [1,2] *Let* $x_1, \ldots, x_n \in X$ *be a collection of elements of a normed space* X *with* $\sum_1^n x_i = 0$. *Then*

a. *For any collection of signs* $\vartheta = (\vartheta_1, \ldots, \vartheta_n)$ *there is a permutation* $\pi : \{1,\ldots,n\} \to \{1,\ldots,n\}$ *such that*

$$\max_{1\le k\le n} \| \sum_1^k x_i \| + \max_{1\le k\le n} \| \sum_1^k \vartheta_i x_i \| \ge 2\max_{1\le k\le n} \| \sum_1^k x_{\pi(i)} \|;$$

There is an explicit one-to-one correspondencebetween $\vartheta = (\vartheta_1, \ldots, \vartheta_n)$ *and* $\pi(\vartheta)$.

b. *(Transference Theorem). There is a permutation* $\sigma : \{1,\ldots,n\} \to \{1,\ldots,n\}$ *such that*

$$\max_{1\le k\le n} \| \sum_1^k x_{\sigma(i)} \| \le \max_{1\le k\le n} \| \sum_1^k \vartheta_i x_{\sigma(i)} \|$$

for any collection of signs $\vartheta = (\vartheta_1, \ldots, \vartheta_n)$.

Literature. [1] S.Chobanyan, G.Giorgobiani, Lecture Notes in Math.,1391, 1987, 33-46;[2] S.Chobanyan, Birkhauser, Prog. Probab. 35, 1994, 3-29;

3. The greedy algorithm is not in general the best. Given vectors $x_1, ..., x_n \in X$ *a greedy algorithm* chooses at each step a vector that minimizes the norm of the next partial sum. In other words, on step 1 it chooses an element x_{n_1} that has a minimum norm. On step 2 it selects an element x_{n_2}, $x_{n_2} \neq x_{n_1}$ such that

$\| x_{n_1} + x_{n_2} \| \leq \| x_{n_1} + x_{n_k} \|$ for any $n_k \neq n_1$, etc.

The following example constructed by Jakub Wojtaszchik (oral communication) shows that a greedy algorithm is not in general the best one even in a two-dimensional space.

Example. Consider *n* groups of vectors of l_∞^2 each consisting of the three following vectors: (1,1), (2,-3), and (-3,2). Obviously, the greedy algorithm chooses at the first *n* steps the vectors (1,1),…, (1,1). Therefore, for the optimal permutation π_o and greedy permutation π_g we have respectively

$$\max_{1 \leq k \leq n} \| x_{\pi_o(1)} + ... + x_{\pi_o(k)} \| = 3;$$

and $$\max_{1 \leq k \leq n} \| x_{\pi_g(1)} + ... + x_{\pi_g(k)} \| = n + 2.$$

In [1] we show that such sort of an example can be constructed in any 2-dimensional normed space.

Literature. [1] G.Chelidze, S.Chobanyan, G.Giorgobiani, V.Kvaratskhelia, Bull. Georgian National Acad. Sci, 4, 2010, 5-7.

4. Corollaries to the Transference theorem (Theorem 1b).

The Transference theorem allows us to get a permutation theorem given a sign theorem. Moreover, as we'll see in Section 5, if the sign algorithm is constructive, then a desired permutation can also be found constructively. As a first example we consider the classical Steinitz permutation theorem that we get from the following Grinberg-Sevostyanov sign theorem .

Theorem 2 [1,2]. *Let* X *be a normed space of dimension* d, $x_1,...,x_n \in X$, $\| x_i \| \le 1$, $i = 1, ..., n$. *Then there exists a collection of signs* $\vartheta = (\vartheta_1,...,\vartheta_n)$ *such that*

$$\max_{1\le k\le n} \| x_1\vartheta_1 + ... + x_k\vartheta_k \| \le 2d .$$

The permutation version of Theorem 2 found by the Transference theorem can be stated as follows.

Corollary. *The Steinitz inequality. Let* X *be a normed space of dimension* d, $x_1,...,x_n \in X$, $\| x_i \| \le 1$, $i = 1, ..., n$ *and* $x_1 + ... + x_n = 0$. *Then there exists a permutation* $\sigma: \{1, ..., n\} \to \{1, ..., n\}$ *such that*

$$\max_{1\le k\le n} \| x_{\sigma(1)} + ... + x_{\sigma(k)} \| \le 2d.$$

Remark. Steinitz [3] proved his inequality straightforwardly, however a proof through the sign version additionally allows to find the desired permutation constructively provided that the collection of signs in the sign version can be obtained constructively, by use of the Transference theorem (see Section 5).

Another sign-permutation duality example is provided by the case of the space l_∞.

Theorem 3 (Spencer [4]). *Let* $X = l_\infty^d$, $x_1,...,x_n \in X$, $\| x_i \| \le 1$, $i = 1, ..., n$. *Then there exists a collection of signs* $\vartheta = (\vartheta_1,...,\vartheta_n)$ *such that*

$$\max_{1\le k\le n} \| x_1\vartheta_1 + ... + x_k\vartheta_k \| \le \sqrt{2n\ln 2d} .$$

Remark. Spencer also gives an effective way of finding thetas. This means that the following dual permutational counterpart also is provided by an effective construction of the permutation (see Section 5).

Corollary. *Let* $X = l_\infty^d$, $x_1, ..., x_n \in X$, $\| x_i \| \le 1$, $i = 1, ..., n$ *and* $x_1 + ... + x_n = 0.$ *Then there exists a permutation* $\sigma: \{1, ..., n\} \to \{1, ..., n\}$ *such that*

$$\max_{1 \le k \le n} \| x_{\sigma(1)} + ... + x_{\sigma(k)} \| \le .\sqrt{2n \ln 2d} .$$

Literature. [1] V.S.Grinberg and S.V.Sevastyanov, Funkts. Analiz I Prilozh, 14, 1980, 56-57 (in Russian); [2] I.Barani and V.S.Grinberg, Linear Algebra Appl., 41, 1981, 1-9; [3] E.Steinitz, J.Reine Ang. Mathematik, 143, 1913, 128-175, [4] J.Spencer, J.Combinatorial Theory, Ser. B, 23, 1977, 68-74.

5. **Corollary to the Transference theorem.** In this section we show that the algorithm for near optimal permutation for

$\Phi_x(\pi) = \max_{1\le k\le n} \| \sum_1^k x_{\pi(i)} \|$, can be reduced to the algorithm for near optimal sign. The reduction is based on the Transference theorem (Theorem 1b).

Theorem 4. *Let* X *be a normed space,*

$$x_1,\dots,x_n \in X,\ \| x_i \| \le 1,\ i=1,\dots,n, \text{ and } x_1+\dots+x_n = 0.$$

Assume that for any permutation $\pi: \{1,\dots,n\} \to \{1,\dots,n\}$ *there is an algorithm with a polynomial complexity to define* $\vartheta = (\vartheta_1,\dots,\vartheta_n)$ *such that*

$$\max_{1\le k\le n} \| \sum_1^k x_{\pi(i)} \vartheta_i \| \le D, \qquad (1)$$

where D does not depend on π.

Then for any $\varepsilon > 0$ *there is an algorithm with a polynomial complexity to define* $\sigma: \{1,\dots,n\} \to \{1,\dots,n\}$ *such that*

$$\max_{1\le k\le n} \| \sum_1^k x_{\sigma(i)} \| \le D + \varepsilon \quad .$$

The complexity of the algorithm is $(C \cdot \log(n/\varepsilon)$*, where* C *is the complexity of the sign algorithm (for finding* ϑ).

Literature. [1] N.Makai, Appl. Math. Comp., 150, 2004, 785-801;

[6] L.Chobanyan, S.Chobanyan, Giorgobiani, Bull. Georgian National Acad. Sci., 5, 2011, 16-21; [7] L.Chobanyan, V.V.Kvaratskhelia, 9th International Conference CSIT-2013, Yerevan, Armenia, Proceedings: pp.58-60, 2013; [8]N.Harvey, S.Samadi, Workshop and Conference Proceedings 35, 1-18;2014;

6. **Applying the Monte-Carlo.** *Let* $X = l_2$,

$$x_1,\dots,x_n \in X,\ \| x_i \| \le 1,\ i=1,\dots,n, \text{ and } x_1+\dots+x_n = 0.$$

Let us introduce the notations:

$$| x_\pi | = \max_{1\le k\le n} \| \sum_1^k x_{\pi(i)} \|, \quad | x_\pi \vartheta | = \max_{1\le k\le n} \| \sum_1^k x_{\pi(i)} \vartheta_i \| \quad ,$$

We choose at random *k* collections $\theta_1^{(1)},\dots,\theta_k^{(1)}$ each of them being a collection of *n* signs and choose among them ϑ_1 , that one, for which $| x\vartheta_i^{(1)} |$ attains its minimum. Then we create the permutation π_2 generated by the initial permutation (denote it by π_1) and ϑ_1 according to Theorem 1(a). Therefore, we'll have

$$|x_{\pi_2}| \le \frac{1}{2}|x_{\pi_1}| + \frac{1}{2}|x_{\pi_1}\vartheta_1|.$$

Then we choose at random (independently) $\theta_1^{(2)},\dots,\theta_k^{(2)}$ and among them choose ϑ_2 minimizing $|x_{\pi_2}\vartheta_i^{(2)}|$. Carrying out these iterations $l-1$ times we find a sequence of permutations $\pi_1,\dots,\pi_l$ such that for the last permutation we get the following inequality

$$|x_{\pi_l}| \le \frac{1}{2^l}|x_{\pi_1}| + (1-\frac{1}{2^l})\ \max_{i\le l}|x_{\pi_i}\vartheta_i|. \qquad (1)$$

We now show that π_l for sufficiently large l is a near optimal permutation. For these purposes let us make sure that the following probability is small enough after an appropriate choice of C, k *and* l :

$$P(\max_{i\le l}|x_{\pi(i)}\vartheta_i| > C\sqrt{n}) \le \sum_{i=1}^{l} P(|x_{\pi(i)}\vartheta_i| > C\sqrt{n}) \le$$
$$\sum_{i=1}^{l} P\{(|x_{\pi(i)}\vartheta_1^i| > C\sqrt{n}) \cap \dots \cap ((|x_{\pi(i)}\vartheta_k^i| > C\sqrt{n})\} =$$
$$\sum_{i=1}^{l} P\{(|x_{\pi(i)}\vartheta_1^i| > C\sqrt{n})\}^k.$$

The next step is to use the estimation of the tail probability for the Rademacher random variables with values in a normed space (in our case it is the space l_2^d) (see the monograph by M.Ledoux and M.Talagrand [1], p.101). Then we get

$$P(\max_{i\le l}|x_{\pi(i)}\vartheta_i| > C\sqrt{n}) \le l\cdot 2^{k+1}\, e^{-\frac{Cn}{32n}}.$$

Up to now l and k were arbitrary. Letting $l=k$, as well as $C=32\cdot\ln 4$ we get

$$P(\max_{i\le l}|x_{\pi(i)}\vartheta_i|>C\sqrt{n})\le\frac{k}{2^{k-1}}\quad(2).$$

These computations along with (1) imply that with large probability (which can be made arbitrarily close to one) the following inequality holds

$$|x_{\pi_l}|\le\frac{1}{2^l}|x_{\pi_1}|+C\sqrt{n}.\quad(3)$$

According to Theorem 1b, the order of $\sqrt{n}$ is correct, and

it is also known that it can not be improved.

Therefore, we proved the following

Theorem 5. The random algorithm described in this section leads to the nearly optimal permutation. The algorithm runs in a polynomial time.

Literature.

M.Ledoux, M.Talagrand, Probability in Banach Spaces, 1991.