

# SOFT CLUSTERING BY CONVEX ELECTORAL MODEL

Yurii Nesterov

REPRINT | 3245

## **CORE**

Voie du Roman Pays 34, L1.03.01

B-1348 Louvain-la-Neuve

Tel (32 10) 47 43 04

Email: [lidam-library@uclouvain.be](mailto:lidam-library@uclouvain.be)

<https://uclouvain.be/en/research-institutes/lidam/core/core-reprints.html>



# Soft clustering by convex electoral model

Yurii Nesterov<sup>1</sup>

Published online: 19 August 2020  
© The Author(s) 2020

## Abstract

In this paper, we suggest a new technique for soft clustering of multidimensional data. It is based on a new convex voting model, where each voter chooses a party with certain probability depending on the divergence between his/her preferences and the position of the party. The parties can react on the results of polls by changing their positions. We prove that under some natural assumptions this system has a unique fixed point, providing a unique solution for soft clustering. The solution of our model can be found either by imitation of the sequential elections, or by direct minimization of a convex potential function. In both cases, the methods converge linearly to the solution. We provide our methods with worst-case complexity bounds. To the best of our knowledge, these are the first polynomial-time complexity results in this field.

**Keywords** Soft clustering · Fuzzy clustering · C-means · Polynomial-time complexity bounds · Electoral models

## 1 Introduction

### 1.1 Motivation

Cluster analysis is one of the most important and natural tools in Data Mining. It was immediately addressed with the first steps in the computer sciences. It is enough to mention the first publications in the middle of twentieth century (see Forgy 1965; Lloyd 1057; MacQueen 1967; Steinhaus 1957). The first proposed algorithms had a combinatorial nature. Given the set of observations  $v_i \in \mathbb{R}^m, i = 1, \dots, N$ , and pre-defined number of clusters  $K$ , we can define the total variance of the partition:

$$\text{Var}(\mathcal{S}) \stackrel{\text{def}}{=} \sum_{k=1}^K \sum_{i \in S_k} \|v_i - c_k\|_2^2, \quad (1.1)$$

where  $\mathcal{S} = \{S_1, \dots, S_K\}$  is a partition of the set  $\{1, \dots, N\}$ , and the center of  $k$ th cluster is given by a least-square estimate:

$$c_k = \arg \min_{c \in \mathbb{R}^m} \sum_{i \in S_k} \|v_i - c\|_2^2 = \frac{1}{|S_k|} \sum_{i \in S_k} v_i, \quad k = 1, \dots, K. \quad (1.2)$$

Thus, it looks natural to find the best clustering by solving the following problem:

$$\min_{\mathcal{S}} \text{Var}(\mathcal{S}). \quad (1.3)$$

The first combinatorial greedy method for minimizing the function  $\text{Var}(\cdot)$  was proposed by Lloyd (1057) and reinvented by Forgy (1965). The algorithm stops when further improvement of the objective function is no longer possible. This approach is usually addressed as *hard K-means clustering*.

However, it was discovered soon in Garey et al. (1982) that the problem (1.3) is NP-hard even for  $K = 2$  (see also Aloise et al. 2009; Dasgupta and Freund 2009; Kleinberg et al. 1998). In order to get rid of the combinatorial nature of problem (1.3), in Dunn (1973), Bezdek (1981), it was suggested to use fuzzy (or *soft*) clustering. In this approach, the positions of the centers  $C = (c_1, \dots, c_K)$  of the clusters become variables. At the same time, each element  $v_i$  participates in cluster  $k$  up to a certain degree  $p_i^{(k)}$ . The smaller this *membership value* is, the lower should be the probability that the item  $i$  must be attributed to cluster  $k$ . Usually, the dependence of these values on the central positions of the clusters is given by the following expressions:

---

Communicated by Yaroslav D. Sergeyev.

✉ Yurii Nesterov  
Yurii.Neterov@uclouvain.be

<sup>1</sup> Center for Operations Research and Econometrics (CORE), Catholic University of Louvain (UCL), Louvain-la-Neuve, Belgium

$$p_i^{(k)} = \pi_i^{(k)}(C) \stackrel{\text{def}}{=} \left[ \sum_{j=1}^K \left( \|v_i - c_j\|_2 \right)^{\frac{2}{\sigma-1}} \right]^{-\sigma},$$

$$k = 1, \dots, K, \quad i = 1, \dots, N,$$

where the parameter  $\sigma > 1$  is called the *fuzzifier* of the system. In the absence of a reliable information on its reasonable value, the usual choice is  $\sigma = 2$ .

Thus, in order to compute so-called *C-means soft clustering*, we need to solve the following nonlinear minimization problem:

$$\min_{C \in \mathbb{R}^{m \times K}} \left\{ \text{Var}_\pi(C) = \sum_{i=1}^N \sum_{k=1}^K \pi_i^{(k)}(C) \|v_i - c_k\|_2^2 \right\}. \quad (1.4)$$

For  $\sigma > 1$ , this can be done by the standard algorithms of general nonlinear optimization. However, note that function  $\text{Var}_\pi(\cdot)$  can have very complicated topological structure. So, the only possible theoretical guarantee for these algorithms is the convergence to local minima, which are most probably not unique.

At this moment, there exist hundreds of different algorithms for computing hard and soft clusterings (see, for example, the surveys Bora and Gupta 2014; Peters et al. 2013; Xu and Wunsch 2005; Yang 1993). However, to the best of our knowledge, all of them a *heuristic*. None of them is supported by rigorous theoretical statements on its efficiency and the quality of the generated solutions. Such a situation does not look very striking since both variants of clustering problems (1.3) and (1.4) look computationally difficult in view of their non-convexity.

The main goal of this paper is to show that a small change in the problem setting makes our goal computationally tractable. In order to explain the origin of our approach, let us try to find some successful examples of clustering in the big systems, existing in real life.<sup>1</sup> For our purposes, the most interesting could be a proper understanding of electoral procedures in modern democratic states.

Since the political parties must reflect the aggregate preferences of a big group of voters, this is indeed a good example of clustering in real life—which has already proved its efficiency by two hundred years of elections in the USA. Note that this procedure should have an enormous stabilizing effect, which allows the country to recover quickly after external and internal shocks and

<sup>1</sup> It is not the first time when the nature helps in finding an efficient optimization algorithm. May be the most striking example is a derivation of the shortest-path Dijkstra’s algorithm from the least time Fermat principle of propagation of light. However, now we cannot discuss this interesting topic in details.

remain functioning in accordance with the interests and experience of the society.<sup>2</sup>

In Operations Research, behavior of a stable system is usually explained by existence of a convex potential function, which must be minimized by the natural evolution of the system. Thus, one of the main goals of this paper is a construction of a *convex model* for sequential elections, which can help in *soft clustering* of voters, taking into account their internal preferences. In our model, the voter is characterized by a set of opinions (features), which can be compared with the positions of parties. The opinions of voters are fixed. However, the positions of parties may be changed in accordance with observed opinions of attracted voters. At the same time, party cannot go too far from its basic declarations (*core values*). For each voter, we compute the probabilities for selecting each party. These probabilities are adjusted after each round of consecutive elections. Thus, in our approach probabilities justify a computational tool for approaching a good clustering.

Our main result is that under some natural behavioral assumptions, the electoral system is stable. This means that the sequential election procedure converges to a unique fixed point. Moreover, this convergence is fast (linear). Thus, we get a computationally efficient algorithm for computing soft clusters. It has a form of an alternating minimization method for a convex potential function. From the technical point of view, the main novelty of our approach consists in a complete elimination of the least squares principle [compared with (1.1), (1.2), and (1.4)]. Instead, the membership values are computed, for example, as

$$p_i^{(k)} = \frac{e^{-\|v_i - c_k\|/\mu}}{\sum_{j=1}^K e^{-\|v_i - c_j\|/\mu}}, \quad k = 1, \dots, K, \quad i = 1, \dots, N, \quad (1.5)$$

where  $\mu > 0$  is a volatility coefficient. Note that the norms in this expression are not squared.

### 1.2 Contents

The paper is organized as follows: In Sect. 2, we develop all necessary mathematical tools for analyzing the global convergence of an alternating minimization scheme as applied to a bivariate potential function with a multiplicative interaction term. We prove sufficient conditions for its convexity and establish a linear rate of convergence for alternating minimization.

<sup>2</sup> An indirect proof of this conjecture is given by economic domination of democratic states in the modern world. At the same time, we can present a mathematically correct sufficient condition for stability of democratic elections (see Theorem 4).

In Sect. 3, we describe our electoral model and introduce the main behavioral assumptions for the voters and political parties. We define the stable systems as systems for which the sequential election procedure has a unique fixed point. In Sect. 4, we justify the conditions for stability, using bivariate potential function studied in Sect. 2. It is important that for measuring divergence in the opinions of voters and positions of the parties we can use an arbitrary Lipschitz continuous function. In any case, the rate of convergence of the whole process is linear.

In Sect. 5, we rewrite the stability condition of the system in terms of the individual tolerance of the party and discuss the computational cost of the proposed model. In the simplest case of measuring the divergence in the opinions by a usual norm, this cost is proportional to  $O(\frac{1}{\epsilon^2})$ , where  $\epsilon$  is the desired accuracy of the computed solution. However, we can use smooth approximations of the norms, or distances based on Huber function. In this case, the computational cost of our model is reduced. It becomes proportional to  $\ln \frac{1}{\epsilon}$  multiplied by some factors dependent on the parameters of our model.

In Sect. 6, we analyze a clustering scheme based on a direct minimization of the bivariate potential. For doing that, we replace the probabilistic assignment rule (1.5) by a Euclidean projection onto the standard simplex. As a consequence, we can apply to our model the standard gradient methods of smooth convex optimization with linear rate of convergence. Consequently, we get the worst-case estimate for the total computational expenses of the order  $O(KmN \ln \frac{1}{\epsilon})$  arithmetic operations. This is the fastest clustering scheme in our paper. Recall that for the best of our knowledge, in our paper we present the first complexity results in this field.

### 1.3 Notation and generalities

In this paper, we denote by  $\mathbb{E}$  with an optional subscript a finite-dimensional real vector space, and by  $\mathbb{E}^*$  the corresponding dual space. For linear function  $s \in \mathbb{E}^*$ , denote its value at  $x \in \mathbb{E}$  by  $\langle s, x \rangle_{\mathbb{E}}$ . If no ambiguity arise, the subscript is usually omitted. For any norm  $\|\cdot\|_{\mathbb{E}}$  in  $\mathbb{E}$ , we define its dual norm in the standard way,

$$\|s\|_{\mathbb{E}}^* = \max_{x \in \mathbb{E}} \{\langle s, x \rangle : \|x\|_{\mathbb{E}} \leq 1\}, \quad s \in \mathbb{E}^*, \tag{1.6}$$

which ensures validity of Cauchy–Schwartz inequality:

$$\langle s, x \rangle_{\mathbb{E}} \leq \|s\|_{\mathbb{E}}^* \cdot \|x\|_{\mathbb{E}}, \quad s \in \mathbb{E}^*, x \in \mathbb{E}. \tag{1.7}$$

If  $\mathbb{E} = \mathbb{R}^n$ , the space of real column vectors  $x = (x^{(1)}, \dots, x^{(n)})^T$ , then  $\mathbb{E}^* = \mathbb{R}^n$  and

$$\langle s, x \rangle = \sum_{i=1}^n s^{(i)} x^{(i)}, \quad s, x \in \mathbb{R}^n.$$

Notation  $\mathbb{R}_+^n$  is used for the positive orthant:

$$\mathbb{R}_+^n = \{x \in \mathbb{R}^n : x^{(i)} \geq 0, i = 1, \dots, n\},$$

and by  $\Delta_n$  we denote the standard simplex:

$$\Delta_n = \left\{ x \in \mathbb{R}_+^n : \sum_{i=1}^n x^{(i)} = 1 \right\}.$$

The standard notation is used for  $\ell_p$  norms:

$$\|x\|_p = \left[ \sum_{i=1}^n |x^{(i)}|^p \right]^{1/p}, \quad x \in \mathbb{R}^n, p \geq 1,$$

with  $\|x\|_{\infty} = \max_{1 \leq i \leq n} |x^{(i)}|$ .

For function  $f(x), x \in \mathbb{E}$ , we denote by  $\nabla f(x) \in \mathbb{E}^*$  its gradient at  $x$ . If  $f$  is a non-differentiable convex function, the same notation is used for its subgradient. If function  $f(\cdot, \cdot)$  has two arguments, notation  $\nabla_2 f(x, y)$  corresponds to its gradient with respect to variable  $y$ . Finally, for twice differentiable function  $f$  we denote by  $\nabla^2 f(x)$  its Hessian at  $x$ . For  $x$  being fixed, this is a linear operator from  $\mathbb{E}$  to  $\mathbb{E}^*$ .

Recall that function  $f$  is called *strongly convex* on convex set  $Q \subseteq \mathbb{E}$  with *convexity parameter*  $\sigma > 0$  if for all  $x, y \in Q$  and  $\alpha \in [0, 1]$  we have

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y) - \frac{\sigma}{2} \alpha(1 - \alpha) \|x - y\|_{\mathbb{E}}^2. \tag{1.8}$$

The most important property of strongly convex function is that it attains a unique minimum on any closed convex set  $Q$ . Therefore, if  $x^* = \arg \min_{x \in Q} f(x)$ , then for any  $x \in Q$  and  $\alpha \in (0, 1)$  we have

$$f(x^*) \leq f(\alpha x^* + (1 - \alpha)x) \stackrel{(1.8)}{\leq} \alpha f(x^*) + (1 - \alpha)f(x) - \frac{\sigma}{2} \alpha(1 - \alpha) \|x - x^*\|_{\mathbb{E}}^2.$$

Dividing the difference of the right- and left-hand side of this inequality  $(1 - \alpha)$  and tending  $\alpha$  to one, we get

$$f(x) \geq f(x^*) + \frac{\sigma}{2} \|x - x^*\|_{\mathbb{E}}^2, \quad x \in Q. \tag{1.9}$$

## 2 Bivariate potential with multiplicative interaction cost

Let  $\mathbb{E}, \mathbb{E}_1$ , and  $\mathbb{E}_2$  be three finite-dimensional real vector spaces. Consider the following function of two variables:

$$\Phi(x, p) = f(x) + g(x, p) + h(p), \quad x \in \mathbb{E}_1, p \in \mathbb{E}_2. \tag{2.1}$$

where functions  $f$  and  $h$  are closed and convex on their domains. Our main structural assumption on function  $\Phi$  is that the interaction term  $g$  has a *multiplicative form*:

$$g(x, p) = \langle G_1(x), G_2(p) \rangle_{\mathbb{E}}, \tag{2.2}$$

$$G_1 : \mathbb{E}_1 \rightarrow \mathbb{E}^*, \quad G_2 : \mathbb{E}_2 \rightarrow \mathbb{E}.$$

In Sect. 4, we will see an application example for a potential having exactly this structure.

The main goal of this section consists in developing natural conditions for convexity of function  $\Phi$  and analyzing the performance of alternating minimization approach for solving the corresponding optimization problems.

Denote by  $Q_1$  and  $Q_2$  the closed convex feasible sets in the spaces  $\mathbb{E}_1$  and  $\mathbb{E}_2$ , respectively. And let  $Q = Q_1 \times Q_2$ .

**Assumption 1** For any  $x \in Q_1$  function  $g(x, \cdot) = \langle G_1(x), G_2(\cdot) \rangle$  is closed and convex on  $Q_2$ , and for any  $p \in Q_2$  function  $g(\cdot, p) = \langle G_1(\cdot), G_2(p) \rangle$  is closed and convex on  $Q_1$ .

In the remaining part of this section, we always assume that Assumption 1 is satisfied.

**Example 1** Assumption 1 is valid, for example, for linear operators:

$$G_1(x) = A_1x + b_1, \quad G_2(p) = A_2p + b_2.$$

Another important example is  $g(x, p) = \langle G_1(x), p \rangle$ , where  $p \in \mathbb{R}_+^n$  and all components of the vector function  $G_1(x)$  are convex.

Assumption 1 has the following important consequence.

**Lemma 1** For any two points  $z_0 = (x_0, p_0)$  and  $z_1 = (x_1, p_1)$  from  $Q$  and any  $\alpha \in [0, 1]$ , we have

$$\langle G_1((1 - \alpha)x_0 + \alpha x_1), G_2((1 - \alpha)p_0 + \alpha p_1) \rangle \leq \langle (1 - \alpha)G_1(x_0) + \alpha G_1(x_1), (1 - \alpha)G_2(p_0) + \alpha G_2(p_1) \rangle. \tag{2.3}$$

**Proof** Denote  $z_\alpha = (x_\alpha, p_\alpha) = (1 - \alpha)z_0 + \alpha z_1$ . Then in view of Assumption 1 we have

$$\langle G_1(x_\alpha), G_2(p_\alpha) \rangle \leq \langle G_1(x_\alpha), (1 - \alpha)G_2(p_0) + \alpha G_2(p_1) \rangle.$$

On the other hand, by the same reason

$$\langle G_1(x_\alpha), G_2(p_i) \rangle \leq \langle (1 - \alpha)G_1(x_0) + \alpha G_1(x_1), G_2(p_i) \rangle, \quad i = 1, 2.$$

Putting all inequalities together, we get (2.3).  $\square$

Now we can justify a sufficient condition for convexity of the potential  $\Phi(\cdot, \cdot)$ .

**Theorem 1** Let function  $f$  be strongly convex on  $Q_1$  with parameter  $\sigma_1$ , and function  $h$  be strongly convex on  $Q_2$  with parameter  $\sigma_2$ . Assume that the operators  $G_1(\cdot)$  and  $G_2(\cdot)$  are Lipschitz continuous:

$$\|G_1(x_0) - G_1(x_1)\|_{\mathbb{E}}^* \leq L_1 \|x_0 - x_1\|_{\mathbb{E}_1}, \quad x_0, x_1 \in Q_1, \\ \|G_2(p_0) - G_2(p_1)\|_{\mathbb{E}} \leq L_2 \|p_0 - p_1\|_{\mathbb{E}_2}, \quad p_0, p_1 \in Q_2. \tag{2.4}$$

If the constants  $L_1$  and  $L_2$  are small enough, namely, if

$$L_1^2 L_2^2 \leq \sigma_1 \sigma_2, \tag{2.5}$$

then function  $\Phi$  is convex on  $Q = Q_1 \times Q_2$ . If inequality (2.5) is strict, then  $\Phi$  is strongly convex on  $Q$ .

**Proof** Let us fix two points  $z_0 = (x_0, p_0)$  and  $z_1 = (x_1, p_1)$  from  $Q$ . Consider the intermediate point  $z_\alpha = (1 - \alpha)z_0 + \alpha z_1$  with  $\alpha \in [0, 1]$ . Then, since  $p_\alpha \in Q_2$ , we have

$$\begin{aligned} \Phi(z_\alpha) &= f(x_\alpha) + \langle G_1(x_\alpha), G_2(p_\alpha) \rangle + h(p_\alpha) \\ &\stackrel{(2.3)}{\leq} f(x_\alpha) + \langle (1 - \alpha)G_1(x_0) + \alpha G_1(x_1), (1 - \alpha)G_2(p_0) + \alpha G_2(p_1) \rangle + h(p_\alpha) \\ &\stackrel{(1.8)}{\leq} (1 - \alpha)f(x_0) + \alpha f(x_1) - \frac{\sigma_1}{2} \alpha(1 - \alpha) \|x_1 - x_0\|^2 \\ &\quad + \langle (1 - \alpha)G_1(x_0) + \alpha G_1(x_1), (1 - \alpha)G_2(p_0) + \alpha G_2(p_1) \rangle \\ &\quad + (1 - \alpha)h(p_0) + \alpha h(p_1) - \frac{\sigma_2}{2} \alpha(1 - \alpha) \|p_1 - p_0\|^2 \\ &= (1 - \alpha)[\Phi(z_0) - \langle G_1(x_0), G_2(p_0) \rangle] \\ &\quad + \alpha[\Phi(z_1) - \langle G_1(x_1), G_2(p_1) \rangle] \\ &\quad + \langle (1 - \alpha)G_1(x_0) + \alpha G_1(x_1), (1 - \alpha)G_2(p_0) + \alpha G_2(p_1) \rangle \\ &\quad - \frac{\alpha(1 - \alpha)}{2} [\sigma_1 \|x_1 - x_0\|^2 + \sigma_2 \|p_1 - p_0\|^2] \\ &= (1 - \alpha)\Phi(z_0) + \alpha\Phi(z_1) - \alpha(1 - \alpha)\delta(z_0, z_1), \end{aligned}$$

where  $\delta(z_0, z_1) = \frac{\sigma_1}{2} \|x_1 - x_0\|^2 + \frac{\sigma_2}{2} \|p_1 - p_0\|^2 + \langle G_1(x_1) - G_1(x_0), G_2(p_1) - G_2(p_0) \rangle$ . Assuming now that  $L_1^2 L_2^2 \leq (\sigma_1 - \epsilon)(\sigma_2 - \epsilon)$  for some small  $\epsilon \geq 0$ , we have

$$\begin{aligned} \delta(z_0, z_1) &\stackrel{(2.4)}{\geq} \frac{\sigma_1}{2} \|x_1 - x_0\|^2 + \frac{\sigma_2}{2} \|p_1 - p_0\|^2 \\ &\quad - (\sigma_1 - \epsilon)^{1/2} (\sigma_2 - \epsilon)^{1/2} \|x_1 - x_0\| \cdot \|p_1 - p_0\| \\ &\geq \frac{\epsilon}{2} \left( \|x_1 - x_0\|^2 + \|p_1 - p_0\|^2 \right). \end{aligned}$$

Thus, if  $\epsilon > 0$ , then function  $\Phi$  is strongly convex in view of definition (1.8). If  $\epsilon = 0$ , then  $\Phi$  is just a convex function.  $\square$

Simple example with  $Q_1 = \mathbb{R}^n$ ,  $Q_2 = \mathbb{R}^m$ , and

$$\Phi(x, p) = \frac{\sigma_1}{2} \|x\|_2^2 + \langle Ax, p \rangle + \frac{\sigma_2}{2} \|p\|_2^2$$

shows that the condition (2.5) cannot be improved. In this example,  $L_1 = \|A\|$  and  $L_2 = 1$ .

Consider the following minimization problem:

$$\min_{z \in Q} \Phi(z), \tag{2.6}$$

where the parameters of function  $\Phi$  satisfy a strict variant of condition (2.5). In this case, function  $\Phi$  is strongly convex. Consequently, there exists a unique solution

$$z^* = (x^*, p^*) \in Q$$

of the problem (2.6). Let us show that it can be found by the following alternating minimization scheme.

**Initialization.** Choose  $x_0 \in Q_1$ .

$$\begin{aligned} \text{Iteration } t \geq 0. \quad & \text{a) Find } p_{t+1} = \arg \min_{p \in Q_2} \Phi(x_t, p). \\ & \text{b) Find } x_{t+1} = \arg \min_{x \in Q_1} \Phi(x, p_{t+1}). \end{aligned} \tag{2.7}$$

In order to analyze the convergence of this scheme, let us define the following operators:

$$u(x) = \arg \min_{p \in Q_2} \Phi(x, p) \in Q_2,$$

$$v(p) = \arg \min_{x \in Q_1} \Phi(x, p) \in Q_1.$$

In view of Assumption 1, the objective function in both optimization problems above is strongly convex. Hence, the points  $u(\cdot)$  and  $v(\cdot)$  are well defined.

**Lemma 2** *Let the conditions of Theorem 1 be satisfied. Then, for any  $x_1, x_2 \in Q_1$  we have*

$$\|u(x_1) - u(x_2)\| \leq \frac{L_1 L_2}{\sigma_2} \|x_1 - x_2\|. \tag{2.8}$$

Similarly, for any  $p_1$  and  $p_2$  from  $Q_2$  we have

$$\|v(p_1) - v(p_2)\| \leq \frac{L_1 L_2}{\sigma_1} \|p_1 - p_2\|. \tag{2.9}$$

**Proof** Indeed, point  $u(x)$  is the solution of the problem  $\min_{p \in Q_2} \{\langle G_1(x), G_2(p) \rangle + h(p)\}$  with strongly convex objective function. Therefore, in view of the property (1.9), we have

$$\begin{aligned} & \langle G_1(x_1), G_2(u(x_2)) \rangle + h(u(x_2)) \\ & \geq \langle G_1(x_1), G_2(u(x_1)) \rangle + h(u(x_1)) \\ & \quad + \frac{\sigma_2}{2} \|u(x_2) - u(x_1)\|^2, \\ & \langle G_1(x_2), G_2(u(x_1)) \rangle + h(u(x_1)) \\ & \geq \langle G_1(x_2), G_2(u(x_2)) \rangle + h(u(x_2)) \\ & \quad + \frac{\sigma_2}{2} \|u(x_2) - u(x_1)\|^2. \end{aligned}$$

Adding these two inequalities, we get

$$\begin{aligned} & \sigma_2 \|u(x_2) - u(x_1)\|^2 \\ & \leq \langle G_1(x_1) - G_1(x_2), G_2(u(x_2)) - G_2(u(x_1)) \rangle \\ & \leq \|G_1(x_1) - G_1(x_2)\|_{\mathbb{E}}^* \cdot \|G_2(u(x_2)) - G_2(u(x_1))\|_{\mathbb{E}}. \end{aligned}$$

Thus, in view of inequality (2.4), we get (2.8). The proof of Inequality (2.9) follows by the same argument.  $\square$

The following theorem is a trivial consequence of Lemma 2. Define

$$T(x) = v(u(x)), \quad S(p) = u(v(p)).$$

**Theorem 2** *Let  $\lambda \stackrel{\text{def}}{=} \frac{L_1^2 L_2^2}{\sigma_1 \sigma_2} < 1$ . Then,  $T(\cdot)$  and  $S(\cdot)$  are contracting mappings:*

$$\begin{aligned} \|T(x_1) - T(x_2)\| & \leq \lambda \|x_1 - x_2\|, \quad x_1, x_2 \in Q_1, \\ \|S(p_1) - S(p_2)\| & \leq \lambda \|p_1 - p_2\|, \quad p_1, p_2 \in Q_2. \end{aligned} \tag{2.10}$$

Note that in terms of operators  $T$  and  $S$  the process (2.7) can be written in the following way:

$$x_{t+1} = T(x_t), \quad p_{t+1} = S(p_t). \tag{2.11}$$

Therefore, for any  $t \geq 0$ , we have

$$\begin{aligned} \|x_{t+1} - x^*\| & = \|T(x_t) - T(x^*)\| \leq \lambda \|x_t - x^*\| \\ & \leq \dots \leq \lambda^{t+1} \|x_0 - x^*\|. \end{aligned}$$

A similar rate of convergence can be established for the sequence  $\{p_k\}_{k \geq 0}$ . We conclude that the process (2.7) converges linearly to the solution of problem (2.6).

Note that from the viewpoint of standard optimization theory, problem (2.6) has possibly non-differentiable objective function with unbounded derivatives (this is allowed by Assumption 1). It has the only favorable property, the strong convexity of the objective. However, this is not enough for justifying linear rate of convergence of any standard optimization scheme.

In the next section, we consider an application example, where the equilibrium state of the system can be found by minimizing bivariate potential with a multiplicative interaction term.

### 3 Electoral model

Let us show how we can use the machinery developed in Sect. 2 for justifying a soft clustering model based on a stable electoral procedure. The basic elements of our model are interpreted as independent voters possessing some features (opinions). They must be attached to different clusters (political parties), which represent these features in the best possible way.

In our model, we have  $N$  independent voters and  $K$  political parties. Our main assumption is that the voting

results are *random*. Voter  $i$  decides to vote for party  $k$  with probability  $p_i^{(k)}$ . It is convenient to put these probabilities in a vector

$$p_i = \left( p_i^{(1)}, \dots, p_i^{(K)} \right)^T \in \Delta_K.$$

These vectors can be unified in a matrix  $P = (p_1, \dots, p_N) \in \mathbb{R}^{K \times N}$ , which we call the *voting matrix*. At the beginning of the voting process, this matrix is unknown. However, it will be computed as an outcome of a sequence of consecutive elections.

Let us try to explain the results of elections by some quantitative parameters. We assume that an opinion of voter  $i$  can be described by  $m$  different real values (personal preferences), which we put in vector  $v_i \in \mathcal{V} \subseteq \mathbb{R}^m, i = 1, \dots, N$ , where  $\mathcal{V}$  is a closed convex set (e.g.,  $\mathcal{V} = \mathbb{R}_+^m$ ). These vectors are fixed during the whole history of consecutive elections.

At the same time, positions  $x_k \in \mathcal{V}$  of political parties are flexible,  $k = 1, \dots, K$ . After each round of elections, these values can be adjusted for better representing the positions of the voters closely attached to this party. It will be convenient to keep these vectors in a matrix  $X = (x_1, \dots, x_K) \in \mathbb{R}^{m \times K}$ .

Let us fix some *distance function*  $\rho(v, x) \geq 0$ , which is used for measuring the distance between the opinion  $v \in \mathcal{V}$  of a voter and current position  $x \in \mathcal{V}$  of a political party. In what follows, we always assume that for any  $v$  being fixed the function  $\rho(v, \cdot)$  is convex. A natural choice for this function would be  $\rho(v, x) = \|x - v\|$ , where  $\|\cdot\|$  is an arbitrary norm in  $\mathbb{R}^m$ . However, in Sect. 5, we will give more examples with some motivation for their use.<sup>3</sup>

Clearly, the bigger the distance between  $v$  and  $x$  is, the smaller should be the probability that this party will be selected by this particular voter. In our electoral model, for this decision we apply the discrete choice probabilities of *logit model* (e.g., Anderson et al. 1992).

**Assumption 2** Voter  $i$  selects the  $k$ th party with probability

$$p_i^{(k)}(X) = e^{-\rho(v_i, x_k)/\mu} / \left[ \sum_{j=1}^K e^{-\rho(v_i, x_j)/\mu} \right], \quad k = 1, \dots, K, \tag{3.1}$$

where  $\mu \geq 0$  is the flexibility parameter, which represents the volatility of opinions of voters.

Denote by  $P_*(X) = (p_1(X), \dots, p_N(X))$  the corresponding voting matrix.

<sup>3</sup> Note that our theory *does not work* for the standard choice  $\rho(v, x) = \frac{1}{2} \|x - v\|_2^2$ .

In Assumption 2, value  $\mu = 0$  corresponds to the deterministic choice: the voter always chooses a party, which is the closest to his/her opinion. However, usually this parameter is strictly positive.

It is important that the probability vector  $p_i(X)$  has an optimization interpretation. Consider the *entropy function*

$$\eta(p) = \sum_{k=1}^K p^{(k)} \ln p^{(k)}, \quad p \in \mathbb{R}_+^K,$$

It is easy to check (see, for example, Lemma 4 in Nesterov 2005a) that

$$p_i(X) = \arg \min_{p \in \Delta_K} \{ \langle g_i(X), p \rangle + \mu \eta(p) \}, \tag{3.2}$$

where  $g_i(X) = (\rho(v_i, x_1), \dots, \rho(v_i, x_K))^T$ . Note that function  $\eta(\cdot)$  is strongly convex on the standard simplex. Indeed, for any  $p \in \text{int } \Delta_K$  and  $h \in \mathbb{R}^K$  we have

$$\langle \nabla^2 \eta(x) h, h \rangle = \sum_{k=1}^K \frac{(h^{(k)})^2}{p^{(k)}}.$$

At the same time,

$$\min_{p \in \Delta_K} \left\{ \sum_{k=1}^K \frac{(h^{(k)})^2}{p^{(k)}} \right\} = \|h\|_1^2. \tag{3.3}$$

Thus, the entropy function is strongly convex on  $\Delta_K$  in  $\ell_1$ -norm with convexity parameter one.

It remains to describe the behavior of political parties. Each party is able to modify its position in order to attract the maximal number of voters. However, they should not go too far from their *core values*, which we denote by  $c_k \in \mathcal{V}, k = 1, \dots, K$ . In order to measure the distance from the current position of the party and its core value, we introduce a prox-function  $d(x, y)$ . It must satisfy the following conditions:

- $d(x, y) \geq 0$  for all  $x, y \in \mathcal{V}$ .
- For each  $x \in \mathcal{V}$ , function  $d(x, \cdot)$  is strongly convex in the second argument with convexity parameter one:

$$d(x, v) \geq d(x, y) + \langle \nabla_2 d(x, y), v - y \rangle + \frac{1}{2} \|v - y\|^2, \quad \forall v, y \in \mathcal{V}, \tag{3.4}$$

where  $\|\cdot\|$  is an arbitrary norm in  $\mathbb{R}^m$ .

Let us give several examples of the most important prox-functions.

- *Kullback–Leibler divergence*

$$\begin{aligned} d_1(x, y) &= \eta(y) - \eta(x) - \langle \eta(x), y - x \rangle \\ &= \sum_{k=1}^K y^{(k)} \ln \frac{y^{(k)}}{x^{(k)}}, \quad x, y \in \Delta_K. \end{aligned}$$

This function is strongly convex in  $y$  in  $\ell_1$ -norm.

- *Euclidean distance*  $d_2(x, y) = \frac{1}{2}\|x - y\|_2^2, x, y \in \mathbb{R}^K$ . This function is strongly convex in  $\ell_2$ -norm.
- We can take  $\tilde{d}_i(x, y) = d_i(x, y) + \epsilon\|x - y\|_i$  with  $\epsilon \geq 0, i = 1, 2$ . The additional linear term gives to a party more chances to keep its core values unchanged.

The behavior of political parties is described by the following assumption.

**Assumption 3** For a given voting matrix  $P = (p_1, \dots, p_N) \in \Delta_K^N$ , each political party chooses its optimal current position by minimizing the function

$$\psi_k(P, x_k) \stackrel{\text{def}}{=} \sum_{i=1}^N p_i^{(k)} \rho(v_i, x_k) + \frac{1}{\tau} d(c_k, x_k), \quad k = 1, \dots, K, \tag{3.5}$$

in  $x_k \in \mathcal{V}$ .

In this definition,  $\tau > 0$  is a *group tolerance* parameter.<sup>4</sup> The objective function (3.5) has an interpretation of the expected distance between the opinions of all attracted voters and the current position of the  $k$ th party, augmented by the discrepancy with its core values.

Since  $\psi_k(P, \cdot)$  is strongly convex, there exists its unique minimum  $x_k^*(P)$  over  $\mathcal{V}$ . Denote  $X_*(P) = (x_1^*(P), \dots, x_N^*(P))$ .

Consider now the following process of sequential elections:

**Set**  $X_0 = (c_1, \dots, c_K)$ . **Repeat** :  $P_{t+1} = P_*(X_t), X_{t+1} = X_*(P_{t+1}), t \geq 0$ . (3.6)

The interpretation of process (3.6) is straightforward:

*Given the current positions of political parties  $X_t$ , voters announce their preferences  $P_{t+1}$  during the electoral poll. After observing the results, parties update their positions  $X_{t+1}$  for the next elections.*

**Definition 1** Electoral system is called *stable* if the process (3.6) has a unique limiting point, which is independent on the starting position  $X_0 \in \mathcal{V}^K$ .

In the next section, we derive some sufficient conditions of the electoral stability. Note that in our model we have two tolerance parameters,  $\mu$  and  $\tau$ .

<sup>4</sup> Later on in Sect. 5, we will replace this parameter by the *individual* tolerance. Note that the group tolerance must be smaller than the individual one in order to resist well to a cumulative opinion of a group.

### 4 Stable electoral systems

Let us describe the electoral process (3.6) using the framework of Sect. 2. Denote  $\mathbb{E}_1 = \mathbb{R}^{m \times K}$  the space for positions  $X = (x_1, \dots, x_K)$  of the parties. Then,  $\mathcal{Q}_1 = \mathcal{V}^K$ . For the voting matrix  $P = (p_1, \dots, p_N)$  we introduce the space  $\mathbb{E}_2 = \mathbb{R}^{K \times N}$  and the feasible set  $\mathcal{Q}_2 = \Delta_K^N$ . For the interaction term, we have  $\mathbb{E} = \mathbb{E}_2, G_2(P) \equiv P$ , and  $g(X, P) = \langle G_1(X), P \rangle_{\mathbb{E}_2}$ , where the distance matrix

$$G_1(X) = (g_1(X), \dots, g_N(X))$$

is component-wise convex in  $X$ . Since all matrices  $P \in \mathcal{Q}_2$  have nonnegative elements, function  $g(\cdot, \cdot)$  satisfy Assumption 1.

In view of the rules (3.2) and (3.5), the electoral process (3.6) can be seen as alternating minimization scheme (2.7) being applied to the following bivariate potential

$$\begin{aligned} \Phi(X, P) &= \frac{1}{\tau} f(X) + \langle G_1(X), P \rangle_{\mathbb{E}_2} + \mu h(P), \\ f(X) &= \sum_{k=1}^K d(c_k, x_k), \\ h(P) &= \sum_{i=1}^N \eta(p_i). \end{aligned} \tag{4.1}$$

Now we need to choose appropriate norms in  $\mathbb{E}_1$  and  $\mathbb{E}_2$ , which ensure good values of convexity parameters for function  $f$  and  $h$ . For  $X \in \mathcal{Q}_1$  and direction  $U \in \mathbb{E}_1$ , we have

$$\langle \nabla^2 f(X) U, U \rangle \stackrel{(4.1)}{=} \sum_{k=1}^K \langle \nabla^2 d(c_k, x_k) u_k, u_k \rangle \stackrel{(3.4)}{\geq} \sum_{k=1}^K \|u_k\|^2.$$

Thus, the natural choice of the norm for  $\mathbb{E}_1$  is as follows:

$$\|U\|_{\mathbb{E}_1} = \left[ \sum_{k=1}^K \|u_k\|^2 \right]^{1/2}, \quad U \in \mathbb{E}_1. \tag{4.2}$$

In this case, the convexity parameter of function  $f$  is one, and we can choose  $\sigma_1 = \frac{1}{\tau}$ .

For  $P \in \mathcal{Q}_2$  and direction  $H \in \mathbb{E}_2$ , we have

$$\langle \nabla^2 h(P) H, H \rangle \stackrel{(4.1)}{=} \sum_{i=1}^N \langle \nabla^2 \eta(p_i) h_i, h_i \rangle \stackrel{(3.3)}{\geq} \sum_{i=1}^N \|h_i\|_1^2.$$

Therefore, the natural choice for the norm in  $\mathbb{E}_2$  is

$$\|H\|_{\mathbb{E}_2} = \left[ \sum_{i=1}^N \|h_i\|_1^2 \right]^{1/2}, \quad H \in \mathbb{E}_2.$$

In this case, function  $h$  has convexity parameter one and we can choose  $\sigma_2 = \mu$ .

It remains to estimate the Lipschitz constants for inequality (2.4). In our case,  $L_2 = 1$ . For estimating  $L_1$ , we

need to compute first the dual norm  $\|\cdot\|_{\mathbb{E}_2}^*$ . For arbitrary  $S \in \mathbb{R}^{m \times N}$ , we have

$$\begin{aligned} & \max_{\|H\|_{\mathbb{E}_2} \leq 1} \langle S, H \rangle \\ &= \max_{H, \tau} \left\{ \sum_{i=1}^N \langle s_i, h_i \rangle : \|h_i\|_1 \leq \tau^{(i)}, i = 1, \dots, N, \|\tau\|_2 \leq 1 \right\} \\ &= \max_{\tau} \left\{ \sum_{i=1}^N \tau^{(i)} \|s_i\|_{\infty} : \|\tau\|_2 \leq 1 \right\}. \end{aligned}$$

Thus, we get the following dual norm:

$$\|S\|_{\mathbb{E}_2}^* = \left[ \sum_{i=1}^N \|s_i\|_{\infty}^2 \right]^{1/2}, \quad S \in \mathbb{R}^{m \times N}.$$

Now we can get the Lipschitz constant of the mapping  $G_1(\cdot)$ . For that, we need one more assumption.

**Assumption 4** For any  $v \in \mathcal{V}$ , function  $\rho(v, \cdot)$  is Lipschitz continuous in its second argument with constant one.

Then, for  $X$  and  $Y$  from  $\mathcal{Q}_1$ , we have

$$\|G_1(X) - G_1(Y)\|_{\mathbb{E}_2}^* = \left[ \sum_{i=1}^N \|g_i(X) - g_i(Y)\|_{\infty}^2 \right]^{1/2}.$$

Note that

$$\begin{aligned} g_i(X) &= (\rho(v_i, x_1), \dots, \rho(v_i, x_K)), \\ g_i(Y) &= (\rho(v_i, y_1), \dots, \rho(v_i, y_K)). \end{aligned}$$

Therefore, in view of Assumption 4,

$$\begin{aligned} & \|g_i(X) - g_i(Y)\|_{\infty}^2 \\ &= \max_{1 \leq k \leq K} |\rho(v_i, x_k) - \rho(v_i, y_k)| \leq \max_{1 \leq k \leq K} \|x_k - y_k\|. \end{aligned}$$

Hence,

$$\begin{aligned} & \|G_1(X) - G_1(Y)\|_{\mathbb{E}_2}^* \\ & \leq N^{1/2} \max_{1 \leq k \leq K} \|x_k - y_k\| \stackrel{(4.2)}{\leq} N^{1/2} \|X - Y\|_{\mathbb{E}_1}. \end{aligned}$$

Thus, mapping  $G_1$  is Lipschitz continuous with constant  $L_1 = N^{1/2}$ . Now, using Theorem 2, we come to the following statement.

**Theorem 3** *Let the behavior of voters and political parties satisfy Assumptions 2, 3 with tolerance parameters*

$$N < \frac{\mu}{\tau}. \tag{4.3}$$

Then, the corresponding electoral system is stable. Moreover, for the stationary voting matrix  $P_* = \lim_{t \rightarrow \infty} P_t$ , where matrices  $\{P_t\}$  are generated by the process (3.6), we have

$$\|P_t - P_*\|_{\mathbb{E}_2} \leq \lambda^t N^{1/2}, \quad t \geq 0. \tag{4.4}$$

**Proof** Note that condition (4.3) is a strict version of condition (2.5). In terms of Theorem 2, we have  $P_{t+1} = S(P_t)$  and

$$\|S(P_1) - S(P_2)\| \leq \lambda \|P_1 - P_2\| \quad \forall P_1, P_2 \in \mathcal{Q}_2.$$

Note that for any  $p_1, p_2 \in \Delta_K$  we have  $\|p_1 - p_2\|_1 \leq 1$ . Therefore, for any  $P \in \mathcal{Q}_2$  it holds  $\|P - P_*\| \leq N^{1/2}$ . Thus, by Theorem 2, for all  $k \geq 1$  we get  $\|P_t - P_*\| \leq \lambda^t N^{1/2}$ .  $\square$

### 5 Computational aspects of alternating minimization

Let us find an interpretation of the stability condition (4.3). For that, we introduce the *individual tolerance* parameter  $\hat{\tau} = \tau N$ .

Then, the optimization problem (3.5), defining the response of the party on the voting probabilities, can be rewritten as

$$\min_{x_k \in \mathcal{V}} \left\{ \hat{\psi}_k(P, x_k) = \frac{1}{N} \sum_{i=1}^N p_i^{(k)} \rho(v_i, x_k) + \frac{1}{\hat{\tau}} d(c_k, x_k) \right\}. \tag{5.1}$$

This means that each party reacts on the average opinion of the attracted voters. With this notation, the stability condition  $\lambda < 1$  can be rewritten as

$$\hat{\tau} < \mu. \tag{5.2}$$

In view of its importance, we rewrite this condition in form of a theorem.

**Theorem 4** *Electoral system is stable if the individual tolerance of parties is smaller than the volatility of voters.*

This condition looks indeed natural for guaranteeing stability of electoral system: parties must be more conservative in changing their opinions than the voters. In other words, an excessive populism of parties may cause political instability.

Let us discuss now the computational complexity of the electoral process. For simplicity, we assume that the set  $\mathcal{V}$  is bounded:  $D \stackrel{\text{def}}{=} \max_{x, y \in \mathcal{V}} \|x - y\| < \infty$ . Each step in the recurrence (3.6) consists of two time-consuming operations.

- *Computation of the optimal voting matrix  $P_*(X)$ .* For that, we need to compute first the distance matrix  $G_1(X)$ . If the cost of computation of the value  $\rho(v, x)$  is  $O(m)$  operation, then the distance matrix can be computed in  $O(KmN)$  operations. After that, we can

apply the rule (3.1), which is cheaper:  $O(KN)$  operations. Thus, in total we need

$$T_1 = O(KmN)$$

operations.

- *Computation of new positions of the parties  $X_*(P)$ .* For that, we need to solve  $K$  optimization problems (5.1). Note that just one computation of the objective functions for all these problems needs  $O(KmN)$  operations. Therefore, the importance of this step for our approach is dominating. Complexity of problem (5.1) depends on efficiency of the applied optimization schemes. And their potential efficiency depends on the properties of function  $\rho(v, \cdot)$ . Let us discuss these aspects in more details. Note that the objective function in this problem is strongly convex with parameter

$$\hat{\sigma} = \frac{1}{\hat{\tau}}.$$

In our approach, in view of Assumption 4, we cannot follow the standard advice to use

$$\rho(v, x) = \frac{1}{2} \|x - v\|_2^2, \quad d(c, x) = \frac{1}{2} \|x - c\|_2^2.$$

Recall that this choice is mainly motivated by an extremely low cost of solving the problem (5.1) (only  $O(mN)$  operations). However, we pay for that by non-convexity of the potential, existence of many local minima and most probably by NP-hardness of finding the global solution.

Thus, we need to choose a Lipschitz continuous distance function. Let us look at the most reasonable variants.

1. *Arbitrary norm*  $\rho_1(v, x) = \|v - x\|$ . This function clearly satisfies Assumption 4. With this choice, problem (5.1) becomes a problem of convex optimization with non-differentiable strongly convex objective function. From the general theory, we know that  $\epsilon$ -solution of this problem can be found in

$$O\left(\frac{\hat{\tau}}{\epsilon} M^2\right) \tag{5.3}$$

iterations, where  $M$  is an upper bound for the norms of subgradients gradients of function  $\hat{\psi}_k(P, \cdot)$ , and  $\epsilon$  is the required accuracy of solving the problem (5.1). Since the objective function of this problem has implicit structure, we can apply the *smoothing technique* (see Nesterov 2005a, b) for finding an  $\epsilon$ -solution of this problem in  $O\left(\sqrt{\frac{\hat{\tau}}{\epsilon}} M\right)$  iterations. This bound is much better than (5.3). In order to estimate the total computational time, we need to multiply this bound by  $K$ , the number of parties, and by  $mN$ , the complexity of computing the value and subgradient of function

$\hat{\psi}_k(P, \cdot)$ . Thus, we get the following estimate for computational time:

$$T_2 = \sqrt{\frac{\hat{\tau}}{\epsilon}} M \cdot KmN.$$

Note that this bound depends on two parameters, which can be potentially big. The accuracy  $\epsilon$  must be chosen proportionally of the required accuracy of approximation of matrix  $P_*$ .<sup>5</sup> Thus, the impact of the term  $\frac{1}{\sqrt{\epsilon}}$  can be significant. The constant  $M$  can be estimated from above as follows:

$$\begin{aligned} M &= \max_{x \in \mathcal{V}} \|\nabla_2 \hat{\psi}_k(P, x)\|^* \\ &\stackrel{(5.1)}{\leq} \frac{1}{N} \sum_{i=1}^N p_i^{(k)} + \frac{1}{\hat{\tau}} \max_{x \in \mathcal{V}} \|\nabla_2 d(c_k, x)\|^* \\ &\leq 1 + \frac{1}{\hat{\tau}} \max_{x \in \mathcal{V}} \|\nabla_2 d(c_k, x)\|^*. \end{aligned}$$

Thus,  $M$  can be big if the diameter  $D$  is big.

2. *Smoothed Euclidean norm*  $\rho_2(v, x) = \sqrt{\delta^2 + \|v - x\|_2^2}$ , where  $\delta > 0$  is a smoothing parameter. In this case, it is natural to choose  $d(c, x) = \frac{1}{2} \|x - c\|_2^2$ . Note that  $\rho_2(v, \cdot)$  is infinitely times differentiable in its second argument. Moreover,

$$\begin{aligned} \nabla_2 \rho(v, x) &= \frac{x - v}{[\delta^2 + \|x - v\|_2^2]^{1/2}}, \quad \|\nabla_2 \rho(v, x)\|_2 \leq 1, \\ \nabla_2^2 \rho(v, x) &= \frac{I}{[\delta^2 + \|x - v\|_2^2]^{1/2}} - \frac{(x - v)(x - v)^T}{[\delta^2 + \|x - v\|_2^2]^{3/2}} \preceq \frac{1}{\delta} I. \end{aligned} \tag{5.4}$$

Thus, function  $\rho_2(v, \cdot)$  satisfies Assumption 4, and it has Lipschitz continuous gradient with constant  $\frac{1}{\delta}$ . Consequently, function  $\hat{\psi}_k(c_k, \cdot)$  has Lipschitz continuous gradient with the constant  $L = \frac{1}{\delta} + \frac{1}{\hat{\tau}}$ . Since this function is strongly convex with parameter  $\frac{1}{\hat{\tau}}$ , we conclude that its condition number is equal to

$$\kappa = L/\hat{\sigma} = 1 + \frac{\hat{\tau}}{\delta}.$$

Thus, problem (5.1) can be solved by the usual gradient method in  $O\left(\kappa \ln \frac{LD^2}{\epsilon}\right)$  iterations. For the fast gradient methods, we get the bound of  $O\left(\kappa^{1/2} \ln \frac{LD^2}{\epsilon}\right)$  iterations (see Sections 2.1 and 2.2 in Nesterov 2004). For the latter case, we get the following estimate of computational time:

<sup>5</sup> It is possible to relate these two accuracies by a rigorous mathematical condition. But its justification is quite technical and we decided to drop it in this paper.

$$T_3 = O\left(\sqrt{1 + \frac{\hat{\tau}}{\delta} \ln \frac{LD^2}{\epsilon}} \cdot KmN\right). \tag{5.5}$$

Clearly, this estimate is much better than  $T_2$ .

Note that Euclidean norm can be smoothed in many different ways. From the statistical point of view, sometimes it is reasonable to apply the *Huber function* which is used in *robust regression* for better resistance to outliers:

$$\rho_H(v, x) = \chi_\delta(\|x - v\|_2),$$

$$\chi_\delta(\tau) = \begin{cases} \tau - \frac{1}{2}\delta, & \text{if } \tau > \delta, \\ \frac{1}{2\delta}\tau^2, & \text{otherwise,} \end{cases} \quad \tau \geq 0.$$

It has the same properties as function  $\rho_2(v, \cdot)$  with the same estimate  $T_3$  for the computational time of one iteration of the process (3.6).

Recall that for estimating the total time, which is necessary for procedure (3.6) to converge to an approximate solution of problem

$$\min_{X, P} \{ \Phi(X, P) : X \in \mathcal{V}^K, P \in \Delta_K^N \}, \tag{5.6}$$

we need to multiply the bounds  $T_i$  by an estimate for the total number of steps of the scheme. In view of Theorem 3, in order to get  $\|P_k - P_*\|_{\mathbb{E}_2} \leq \epsilon_P$ , we need at most

$$1 + \frac{1}{\ln \frac{1}{\lambda}} \ln \frac{N^{1/2}}{(1 - \lambda)\epsilon_P} \tag{5.7}$$

iterations of the process (3.6).

### 6 Simplified soft clustering

In the previous section, we derived complexity bounds for the iterative process (3.6), which has interpretation of the sequential election procedure. However, the best complexity estimate (5.5), (5.7) is still quite heavy. One of the possibilities for improving the complexity bounds consists in considering direct methods for solving the minimization problem (5.6). Note that in its current form this problem is not simple. The reason is that function  $h(\cdot)$  in (4.1) has unbounded derivatives at the feasible set. Therefore, in this section we study a possibility of replacing it by a squared Euclidean norm. Of course, in this case we will lose the probabilistic interpretation of the soft assignment (3.1). However, from the viewpoint of clustering, the new potential is still useful for computing the attachment coefficients  $p_i^{(k)}(X)$ .

Consider the following minimization problem:

$$\min_{X, P} \{ \Phi_2(X, P) : X \in \mathcal{V}^K, P \in \Delta_K^N \},$$

$$\Phi_2(X, P) = \frac{1}{\hat{\tau}} f(X) + \frac{1}{N} \langle G_1(X), P \rangle_{\mathbb{E}_2} + \frac{\mu}{N} h(P), \tag{6.1}$$

where  $f(X) = \frac{1}{2} \sum_{k=1}^K \|x_k - c_k\|_2^2$ ,  $h(P) = \frac{1}{2} \sum_{i=1}^N \|p_i\|_2^2$ , and<sup>6</sup>

$$G_1(X)^{(k,i)} = \rho_2(v_i, x_k), \quad k = 1, \dots, K, \quad i = 1, \dots, N.$$

This potential has three positive parameters  $\hat{\tau}$ ,  $\mu$  and  $\delta$  (see Item 2 above), which we assume to satisfy the following relations:

$$\hat{\tau} < \delta < \mu. \tag{6.2}$$

Let us prove that under these conditions the potential  $\Phi_2(\cdot, \cdot)$  is a strongly convex function with Lipschitz continuous gradient.

Let us fix a direction  $Z = (U, H) \in \mathbb{E} = \mathbb{E}_1 \times \mathbb{E}_2$  and choose

$$\|U\|_{\mathbb{E}_1}^2 = \sum_{k=1}^K \|u_k\|_2^2, \quad \|H\|_{\mathbb{E}_2}^2 = \sum_{i=1}^N \|h_i\|_2^2.$$

We need to estimate from above and below the second derivative of the potential  $\Phi_2(X, P)$  along direction  $Z$ . Note that<sup>7</sup>

$$\begin{aligned} & \langle \nabla^2 \Phi(X, P) Z, Z \rangle_{\mathbb{E}} \\ &= \frac{1}{\hat{\tau}} \|U\|_{\mathbb{E}_1}^2 + \frac{1}{N} \sum_{k=1}^K \sum_{i=1}^N P^{(k,i)} \langle \nabla_2^2 \rho_2(v_i, x_k) u_k, u_k \rangle \\ & \quad + \frac{2}{N} \sum_{k=1}^K \sum_{i=1}^N H^{(k,i)} \langle \nabla_2 \rho_2(v_i, x_k), u_k \rangle + \frac{\mu}{N} \|H\|_{\mathbb{E}_2}^2. \end{aligned}$$

Since  $P \geq 0$ , we have

$$\begin{aligned} & \sum_{k=1}^K \sum_{i=1}^N P^{(k,i)} \langle \nabla_2^2 \rho_2(v_i, x_k) u_k, u_k \rangle \\ & \stackrel{(5.4)}{\leq} \sum_{k=1}^K \sum_{i=1}^N \frac{1}{\delta} \|u_k\|_2^2 = \frac{N}{\delta} \|U\|_{\mathbb{E}_2}^2. \end{aligned}$$

At the same time,

$$\begin{aligned} & \left| \sum_{k=1}^K \sum_{i=1}^N H^{(k,i)} \langle \nabla_2 \rho_2(v_i, x_k), u_k \rangle \right| \\ & \stackrel{(5.4)}{\leq} \sum_{k=1}^K \sum_{i=1}^N |H^{(k,i)}| \cdot \|u_k\|_2 \\ & \leq \left[ \sum_{k=1}^K \sum_{i=1}^N (H^{(k,i)})^2 \right]^{1/2} \left[ \sum_{k=1}^K \sum_{i=1}^N \|u_k\|_2^2 \right]^{1/2} \\ & = N^{1/2} \|U\|_{\mathbb{E}_1} \|H\|_{\mathbb{E}_2}. \end{aligned} \tag{6.3}$$

<sup>6</sup> All our further conclusions are also valid for Huber distance function  $\rho_H(\cdot, \cdot)$ .

<sup>7</sup> For the interaction term of the potential  $\Phi_2$ , we use the standard differentiation rule

$(g(x)p)'' = g''(x)(x')^2 p + 2g'(x)x'p' + g(x)p''$ , taking into account that  $x' = u, p' = h$  and  $x'' = 0, p'' = 0$ .

Thus,

$$\begin{aligned} &\langle \nabla^2 \Phi(X, P)Z, Z \rangle_{\mathbb{E}} \\ &\leq \left(\frac{1}{\hat{\tau}} + \frac{1}{\delta}\right) \|U\|_{\mathbb{E}_1}^2 + \frac{2}{N^{1/2}} \|U\|_{\mathbb{E}_1} \|H\|_{\mathbb{E}_2} + \frac{\mu}{N} \|H\|_{\mathbb{E}_2}^2 \\ &\leq \left(\frac{1}{\hat{\tau}} + \frac{2}{\delta}\right) \|U\|_{\mathbb{E}_1}^2 + \frac{\mu + \delta}{N} \|H\|_{\mathbb{E}_2}^2 \end{aligned} \tag{6.4}$$

On the other hand, since  $P \geq 0$  and function  $\rho_2(v_i, \cdot)$  is convex, we have

$$\begin{aligned} &\langle \nabla^2 \Phi(X, P)Z, Z \rangle_{\mathbb{E}} \\ &\geq \frac{1}{\hat{\tau}} \|U\|_{\mathbb{E}_1}^2 - \frac{2}{N^{1/2}} \|U\|_{\mathbb{E}_1} \|H\|_{\mathbb{E}_2} + \frac{\mu}{N} \|H\|_{\mathbb{E}_2}^2 \\ &\geq \left(\frac{1}{\hat{\tau}} - \frac{1}{\delta}\right) \|U\|_{\mathbb{E}_1}^2 + \frac{\mu - \delta}{N} \|H\|_{\mathbb{E}_2}^2. \end{aligned}$$

In view of (6.2), we can choose in  $\mathbb{E}$  the following Euclidean norm:

$$\begin{aligned} \|Z\|_{\mathbb{E}}^2 &= \left(\frac{1}{\hat{\tau}} - \frac{1}{\delta}\right) \|U\|_{\mathbb{E}_1}^2 + \frac{\mu - \delta}{N} \|H\|_{\mathbb{E}_2}^2, \\ Z = (U, H) \in \mathbb{E} &= \mathbb{E}_1 \times \mathbb{E}_2. \end{aligned} \tag{6.5}$$

As we have seen, with respect to this norm this norm, potential  $\Phi_2$  is strongly convex on  $\mathbb{E}$  with convexity parameter one. On the other hand, it has Lipschitz continuous gradient with constant

$$L_2 = \max \left\{ \frac{\delta + 2\hat{\tau}}{\delta - \hat{\tau}}, \frac{\mu + \delta}{\mu - \delta} \right\}. \tag{6.6}$$

Consequently, the condition number of this function is  $\kappa_2 = L_2$ .

If there is no other reason for choosing the parameter  $\delta$ , satisfying condition (6.2), we can try to use this degree of freedom for making the value  $\kappa_2$  as small as possible. For that, it is convenient to introduce another representation of our parameters. Let us choose two factors  $\gamma$  and  $\gamma_1$  from  $(0, 1)$  and define

$$\delta = \gamma\mu, \quad \hat{\tau} = \gamma_1\delta. \tag{6.7}$$

Then,  $\frac{\delta + 2\hat{\tau}}{\delta - \hat{\tau}} = \frac{\mu + \delta}{\mu - \delta}$  if and only if  $\frac{1 + 2\gamma_1}{1 - \gamma_1} = \frac{1 + \gamma}{1 - \gamma}$ . Thus, we can choose  $\gamma_1 = \frac{2\gamma}{3 - \gamma}$ . Then,

$$\delta = \gamma\mu, \quad \hat{\tau} = \frac{2\gamma^2\mu}{3 - \gamma}, \quad \kappa_2^* = \frac{1 + \gamma}{1 - \gamma}. \tag{6.8}$$

Thus, under the choice of parameters (6.8), problem (6.1) is very easy for the gradient methods. It can be solved by the usual gradient method up to accuracy  $\hat{\epsilon}$  in the function value in

$$O\left(\kappa_2^* \ln \frac{\kappa_2^* \delta_f}{\hat{\epsilon}}\right)$$

iterations, where  $\delta_f$  is the initial residual in function value. If we apply the fast gradient methods, the efficiency estimate is even better:

$$O\left([\kappa_2^*]^{1/2} \ln \frac{\kappa_2^* \delta_f}{\hat{\epsilon}}\right)$$

(see Section 2 in Nesterov 2004). Note that the cost of each iteration of such schemes is  $O(KmN)$  operations. Thus, we get indeed very efficient clustering procedures.

Let us look at the rule of forming the soft membership vectors  $p_i(X) \in \Delta_K, i = 1, \dots, N$ , imposed by the potential  $\Phi_2$ . For that, given the position matrix  $X$ , we need to minimize the potential  $\Phi_2(X, \cdot)$ :

$$\min_{P \in \Delta_K^N} \left\{ \frac{1}{N} \langle G_1(X), P \rangle_{\mathbb{E}_2} + \frac{\mu}{2N} \|P\|_{\mathbb{E}_2}^2 \right\}.$$

This means that for each element  $i, 1 \leq i \leq N$ , we need to solve the problem

$$\min_{p_i \in \Delta_K} \left\{ \langle g_i(\bar{X}), p_i \rangle_{\mathbb{R}^K} + \frac{1}{2} \mu \|p_i\|_2^2 \right\}, \tag{6.9}$$

where  $g_i^{(k)}(\bar{X}) = \rho_2(v_i, \bar{x}_k), k = 1, \dots, K$ . Note that the objective function in the latter problem is as follows:

$$\frac{\mu}{2} \left\| p_i + \frac{1}{\mu} g_i(\bar{X}) \right\|_2^2 - \frac{1}{2\mu} \|g_i(\bar{X})\|_2^2.$$

Thus, the solution of this problem is a Euclidean projection of the point

$$\tilde{p}_i = -\frac{1}{\mu} g_i(\bar{X})$$

onto the standard simplex  $\Delta_K$ . This point can be found in  $O(K \ln K)$  operations since we need to order the entries of vector  $\tilde{p}_i$ . If  $K$  is not big, the complexity of straightforward ordering  $O(K^2)$  is also acceptable. Note that for  $\mu \rightarrow 0$  the rule (6.9) converges to the deterministic choice of the entry of vector  $g_i(X) \in \mathbb{R}^K$  with the minimal value.

**Acknowledgements** The author would like to thank three anonymous referees for very useful comments and suggestions.

**Funding** This study was funded by the Advanced Grant 788368 of the European Research Council.

### Compliance with ethical standards

**Conflict of interest** The author declares that he has no conflict of interest.

**Ethical approval** This article does not contain any studies with human participants or animals performed by any of the authors.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Aloise D, Deshpande A, Hansen P, Popat P (2009) NP-hardness of Euclidean sum-of-squares clustering. *Mach Learn* 75(2):245–249
- Anderson SP, de Palma A, Thisse J-F (1992) *Discrete choice theory of product differentiation*. MIT Press, Cambridge
- Bezdek JC (1981) *Pattern recognition with fuzzy objective function algorithms*. Plenum Press, New York
- Bora DJ, Gupta AK (2014) A comparative study between fuzzy clustering algorithm and hard clustering algorithm. *Int J Comput Trends Technol* 10(2):108–113
- Dasgupta S, Freund Y (2009) Random projection trees for vector quantization. *IEEE Trans Inf Theory* 55(7):3229–3242
- Dunn JC (1973) A fuzzy relative of the ISODATA process and its use in detecting well-separated clusters. *J Cybern* 3(3):32–57
- Forgy EW (1965) Cluster analysis of multivariate data: efficiency versus interpretability of classifications. *Biometrics* 21:768–769
- Garey M, Johnson D, Witsenhausen H (1982) The complexity of the generalized Lloyd-max problem. *IEEE Trans Inf Theory* 28(2):255–256
- Kleinberg J, Papadimitriou C, Raghavan P (1998) A microeconomic view of data mining. *Data Min Knowl Disc* 2(4):311–324
- Lloyd SP (1957) *Least square quantization in PCM*. Bell Telephone Laboratories Press, Murray Hill
- MacQueen JC (1967) Some methods for classification and analysis of multivariate observations. In: *Proceedings of 5th Berkeley symposium on mathematical statistics and probability*, vol 1. University of California Press, pp 281–291
- Nesterov Y (2004) *Introductory lectures on convex optimization*. Kluwer, Boston
- Nesterov Y (2005) Smooth minimization of non-smooth functions. *Math Program* 103(1):127–152
- Nesterov Y (2005) Excessive gap technique in nonsmooth convex minimization. *SIAM J Optim* 16(1):235–249
- Peters G, Crespo F, Lingras P, Weber R (2013) Soft clustering—fuzzy and rough approaches and their extensions and derivatives. *Int J Approx Reason* 54:317–322
- Steinhaus H (1957) Sur la division des corps matériels en parties. *Bull Acad Polon Sci* 4(12):801–804 (in French)
- Xu R, Wunsch D (2005) Survey of clustering algorithms. *IEEE Trans Neural Netw* 16(3):645–678
- Yang M-S (1993) A survey on fuzzy clustering. *Math Comput Model* 18(11):1–16

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.