On The Maximum Principle for Some Set Functions [1]

Abstract. This article discusses the problem of finding extreme points for functions defined on all subsets of some large or general finite set. The construction method leads to the detection of extreme subsets. The main feature of the construction method is based on the assumption that on each subset, and for each of its elements, some numbers are given, i.e. credentials or weights, satisfying the monotonicity conditions p.1 and p.2.
Keywords: classification; graphs; convex functions; algorithm

1. Introduction [NB!]

In our study, we consider the problem of finding the global extremum of a function defined on all subsets of a given finite set. The described construction algorithm was used to solve some problems of object classification using the technique of homogeneous Markov chains. In general terms, the proposed construction allows one to solve some problems on graphs, for example, to single out, in a sense, "connected" subsets of the vertices of the graph. We formulate the theoretical foundations of our construction in terms of transparent rules for choosing subsets in a given finite set and some sequences of the same elements of a finite set. The result will be extracting the extreme subsets.

The types of problems of similar nature have a combinatorial character and do belong mostly to the discrete programming problems. Cherenin (1962), Cherenin and Hachaturov (1965) have successfully solved a preeminent class of similar problems on the finite sets. In the framework of these papers a functions have been considered satisfying condition, which can be formulated as follows. If ω_1 and ω_2 are two representatives for subsets of a given finite set then $f(\omega_1)+f(\omega_2)\le f(\omega_1\cup\omega_2)+f(\omega_1\cap\omega_2)$.

This condition with some reservation reflects the convexity of the function f.

---

[1] This idea at the moment, perhaps invisible from the first glance, is incorporated into "Left- and Right-Wing Political Power Design" as political parties bargaining game. Reg. "data analysis", see also, J. E. Mullat (1976-1977) Extremal Subsystems of Monotonic Systems, I,II,III, Automation and Remote Control, 37, pp. 758-766, 37, pp. 1286-1294; 38. pp. 89-96.

The main property or requirement for the class of functions considered in the manuscript is the assumption of the existence of some numbers or weights that reveal for each element of a finite set the degree of its occurrence in the subset. The degree of occurrence must satisfy conditions (1) and (2), see below.

Concerning the current investigation it is worthwhile also to pay attention to Mirkin's (1970) work. In this work, a problem of optimal classification is reduced to finding special "painting" on a non-ordered graph. The optimal classification there is characterized by some maximum value of a function, corresponding in its form to the definition (1), however hereby we interpret (1) in a different sense. We do not consider in our function definition a decomposition of a given set into two non-intersecting subsets what was the main concern of Mirkin's work.

2. The Model

Let $\{H\}$ is a set of subsets of some finite set W. Suppose that we introduce a π_H function for each set $H \subseteq W$ of its elements as arguments. Below by the collection $\{\pi_H\}$ we entitle a system of weights on the set H. The main supposition concerning the weight systems $\{\{\pi_H\}\}$ is as follows:

p.1 the credential $\pi_H(\alpha)$ of the element $\alpha \in H$ is a real number.

p.2 Following dependencies inhere between different credential, i.e., credential systems for different subsets of the set M: for each element $\alpha \in H$ and each $\beta \in H \setminus \{\alpha\}$ yields that $\pi_{H\setminus\alpha}(\beta) \le \pi_H(\alpha)$.

In other words, following p.2, the requirement is that a removal of an arbitrary element α from a set H results in a new credential system $\{\pi_{H\setminus\alpha}\}$ and the effect of the removed element α on the credentials within the remaining part $H \setminus \{\alpha\}$ is only towards the direction of a decrease. We explain these two conditions by examples from the graph theory, although there are examples from other jurisdictions, however less convenient for a short discussion. Let consider non-oriented graphs, i.e., graphs with the property when a relation of a vertex x to y implies a reverse relation of vertex y to x.

Example 1. [2 3]

Let W is a vertex set of a graph G. We define a credential system $\{\pi_H\}$ on each subset of vertexes H as a collection of numbers $\{\pi_H(\alpha)\}$, where the number $\pi_H(\alpha)$ is equal to the number of vertexes in H related to the vertex α. The truthfulness of the pp. 1 and 2 is easily checked, if one only remembers to recall that together with the removal of a vertex α all connected to it edges have to be removed concurrently.

Example 2.

Let W is a set of edges in a graph G or the set of pairs of vertexes related by the graph G. We define a credential system $\{\pi_H\}$ on arbitrary subset H of edges in the graph G as a collection of numbers $\{\pi_H(\alpha)\}$, where $\alpha \in H$ and $\pi_H(\alpha)$ is a number of triangles in the set of edges H, containing the edge α. The number $\pi_H(\alpha)$ is equal to the number of those vertexes on which the set H resides such, that if x is a pointed vertex and the edge $\alpha = [b,e]$, then it ensues that $[b,x] \in H$ and $[e,x] \in H$.

In the examples, we have exploited the fact, that a graph is a topological object from one side and a binary relation from the other side. Let now consider the following set function

$$f(H) = \min_{\alpha \in H} \pi_H(\alpha), \tag{1}$$

where $H \subseteq W$. We suggest below a principle, valid for the subset H, on which the global maximum of a type (1) function is reached. We formulate this principle in terms of some sequences of the set W elements and the sequences of the subsets of the same set W.

---

2 Kempner Y., Mirkin B. and I. Muchnik (1997) have given another example in Monotone Linkage Clustering and Quasi-Convex Set Functions, Appl. Math. Letters, v. 10, issue no. 4, pp. 19-24. Mirkin B. and I. Muchnik. (2002) Layered Clusters of Tightness Set Functions, Applied Mathematics Letters, v. 15, issue no. 2, pp. 147-151.

3 Yet another examples, Kuznetsov E.N. and I.B. Muchnik, Moscow (1982) Analysis of the Distribution Functions in an Organization, Automation and Remote Control, Plenum Publishing Corporation, pp. 1325-1332; Kuusik R. (1993) The Super-Fast Algorithm of Hierarchical Clustering and The Theory of Monotonic Systems, Data Processing, Problems of Programming, Transactions of Tallinn Technical University, No. 734, pp. 37-61; Mullat J.E., (1995) A Fast Algorithm for Finding Matching Responses in a Survey Data Table, Mathematical Social Sciences 30, pp. 195-205; Genkin A.V. and I. B. Muchnik (1993) Fixed Approach to Clustering, Journal of Classification, Springer, 10, pp. 219-240,.

Let $\overline{\alpha} = \{\alpha_0, \alpha_1, ..., \alpha_{k-1}\}$ is a sequence of elements of the set W and $k = |W|$. We define using the sequence $\overline{\alpha}$ a sequence of sets $\overline{H}(\overline{\alpha}) = \{H_0, H_1, ..., H_{k-1}\}$: as $H_0 = W$ and $H_{i+1} = H_i \setminus \{\alpha_i\}$.

Definition 1. We call a sequence of elements $\overline{\alpha}$ from the set W a defining sequence, if in the sequence of sets $\overline{H}(\overline{\alpha})$ there exists a sub sequence $\overline{\Gamma} = \{\Gamma_0, \Gamma_1, ..., \Gamma_p\}$ such that:

1°. The credential $\pi_{H_i}(\alpha_i)$ of an arbitrary element, belonging to Γ_j, but not belonging to Γ_{j+1}, is strictly less than $f(\Gamma_{j+1})$;

2° In Γ_p there do not exists such a strict subset L that $f(\Gamma_p) < F(L)$.

Definition 2. We call a subset H of the set W a definable, if there exists a defining sequence such that $H = \Gamma_p$.

Below, we simply refer to the notification $\{\pi_H\}$ as a credential system with respect to the set H.

Theorem. On the definable set H the function $f(H)$ reaches its global maximum. The definable set is unique. All sets, where the global maximum has been reached, lie within the definable set.

Proof. Let H is a definable set. Assume, that there exists L such that $f(H) \leq f(L)$. Suppose that $L \setminus H \neq \varnothing$, [4] otherwise we have just to proof the uniqueness of H, what we will accomplish below. Let H_t is the smallest from the sets H_i $(i = 0,1,...,k-1)$, which include in it the set $L \setminus H$. From this fact one can conclude, that there exists an element $\ell \in L$ such, that $\ell \in H_t$, but $\ell \notin H_{t+1}$. Moreover, in combination with $L \setminus H \neq \varnothing$ the last conclusion ensues $t < p$. Inequality $t < p$ disposes to an existence of at least one a subset in the sequence of sets $\overline{\Gamma}$ such, that

$$\pi_{H_t}(\ell) < f(\Gamma_j) \qquad (2)$$

and $j \geq t+1$. Since $\ell \notin H_{t+1}$ and $\Gamma_j \subseteq H_{t+1}$ are true, it follows that $\ell \notin \Gamma_j$. Thus, the inequality

$$f(\Gamma_j) \leq f(\Gamma_p) \qquad (3)$$

is valid as a consequence of the property 2° for the defining sequence.

---

[4] Here $\varnothing$ symbolizes an empty set.

Now, let $\ell \in \mathrm{L}$ and the credential $\pi_{\mathrm{L}}(\ell)$ is at the minimum in credential system with the respect to the set L. Inequalities (2) and (3) allow us to conclude, that $\pi_{\mathrm{H}_t}(\ell) < \pi_{\mathrm{L}}(\ell)$. Above we selected H_t on the condition that $\mathrm{L} \subset \mathrm{H}_t$. Hereby, recalling the main property p.2 of the credential system (the removal of elements), it is easily to establish that $\pi_{\mathrm{L}}(\ell) \le \pi_{\mathrm{H}_t}(\ell)$, i.e., in the credential system with the respect to the set L, there exists a credential, which is strictly less than the minimal. We came to a contradiction and by this, we have proved that on H the global maximum has been reached. Further, all such sets, different from H, where the global maximum is likewise reached, might really be located within H. It remains to be proved the uniqueness of the definable set. In connection of what we proved above, one might suppose that a definable set H' is located within H, however, proceeding with the line of reasoning towards H' similar to those we proposed above for L, we conclude, that $\mathrm{H} \subset \mathrm{H}'$. ■

Corollary. Let $\{\mathrm{R}\}$ is a system of sets, where the function of type (1) reaches its global maximum. Hereby, if $\mathrm{H}_1 \in \{\mathrm{R}\}$ and $\mathrm{H}_2 \in \{\mathrm{R}\}$ are valid, then $\mathrm{H}_1 \cup \mathrm{H}_2 \in \{\mathrm{R}\}$.

Proof. Following the p.2 (the main property) $f(\mathrm{H}_1) \le f(\mathrm{H}_1 \cup \mathrm{H}_2)$, but in addition $f(\mathrm{H}_1 \cup \mathrm{H}_2) \le f(\mathrm{H}_1)$, consequently $\mathrm{H}_1 \cup \mathrm{H}_2 \in \{\mathrm{R}\}$. ■

Below we introduce an actual algorithm for constructing the defining sequences of elements of a set W. For the availability of the algorithm is exposed in the form of a block-scheme similar to some extent of a computer program.

3. ALGORITHM [5]

a.1. Let the set $\mathrm{R} = \mathrm{W}$ and sequences $\overline{\alpha}$ and $\overline{\beta}$ [6] be empty sets in the beginning, and let the index $i = 0$.

a.2. Find an element μ at the least credential with the respect to the set R, record the value $\lambda = \pi_{\mathrm{R}}(\mu)$ and constitute $\overline{\alpha} = \overline{\alpha}, \overline{\beta}, \mu$ and thereafter $\overline{\beta} = \varnothing$.

---

[5] Further developments, see Muchnik, I., and Shvartser, L. (1990) Maximization of generalized characteristics of functions of monotone systems, Automation and Remote Control, 51, pp. 1562-1572,

[6] Hereby $\overline{\beta} = \{\beta_1, \beta_2, \ldots, \beta_i, \ldots\}$

a.3. Exclude the element μ from the set R and take into account the influence of the removed element $\mu \in R$ on remaining elements, i.e., recalculate all values $\pi_{R\setminus\mu}(\beta)$ for all $\beta \in R \setminus \{\mu\}$.

a.4. In case, among the remaining elements there exist such γ, that

$$\pi_{R\setminus\mu}(\gamma) \le \lambda \tag{4}$$

compose a sequence from those elements $\overline{\gamma} = \{\gamma_1, \gamma_2, \ldots, \gamma_s\}$ and substitute $\overline{\beta} = \overline{\beta}, \overline{\gamma}$.

a.5. Substitute the set $R = R \setminus \{\mu\}$ and the element $\mu = \beta_{i+1}$. Return to the a.3 in case the element β_{i+1} is the element for the sequence $\overline{\beta}$ increasing in this moment the index i by one.

a.6. In case, when the sequence $\overline{\alpha}$ has utilized the whole set W, the construction is finished. Otherwise, return to a.2 initializing first $i = 0$.

Let us prove that the sequence $\overline{\alpha}$ just constructed by the proposed algorithm is defining. We consider the sequence $\overline{H}(\overline{\alpha})$ and let one selects in the role of the sequence $\overline{\Gamma}$ those sets, which start by the element μ found at the moment the algorithm is crossing the step a.2. The fact of crossing the a.2 of the algorithm guarantees, that the condition (4) is not valid before the cross was occurred, and the element β_{i+1} is not in the sequence β at this stage. The above guarantees as well the condition 1° fulfillment for the defining sequences. Suppose, that the condition 2° in the definition 1 do not hold, i.e., in the last set Γ_p in the sequence $\overline{\Gamma}$, there exists such a subset L, that $f(\Gamma_p) < f(L)$. Let us consider the sequence $\overline{\beta}$, which is generated at the last crossing through the a.2 of the above-described algorithm and let λ symbolize the highest value among all such λ. One has to conclude, that $\lambda_p < f(\Gamma_p)$, and, from the supposition of an existence of a set L, we come to the inequality $\lambda_p < f(L)$. By the construction, the sequence $\overline{\alpha}$ and together with the sequence $\overline{\beta}$ (both of them), which is generated at last crossing though the a.2 of the algorithm has utilized all elements in W. Consequently, we can consider a set of elements K in the sequence $\overline{\beta}$, which start from the first confronted element $\ell \in L$, where $L \subset K$. On the basis justified above, we have $\pi_K(\ell) = \lambda_p$ and, recalling the main property of the credential system p.2 (the removal of elements), we conclude moreover that $\pi_L(\ell) \le \lambda_p$. We reached to a contradiction and by that we have proved the property 2° of the definition 1 for the sequence $\overline{\alpha}$. On that account, the construction of defining sequences is possible by the pointed above algorithm.

We emphasize the necessity of concretizing the notion of credential system with the respect to a subset of a given finite set for solving some of the pattern recognition problems, what should be the subject for further investigation.

In conclusion, we will point out, that the construction of defining sequences has been realized in practice on a computer for one problem in graph theory, related to an extraction of "almost totally connected" sub-graphs in a given graph. The number of edges in such graphs has been around 10^4.

LITERATURE

1. Cherenin, V.P., a) (1962) Solution of some Combinatorial Problems of optimal Scheduling by the Method of Successive Computations, "Proc. of Conf. on Experience and Prospective Applications of Mathematical Methods in Planning, (in Russian), *Isd. SO AN SSR*, Novosibirsk, pp. 111-113;
b) (1962) Solution of some Combinatorial Problems of optimal Scheduling by the Method of Successive Computations," (in Russian), *Scien.-Method Proc. of econ.-math. seminar*, publ. 2, LEMM and VC AN SSR, M.

2. Cherenin V.P. and B.R. Khachaturov, a) (1965) The Solution by the Method of Successive Computations of some Problems in Plant Locations, (in Russian), Implementation of Math. Methods and EVM in Econ. Investigations, *Nauka*, Uzb. SSR, Tashkent; b) (1965) The Solution by the Method of Successive Computations of some Problems in Plant Locations, (in Russian), "*Econ.-Math. methods*", publ.2, Isd. "Nauka", M.

3. Mirkin, B.G. (1970) The Classification Problem for Qualitative Data, (in Russian), *Math. Questions of Econ. Models Formation*, Novosibirsk.

NB! In his work "Cores of Convex Games" Shapley investigated a class of n-person's games with special convex (supermodular) property, International Journal of Game Theory, Vol. 1, 1971, pp. 11-26. When writing current paper, in that time in the past, the author was not familiar with this work and could not predict the close connection between the basic monotonicity property pp.1-2, see above, and that of supermodular characteristics functions in convex games induce the same property upon marginal utilities. We are going to explain the connection. We will consequently do it in Shapley's own words to make the idea crystal clear.

The core of a n-person game is the set of feasible outcomes that cannot be improved upon by any coalition of players. A convex game is one that is based on a convex set function; intuitively this means that the incentives for joining a coalition increase as the coalition grows, so that one might expect a "snowballing" or "band-wagon" effect when the game is played cooperatively… In Shapley's paper a coalition game is a function v mapping a Ring of subsets from some set called a grand coalition $\mathcal{N}$ to the real numbers, satisfying $\mathrm{v}(\varnothing)=0$. *The function* v *is supperadditive if*

$\mathrm{v(S)+v(T)\leq v(S\cup T)}$, i.e., all $\mathrm{S,T}\in\mathcal{N}$, with $\mathrm{S\cap T}=\varnothing$.

It is convex if $\mathrm{v(S)+v(T)\leq v(S\cup T)+v(S\cap T)}$

for all $\mathrm{S,T}\in\mathcal{N}$, p.12.

In the standard form in game theory, the elements of $\mathcal{N}$ are "players", the subsets of $\mathcal{N}$ are "coalitions"; $\mathrm{v(S)}$ is called the "characteristic function", which gives each coalition the best payoff that it can get without the help of other players.

Supper-additivity arises naturally in this interpretation, but convexity is another matter. For example, in voting situation S and T, but not $\mathrm{S\cap T}$, might be winning coalitions, causing "convexity" to fail. To see what convexity does entail, consider the function m:

$$\mathrm{m(S,T)=v(S\cup T)-v(S)-v(T)},$$

as defining the "incentive to merge" between disjoint coalitions S and T. Then it is a simple exercise to verify that convexity is equivalent to the assertion that $\mathrm{m(S,T)}$ is no decreasing in each variable – whence the "snowballing" or "band wagon" effect mentioned in the introduction.

Another condition that is equivalent to convexity (provided $\mathcal{N}$ is finite) is to require that

$$\mathrm{v(S\cup\{i\})-v(S)\leq v(T\cup\{i\})-v(T)}$$

for all individuals $\mathrm{i}\in\mathcal{N}$ and all $\mathrm{S\subseteq T\subseteq}\mathcal{N}\setminus\{\mathrm{i}\}$. This expresses a sort of increasing marginal utility for coalition membership, and is analogous to *"increasing the returns to scale associated with convex production functions in economics.",* p.13

We return now back from the "expedition" into Shapley's work and make some comments. The latter condition, which is equivalent to convexity, is an exact, we repeat it once again, an exact utilization of our basic monotonicity property pp.1-2. Set functions of this type are also known in the literature as "suppermodular". As it turns out now the author knew such functions. To the knowledge of the author Cherenin was first who introduced functions of this type already in 1948. Nemhauser et al, also used $\mathrm{v(S)+v(T)\geq v(S\cup T)+v(S\cap T)}$ but an inverse property introduced in 1978 for computational optimization problems in "An Analysis of Approximation for Maximizing Submodular Set Functions", Mathematical Programming 14, 1978, 265-294. Shapley also notes the latter inverse property in connection with rank function of a matroid known as "submodular" or "lower semi-modular." Besides, in Nemhauser et al paper, the reader may find the proof of the conditions:

$$v(S) + v(T) \leq v(S \cup T) + v(S \cap T) \text{ and}$$
$$v(S \cup \{i\}) - v(S) \leq v(T \cup \{i\}) - v(T) \text{ equivalency.}$$

However, the connection between the convex games and the monotonicity property pp.1-2 is invisible. Only recently Genkin and Muchnik pointed out (not in the connection with game theoretical models, but actually in connection with the problems of object classification, see "Submodular Set Functions and Monotone Systems in Aggregation Problems I,II," Translated from Automat. Telemekhanika No.5, pp.135-148, © 1987 0005-1179/87/4805-0679, Plenum Publishing Corporation), that the functions family $\pi_H(\alpha) = v(H) - v(H \setminus \{\alpha\})$ represent a derivatives of supper-modular set functions in the form just exhibited in Shapley's work.

SUMMARIZING

In convex games, following the theory developed in this work from 1971, one can always find a coalition, where it members will be awarded individually at least by some maximum payoff of guaranteed marginal utility, see the Theorem. We call this coalition the largest kernel (nuclei) or the definable set. A good example and its like, is the Example 1. Here, in economic terms, the marginal utility highlights the number of direct dealers with the player $i \in S$ (number of direct contacts, buyers, sellers, direct suppliers, etc.). On the contrary, the Example 2 is not its like and goes beyond the Shapley's Convex Game idea.