BOOK REVIEWS

BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 29, Number 1, July 1993

0273-0979/93 $1.00 + $.25 per page

Submodular functions and optimization theory, by S. Fujishige. Annals of Discrete Mathematics, no. 47, North Holland, Amsterdam, 270 pp., 1991, $97.25. ISBN 0-444-88556-0

Let $N = \{1, \dots, n\}$ be a finite set, and let $\mathscr{D}$ denote a collection of subsets of N, containing N and $\varnothing$, which is closed under unions and intersections. A set function $f: \mathscr{D} \to \Re$ with $f(\varnothing) = 0$ is submodular if

$$f(S) + f(T) \geq f(S \cup T) + f(S \cap T)$$

for every S, T and element of $\mathscr{D}$. A function f is supermodular if $-f$ is submodular. Let $\Gamma(\mathscr{D})$ denote the set of all submodular functions on $\mathscr{D}$. A submodular system is a pair $(\mathscr{D}, f)$ where f is an element of $\Gamma(\mathscr{D})$. Submodular functions and systems have numerous applications in combinatorial theory (where they arise as rank functions of matroids), game theory (in which a supermodular function defines a convex game), and optimization theory (especially network problems). Not surprisingly, one can study submodularity from many directions; the general theme of Santoru Fujishige's book is, as the title indicates, the interplay of submodularity and optimization, which the author accomplishes by presenting many aspects of the polyhedral theory for submodular functions.

To begin our review, we present some notation and definitions. In the discussion that follows, we will not present the author's most general results for submodular systems but will instead assume throughout that $\mathscr{D} = 2^N$ for expositional simplicity. For any set function f, two associated polyhedra with interesting structural properties are defined by

$$P(f) = \{x \in \Re^n \mid x(S) \leq f(S), \text{ for all } S \subseteq N\}$$

and

$$B(f) = \{x \in P(f) \mid x(N) = f(N)\}$$

where $x(S) = \sum_{i \in S} x_i$. If f is submodular, then $P(f)$ is called the submodular polyhedron for f, and $B(f)$ is called the base polyhedron. If f is submodular and monotone increasing (i.e., $f(T) \geq f(S)$ if $T \supseteq S$), then f is the rank function of the polymatroid (N, f).

$B(f)$ is related to two other set functions f_p and $\hat{f}$ on 2^N defined respectively as

$$f_p(S) = \min \sum_{k=1}^{r} f(T_k) \quad \text{s.t. } \{T_1, \dots, T_r\} \text{ is a partition of } S$$

and

$$\hat{f}(S) = \max_{x \in P(f)} x(S).$$

f_p is called the Dilworth truncation of f. For any set function f, $B(f) \neq \varnothing$ if and only if $\hat{f}(N) = f(N)$. Many of the results and applications of submodular functions also rely on properties of the dual function $f^{\#}$ which is defined by

$$f^{\#}(S) = f(N) - f(N \backslash S) \quad \text{for all } S \subseteq N.$$

It is clear that f is submodular if and only if $f^{\#}$ is supermodular and, for any f, $B(f) = -B(-f^{\#})$.

Chapter I contains some notation and definitions from set theory, graph theory, algebra, and convex analysis that are used throughout the book. Chapter II begins with a presentation of the well-known relationship between matroids and submodularity and introduces the various polyhedra associated with a submodular function f. The main result in this chapter is that $B(f)$ is nonempty if f is submodular, but this need not be the case for general set functions f. There are many generalizations of submodularity that also guarantee that $B(f) \neq \varnothing$, and we will discuss some of these below. Chapter II continues with a discussion of the extreme point structure of $B(f)$ and the greedy algorithm for linear programs, whose feasible set is $B(f)$ for some submodular function f. The central idea here, due to Edmonds (1970) and Shapley (1971), is that the extreme points of the base polyhedron for a submodular function f are precisely the "marginal worth vectors" for f. Let R be one of the $n!$ orderings of N. R induces a binary precedence relation on the members of N. For each $j \in N$ define the set $X_j(R) = \{i \in N \mid i \text{ precedes } j \text{ in } R\}$. (If j is the first element in the ordering R, then $X_j(R) = \varnothing$.) Then the marginal worth vector for R is defined to be the vector $\hat{x}(R) \in \mathfrak{R}^n$ whose jth coordinate is

$$\hat{x}_j(R) = f(X_j(R) \cup \{j\}) - f(X_j(R)).$$

The classic result states that, for submodular f, $B(f) = W(f)$ where $W(f)$ is the convex hull of the set $\{\hat{x}(R) \mid R \text{ is an ordering of } N\}$. As a result, a submodular function can be characterized in terms of the success of the so-called "greedy algorithm". This result can also be established using results not discussed in the book but which are of independent interest. Weber (1988) has shown that, for *any* set function f, $B(f) \subseteq W(f)$. Ichiishi (1981) has shown that if $W(f) \subseteq B(f)$, then f is submodular. Hence, submodularity is completely characterized by the equality of $B(f)$ and $W(f)$.

In many combinatorial optimization problems submodular (or supermodular) functions arise in a natural way. For example, Curiel, Pederzoli, and Tijs (1988) consider a scheduling problem in which n customers must wait in a queue for service by a single server. Each customer has a deterministic service time s_i and a linear cost function $c_i(t_i) = c_i t_i$, when waiting plus service time is equal to t_i. If the customers are initially arranged in order R, the total cost is given by $\sum_{i=1}^{n}[c_i(s_i + \sum_{j \in X_i(R)} s_j)]$. By a well-known result in scheduling theory, the total cost can be minimized if customers are rearranged in decreasing order of their "urgency" index defined by c_i/s_i. Curiel et al. demonstrate that a function f, which defines the cost savings achievable by any set of customers, is supermodular. Another example is the class of simultaneous network synthesis optimization problems, which arise from the minimization of the total cost of establishing communications links among nodes in a network. Let r_{ij} represent the flow requirements between nodes i and j, and let d_{ij} represent the cost per unit of traffic carried on a link between i and j. Granot and Hojati (1990) demonstrate that the function f, which defines the minimum cost of satisfying the flow requirements for every subset of customers, is submodular.

Many of the results of Chapter II are generalizations of the basic theory of polymatroids, and the author focuses on generalizations to arbitrary submodular systems. One could also generalize the notion of submodular set function; a number of such extensions have appeared in the literature. Qi (1988) defines a set function f to be "odd submodular" if there exists a partition $\mathcal{J}_1$ and $\mathcal{J}_2$ of 2^N such that for each $p = 1, 2$ and any two intersecting sets $S, T \in \mathcal{J}_p$ (i.e., $S\backslash T \neq \varnothing$, $T\backslash S \neq \varnothing$, and $S \cap T \neq \varnothing$), $f(S)+f(T) \geq f(S \cup T)+f(S \cap T)$. Every submodular function is odd submodular. Furthermore, if f is odd submodular, then $f(S) = f_p(S)$ for every $S \subseteq N$ because (as Qi shows) the polyhedron $P(f)$ is totally dual integral. Since $f_p = f$ for any subadditive set function, it follows that $B(f) \neq \varnothing$ if f is a subadditive odd submodular function.

Another generalization of submodularity is permutational submodularity, introduced in Granot and Huberman (1982), where it is called permutational convexity. A function f is permutationally submodular if there exists an ordering R such that for each pair j, k and $j \in X_k(R)$ and each $S \subseteq N\backslash X_k(R)$

$$f(X_k(R) \cup S) - f(X_k(R)) \leq f(X_j(R) \cup S) - f(X_j(R)).$$

Every submodular function is permutationally submodular, and Granot and Huberman show that the marginal worth vector $\hat{x}(R)$ for an ordering R satisfying the above conditions is an element of $B(f)$ so that $B(f) \neq \varnothing$.

Permutationally submodular set functions arise from an interesting class of combinatorial optimization problems that are important in the theory of games. Consider the problem of connecting n nodes to some "source" node labelled 0. For each $(i, j) \in (N \cup \{0\}) \times N$, c_{ij} is the cost of connecting nodes i and j. For each $S \subseteq N$, let $f(S)$ be the cost of the minimum cost spanning tree for the node set $S \cup \{0\}$. The resulting set function f is permutationally submodular, and an element of $B(f)$ can be constructed in the following manner: let R be any ordering with the property that $j \in X_i(R)$ if j lies on the unique path connecting i to 0 in the minimum cost spanning tree. Then the marginal worth vector for R is the desired member of $B(f)$.

Chapter III is devoted to neoflows, which are generalizations of classical flow problems. The chapter begins with an algorithm for the intersection problem

$$\max \sum_{i \in N} x_i \quad \text{s.t. } x \in P(f_1) \cap P(f_2)$$

where f_1 and f_2 are submodular. The author then provides a generalization of the polymatroid intersection theorem of Edmonds and uses this result to prove a discrete separation theorem of Frank (1982). A problem related to the intersection theorem is the common base problem—i.e., when is $B(f_1) \cap B(f_2) \neq \varnothing$? An affirmative answer can be given when f_1 majorizes the dual of f_2. In particular, let f_1 and f_2 be submodular with $f_1(N) = f_2(N)$. Then $B(f_1) \cap B(f_2) \neq \varnothing$ if and only if $f_2^{\#}(S) \leq f_1(S)$ for all $S \subseteq N$. The remainder of Chapter III is primarily devoted to an analysis of submodular flow problems first presented in Edmonds and Giles (1977), with special emphasis on the algorithmic aspects.

Chapter IV, entitled "Submodular Analysis", develops a Fenchel-type duality theory for submodular functions and applies the theory to constrained and unconstrained optimization problems involving submodular functions. Briefly, let f be submodular and define the conjugate function $f^*: \mathfrak{R}^n \to \mathfrak{R}$ as $f^*(x) = \max_{S \in 2^N}\{x(S) - f(S)\}$. Note that f^* is convex and if f submodular, then $f(S) = \max_{x \in \mathfrak{R}^n}\{x(S) - f^*(x)\}$. Furthermore, a subgradient theory can be developed in a natural way. A vector $u \in \mathfrak{R}^n$ is a subgradient of f at $S \in 2^N$ if $f(S) + u(T) - u(S) \le f(T)$ for each $T \in 2^N$, and the set of all subgradients at S, the subdifferential at S, is denoted $\partial f(S)$. The subgradient theory for submodular functions is analogous to that for convex functions as developed by Rockafellar (1970), and there are numerous interesting parallel results. For example, if f is submodular and $S \in 2^N$, then $u \in \partial f(S)$ if and only if $S \in \partial_2 f^*(u)$ where $\partial_2 f^*(u) = \{T \in 2^N \mid x(T) - u(T) \le f^*(x) - f^*(u)$ for all $x \in \mathfrak{R}^n\}$.

As noted above, many of the results from submodular analysis can be applied to optimization problems involving submodular functions. For the simplest unconstrained problem with submodular objective function $\min f(S)$ s.t. $S \in 2^N$, it follows that T is a solution if and only if $0 \in \partial f(T)$. Fujishige has generalized this result to obtain a Kuhn-Tucker theorem for submodular optimization. An alternative definition of subgradient is found in Qi (1988).

Chapter V concerns optimization problems whose feasible sets are the polyhedra associated with a submodular function f but whose objective functions are nonlinear. In particular, the author considers several variants of the minimization problem

$$\min G(x_1, \dots, x_n) \quad \text{s.t. } x \in B(f).$$

The first case to be considered is that of separable convex minimization for which $G(x_1, \dots, x_n) = \sum_{i=1}^n g_i(x_i)$, where each $g_i: \mathfrak{R} \to \mathfrak{R}$ is convex. For this problem, a "greedy algorithm" characterization of an optimal solution is provided that generalizes the aforementioned result when each g_i is linear. In addition, a decomposition procedure [Groenevelt, (1985)] for computing a solution is provided. Several other objective functions are treated, such as the "max-min" problem in which $G(x_1, \dots, x_n) = \min_{i \in N}\{g_i(x_i)\}$ and the so-called "fair resource allocation" problem whose objective function is $G(x_1, \dots, x_n) = \varphi(\max_{i \in N}\{g_i(x_i)\}, \min_{i \in N}\{g_i(x_i)\})$, where φ is increasing in its first argument and decreasing in its second argument. Furthermore, an analysis of lexicographic optimization on $B(f)$ is given. For each of these problems the author also discusses the version in which integer constraints on the x_i are imposed.

An important recent application of this methodology, not mentioned in the book, concerns the optimization of performance measures for queueing systems. In particular, consider a queueing system with n customer classes. Let $\rho_i = \lambda_i/\mu_i$ be the utilization rate for class i where λ_i is the mean arrival rate of the Poisson process generating arrivals of type i and $1/\mu_i$ is the expected value of the exponentially distributed service time for class i. Finally, let W_i be the expected waiting time for a customer of class i. The performance vector $W = (W_1, \dots, W_n)$ is a function of the priority discipline. Federgruen and Groenevelt (1988) have demonstrated that for certain queueing systems the set of performance vectors obtainable from some preemptive work-conserving discipline has the mathematical structure of the base polyhedron of a submodular function.

More precisely, W is attainable using a work-conserving preemptive discipline if, for every $S \subseteq N$, $\sum_{i \in S} \rho_i W_i \geq f(S)$ and $\sum_{i \in N} \rho_i W_i = f(N)$, where $f(S)$ represents the average work in the system when only the customer classes in S are served. Hence, the attainable performance vectors correspond to the base vectors in the polyhedron $B(f^{\#})$. That is, if $(x_1, \ldots, x_n) \in B(f^{\#})$, then $(W_1, \ldots, W_n) = (x_1/\rho_1, \ldots, x_n/\rho_n)$ is an attainable performance vector. With this characterization, one can choose priority disciplines so as to optimize some function of the associated performance vector. If f is supermodular (so that $f^{\#}$ is submodular), then the techniques of Chapter V are applicable.

This book is a well-written and thorough treatment of the relationships of submodularity to optimization theory that have been studied in the last decade. It will be a very useful reference for researchers and graduate students working in discrete mathematics, operations research, and game theory. One minor caveat should be mentioned. A number of theorems have maintained hypotheses that are stated only at the beginning of the sections in which the theorem is found. As a result, a reader using the book as a reference source should be careful to identify the context in which results are presented.

We close with a few observations regarding the (nonempty) intersection of game theoretic concepts and ideas that are important in submodularity theory. The important (game theoretically motivated) paper of Shapley (1971) is cited throughout the submodularity literature, but many more concepts have been independently studied by researchers in both areas using techniques that are often similar. In game theory, a set function f with $f(\varnothing) = 0$ is called a "worth function" and a pair (N, f) is a cooperative game in coalitional form. f is superadditive if $f(S) + f(T) \leq f(S \cup T)$ whenever $S \cap T = \varnothing$. The "core" of f is the set

$$C(f) = \{x \in \Re^n \mid x(S) \geq f(S) \text{ and } x(N) = f(N)\}.$$

(In the game theory literature, the polyhedron $B(f)$ as defined in this review would be referred to as the "anticore".) One of the fundamental problems in game theory is to identify classes of games for which the core is nonempty. A supermodular set function defines a "convex game", since such a set function exhibits "increasing returns" to coalition size. It follows from the results described above that the core of a convex game is nonempty. The function f^{SA} defined by

$$f^{\mathrm{SA}}(S) = \max \sum_{k=1}^{m} f(T_k), \quad \text{where } \{T_1, \ldots, T_m\} \text{ is a partition of } S,$$

is called the "superadditive cover" for f and is an obvious analogue of the Dilworth truncation.

The function f^{b} defined by

$$f^{\mathrm{b}}(S) = \max \sum_{T \subseteq S} \delta_T f(T) \text{ s.t. } \delta_T \geq 0 \quad \text{and} \quad \sum_{S \subseteq N : i \in S} \delta_s = 1 \text{ for each } i \in N$$

is called the "balanced cover" of f. A game is "balanced" if $f^{\mathrm{b}}(N) = f(N)$ and "totally balanced" if $f^{\mathrm{b}}(S) = f(S)$ for every S. A fundamental result in cooperative game theory, due to Bondareva (1962) and Shapley (1967), states that a game has a nonempty core if and only if it is balanced and every "subgame" has a nonempty core if and only if the game is totally balanced. Analogous results for base polyhedra have been derived by Qi (1988). Qi calls a function f "discrete convex" if $f(S) = \hat{f}(S)$ for every $S \subseteq N$ and "generally subadditive" if $-f$ is totally balanced. As a consequence of LP duality, he proves that f is discrete convex if and only if it is generally subadditive. This is a restatement of the Bondareva-Shapley theorem.

In the game theory literature there are several classes of functions more general than supermodular functions that are contained in the class of totally balanced set functions. For example, Schmeidler (1972) defines the function f to be exact if $f(S) = \tilde{f}(S)$ for every $S \subseteq N$ where $\tilde{f}(S) = \min_{x \in C(f)} x(S)$. It is clear that f is exact if and only if $f^{\#}(S) = \max_{x \in B(f^{\#})} x(S)$ for every S. If f is exact and superadditive, it follows that both f and $-f^{\#}$ are totally balanced, and Shapley (1971) has demonstrated that a supermodular function is exact. Other interesting classes of functions related to supermodular functions have been studied by Sharkey (1982) and Ichiishi (1990). These results, which were derived independently of the literature on submodular optimization, could benefit from a reexamination in terms of the results that Fuijishige presents. At the same time, researchers who are active in optimization theory might benefit from a thorough study of related results in the game theory literature.

References

O. Bondareva (1962), *The core of an N-person game*, Vestnik Leningrad Univ. 17, 141–142.

I. Curiel, G. Pederzoli, and S. Tijs (1989), *Sequencing games*, European J. Operational Res. **40**, 344–351.

J. Edmonds (1970), *Submodular functions, matroids, and certain polyhedra*, Proceedings of the Calgary International Conference on Combinatorial Structures and Their Applications (R. Guy, H. Hanani, N. Sauer, and J. Schönheim, eds.), Gordon and Breach, New York.

J. Edmonds and R. Giles (1977), *A min-max relation for submodular functions on graphs*, Ann. of Discrete Math. vol. 1, North-Holland, Amsterdam, 185–204.

A. Federgruen and H. Groenevelt (1988), *Characterization and optimization of achievable performance in general queueing systems*, Operations Res. **36**, 733–741.

A. Frank (1982), *An algorithm for submodular functions on graphs*, (A. Bachem, M. Grötschel, and B. Korte, eds.), Ann. of Discrete Math., vol. 16, North-Holland, Amsterdam, pp. 189–212.

D. Granot and M. Hojati (1990), *On cost allocation in communication networks*, Networks **20**, 209–229.

D. Granot and G. Huberman (1981), *Minimum cost spanning tree games*, Math. Programming **21**, 1–18.

——— (1982), *The relationship between convex games and minimal cost spanning tree games: a case for permutationally convex games*, SIAM J. Algebra Discrete Math. **3**, 288–292.

H. Groenevelt (1985), *Two algorithms for maximizing a separable concave function over a polymatroid feasible region*, working paper, The Graduate School of Management, Univ. of Rochester.

T. Ichiishi (1981), *Supermodularity: applications to convex games and to the greedy algorithm for LP*, J. Economic Theory **25**, 283–286.

——— (1990), *Comparative cooperative game theory*, Internat. J. Game Theory **19**, 139–152.

L. Qi (1988), *Odd submodular functions, Dilworth functions and discrete convex functions*, Math. Oper. Res. **13**, 435–446.

R. T. Rockafellar (1970), *Convex analysis*, Princeton Univ. Press, Princeton, NJ.

D. Schmeidler (1972), *Cores of exact games*, J. Math. Anal. Appl. **40**, 214–225.

L. S. Shapley (1967), *On balanced sets and cores*, Naval Res. Logistics Quart. **14**, 453–460.

W. W. Sharkey (1982), *Cooperative games with large cores*, Internat. J. Game Theory **11**, 175–182.

R. Weber (1988), *Probabilistic values for games*, The Shapley Value (A. Roth, ed.), Cambridge Univ. Press, Cambridge.

RICHARD P. MCLEAN AND WILLIAM W. SHARKEY
RUTGERS UNIVERSITY AND BELL COMMUNICATIONS RESEARCH
E-mail address: mclean@zodiac.rutgers.edu
E-mail address: wws@bellcore.com

BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 29, Number 1, July 1993

0273-0979/93 $1.00 + $.25 per page

Finite soluble groups, by Klaus Doerk and Trevor Hawkes. De Gruyter Expositions in Mathematics, vol. 4, de Gruyter, Berlin and New York, 1992, x + 891 pp., $148.00. ISBN 3-11-012892-6

This is an account of the development of the theory of finite soluble groups over the last thirty years, concentrating on "those parts of the subject where a coherent and unified body of knowledge has emerged": the theory of Schunck classes, formations, and Fitting classes of soluble groups with their associated subgroups. My first impression of the book was of its size. It contains almost 900 pages, including indices and appendices. It is written carefully and clearly and will become, as the authors hope, "a basic reference [for soluble groups], a text for postgraduate teaching, and ... a source of research ideas and techniques". The book is about *finite* groups, but of course infinite groups turn up at times, for example, when working with groups associated with infinite fields or with Fitting classes. Again, the book is about *soluble* groups, but the authors often explore "insoluble territory to see where the soluble theory leads".

A group G is said to be *soluble* if it has a series of subgroups $G = G_0 > G_1 > \cdots > G_n = \{1\}$ such that G_i is a normal subgroup of G and G_{i-1}/G_i is abelian for all i.

The two fundamental results on which the theory of finite soluble groups has been built are the theorems of Sylow and Hall. Sylow's theorems state that, for a prime p and a finite group G of order $p^a m$ with p not dividing m, G has a unique conjugacy class of subgroups of order p^a, now called *Sylow p-subgroups* of G, and any p-subgroup of G is contained in a Sylow p-subgroup. In 1928 Hall [6] published a far-reaching generalisation of Sylow's theorems for soluble groups. Let π be a set of prime numbers, and let G be a finite group. The order $|G|$ can be factorised as nm where n is a π-number (that is, all its prime factors lie in π) and m is a π'-number (none of its prime factors lie in π). A subgroup whose order divides n is known as a π-subgroup, and a subgroup of order n is now known as a *Hall π-subgroup.* What Hall proved was that if G is soluble, then G has Hall π-subgroups, they form a single conjugacy class, and any π-subgroup is contained in a Hall π-subgroup. In 1937, nine years later, Hall went on to show that this fact actually characterises finite soluble