The Coming of the Matroids

William H. Cunningham

2010 Mathematics Subject Classification: 90C27, 05B35
Keywords and Phrases: Matroid, submodular function, polymatroid

Introduction

In 1964, thirty years after their introduction, and having lived a quiet life until then, matroids began to get the attention of optimizers. Just a few years later, as a result of exciting research achievements as well as enthusiastic promotion, the theory of matroids and submodular functions had become an integral part of discrete optimization.

Whitney

Matroid theory starts with the paper [22] of Hassler Whitney in 1935. A *matroid* may be defined to be a family of "independent" subsets of a finite ground set S, satisfying

- Every subset of an independent set is independent
- For any $A \subseteq S$ all maximal independent subsets of A (called *bases* of A) have the same cardinality (called the *rank* $r(A)$ of A).

Of course, if we take S to be the set of columns of a matrix, and the independent sets to be the ones that are linearly independent, we get a first example, called a *linear* matroid. Another important class consists of the *graphic* ones – here S is the set of edges of a graph G and a subset is independent if it forms a forest.

Whitney established some equivalent versions of the axioms, highlighted the above two examples, and proved several basic results. In particular, he showed that, given a matroid M, one gets a second *dual* matroid M^* by declaring independent all the sets whose deletion from S do not lower its rank. This generalizes the notion of duality in planar graphs. In addition, he observed that the rank function r satisfies what we now call the *submodular* property: For all subsets A, B of S

$$r(A) + r(B) \geq r(A \cup B) + r(A \cap B).$$

There were other researchers who suggested ideas similar to Whitney's. None of these early papers appears to contain any suggestion of a connection with optimization. In retrospect, one might observe that the definition implies that a certain trivial algorithm solves the optimization problem of finding a largest independent set.

Rado

In the next twenty years, there was little in the way of followup work to Whitney's paper. One exception, not widely appreciated at the time, was a paper [14] of Richard Rado in 1942. Rado gave a matroid generalization of Hall's theorem on matching. This famous theorem says that if G is a bipartite graph with parts S, T, then T can be matched into S if and only if for every subset A of T, $|N(A)| \geq |A|$. (Here $N(A)$ denotes the neighbourset of A.) Rado's "Independent Transversal Theorem" is perhaps the first significant result in matroid theory.

Theorem 1. *Let G be a bipartite graph with parts S, T, and let M be a matroid on S. Then T can be matched to an* independent *set of M, if and and only if, for every subset A of T, $r(N(A)) \geq |A|$.*

Tutte

In the late fifties Bill Tutte published several deep results on matroid theory [18], [19]. Tutte's background is interesting. A chemistry student at the beginning of the war, he was recruited to the Bletchley Park codebreaking project. His brilliant contributions to that effort were kept secret for more than fifty years. See Copeland [1] for details. At the end of the war Tutte returned to Cambridge as a mathematician, and a Fellow of Trinity College; the fellowship was a partial reward for his war work. In his thesis he studied "nets", a generalizations of graphs, which he has described [21] as being "half-way to matroids". He eventually published much of this work in the setting of matroid theory.

Tutte solved several of the fundamental problems suggested by the work of Whitney. These included characterizing the matroids that are graphic, those that arise from matrices over the binary field, and those that are *regular* (that is, arise from matrices over *every* field). These basic results are already of importance to optimizers. Understanding the graphic matroids, is the key to understanding which linear programming problems are reducible, by row operations and variable-scaling, to network flow problems. Moreover, as Tutte showed, the regular matroids are precisely the ones realizable by totally unimodular matrices, which Tutte characterized. However, Tutte's matroid papers were difficult and their connections with optimization were not immediately recognized.

The Sixties

It was in the 1960's that matroids and submodularity became an important subject in optimization. The dominant figure of the period was Jack Edmonds. Not only did he discover beautiful theorems and algorithms. He also championed his subject tirelessly, defining a vocabulary that is still in use, and an agenda – efficient algorithms and polyhedral descriptions – that is still being followed. By 1969 Edmonds and his work had been featured at a major international conference, and he had written for its proceedings the milestone paper [2].

Edmonds, Lehman, and Matroid Partition

Like Tutte, Jack Edmonds had an interesting background; see his own lively account in [3]. After his undergraduate years, which included study at two universities and a year out of school, he enrolled in the graduate program in mathematics at the University of Maryland. He completed a master's thesis, in which he proved a fundamental result in topological graph theory, but left Maryland before completing the doctoral program. He was fortunate to obtain a position in the Applied Mathematics Division of the National Bureau of Standards in Washington. Here, in an operations research group headed by Alan Goldman, he was exposed to problems in combinatorial optimization. Edmonds has written "That is where graduate school started for me, with Alan Goldman".

In 1961, while participating in a workshop at the Rand Corporation, he discovered the key idea that led to his solution of the matching problem. Over the next couple of years, he worked out algorithms and polyhedral descriptions for matching and degree-constrained subgraphs (for more on this, see Pulleyblank [13]). Since Tutte had proved the basic existence theorem in matching theory, Edmonds was certainly aware of his work. However, he credits Alfred Lehman for inspiring him to consider matroids as a natural setting for posing and attacking algorithmic problems. The two met in spring 1964, shortly after Lehman solved the Shannon switching game, a game played on a graph. In fact, Lehman [10] had invented and solved a more general game, played on a matroid. His solution did not however, provide efficient algorithms to decide which player had the winning strategy.

For one variant of Lehman's game, the condition for a certain player to have a winning strategy is that the ground set have two disjoint bases. Edmonds characterized this property, and more generally solved the problem of finding in a matroid M a largest set that is the union of k independent sets, at the same time providing an algorithm. The algorithm is efficient, assuming that there is an efficient algorithm to recognize independence in M. This and related results completed the solution of Lehman's game. Then with Ray Fulkerson, Edmonds solved a yet more general problem, as follows. Suppose that we are given matroids $M_1, \dots, M_k$ on S. Call a set I *partitionable* if it can be expressed as the union of k sets I_i, where I_i is independent in M_i for each i.

Theorem 2 (Matroid Partition Theorem). *The maximum size of a set I partitionable with respect to $M_1, \dots, M_k$ is equal to the minimum, over subsets A of S, of*

$$|S \backslash A| + \sum_{i=1}^{k} r_i(A).$$

Here r_i denotes the rank function of M_i. Their proof is an efficient algorithm to find the optimal I and A. It is easy to obtain from the Matroid Partition Theorem a formula for the maximum number of disjoint bases of a given matroid, and for the minimum number of independent sets that cover S. In fact, the technique provides many applications to packing and covering.

The First Conference

Jack Edmonds organized the first conference on matroids. It was called a "Seminar on Matroids" and was held at NBS August 31 to September 11, 1964. He has written [4] that, when organizing the meeting, he "could not find more than six people who had heard the term" matroid. But there, according to Tutte [21], "the theory of matroids was proclaimed to the world". Edmonds arranged for Tutte to give a series of lectures on his work, and to write for publication a new exposition [20] of his main structural results. Edmonds presented his own work related to partitioning and Lehman's game. Participants included Ray Fulkerson and Gian-Carlo Rota; the latter campaigned to change the term "matroid" to "combinatorial geometry". Tutte and Edmonds were not convinced, and the movement was ultimately not successful, but there was a period in the seventies when it seemed the new term might be winning out. One paper [9] suggested that was the case, and tut-tutted that the term "matroid" was "still

Figure 1: The Seminar on Matroids, NBS, 1964. First row, second from left, Ray Fulkerson, third from left, Bill Tutte. (Photo courtesy of William Pulleyblank)

Figure 2: The Seminar on Matroids, NBS, 1964. First row, right, Jack Edmonds, third from right, Gian-Carlo Rota. (Photo courtesy of William Pulleyblank)

used in pockets of the tradition-bound British Commonwealth". (At that time both Tutte and Edmonds were in Waterloo.)

Matroid Intersection

There are several theorems essentially equivalent to the Matroid Partition Theorem, and they are important in their own right. These equivalent statements serve to emphasize the power of the theorem and algorithm. However, almost inevitably there have been independent discovery and rediscovery of results. In fact Rado's Theorem 1 is one of these. Another of the equivalent theorems is known as Tutte's Linking Theorem; see [12]. Tutte called it Menger's Theorem for Matroids. But for optimizers, undoubtedly the most important of these versions is Edmonds' Matroid Intersection Theorem, which he discovered by applying the Matroid Partition Theorem to M_1 and the dual of M_2.

Theorem 3 (Matroid Intersection Theorem). *Let M_1, M_2 be matroids on S. The maximum size of a common independent set is equal to the minimum over subsets A of S of*

$$r_1(A) + r_2(S\backslash A).$$

This theorem generalizes the famous Kőnig min-max theorem for the maximum size of a matching in a bipartite graph. Since the more general weighted version of that problem (essentially, the optimal assignment problem) was well known to be solvable, Theorem 3 cries out for a weighted generalization. So, given two matroids on S and a weight vector $c \in \mathrm{R}^S$, can we find a common independent set of maximum weight? Or, can we describe the convex hull of common independent sets? First, let's back up and deal with a single matroid.

The Matroid Polytope

By 1964 Jack Edmonds had already solved the weighted matching problem, in the process, proving the matching polyhedron theorem. The fact that a greedy algorithm finds an optimal spanning tree of a graph was well known. Its proof did not require polyhedral methods, but Alan Goldman asked a natural question – can we describe the convex hull of spanning trees? By this time Edmonds was well into matroids, and realized (this was also known to Rado [15]) that the greedy algorithm finds a maximum weight basis of a matroid. So getting the polytope of independent sets was a breeze.

Theorem 4 (Matroid Polytope Theorem). *Let M be a matroid on S with rank function r. The convex hull of characteristic vectors of independent sets is*

$$P(M) = \{x \in \mathrm{R}^S : x \geq 0,\ x(A) \leq r(A) \text{ for all } A \subseteq S\}.$$

Edmonds proved the theorem by proving that, for any weight vector $c \in \mathrm{R}^S$, the LP problem maximize $c^T x$ subject to $x \in P(M)$ is solved by the greedy algorithm. We will see his method in more detail shortly.

Edmonds' Amazing Theorem

Now suppose we have two matroids M_1, M_2 on S and we want to describe the convex hull of common independent sets, which we write, with abuse of notation, as $P(M_1 \cap M_2)$. Clearly, every common extreme point of any two polyhedra is an extreme point of their intersection. In general, there will be other extreme points as well. It would be a rare situation indeed for the two polyhedra to fit together so neatly, that the only extreme points of the intersection were the common extreme points. But this is the case if the two polyhedra are matroid polyhedra! In lectures, Edmonds sometimes referred to his result – indeed, deservedly – as "my amazing theorem".

Theorem 5 (Matroid Intersection Polytope Theorem). *Let M_1, M_2 be matroids on S. Then*

$$P(M_1 \cap M_2) = P(M_1) \cap P(M_2).$$

Now, having generalized from one matroid to two, and from maximum cardinality to maximum weight, Edmonds went further, generalizing the matroid concept. The polyhedron $P(M)$ has the property that for every weight vector c, the greedy algorithm optimizes $c^T x$ over $P(M)$. Edmonds discovered a more general class of polyhedra having this property. And, one that permits generalization of the Amazing Theorem, too.

Polymatroids

Edmonds considered nonempty polyhedra of the form $P(f) = \{x \in \mathrm{R}^S : x \geq 0,\ x(A) \leq f(A) \text{ for all } A \subseteq S\}$, where f is submodular. He called such a polyhedron a *polymatroid*. It turns out that any such $P(f)$ can be described by an f which is also increasing and satisfies $f(\emptyset) = 0$. Such functions are now called *polymatroid functions*. Of course, matroid rank functions are polymatroid functions, and matroid polyhedra are polymatroids.

Generalizing his method for matroids, he considered the dual LP problems

$$\max\ c^T x:\ x \geq 0,\ x(A) \leq f(A) \text{ for all } A \subseteq S \tag{1}$$

$$\begin{gathered} \min\ \sum(f(A)y_A:\ A \subseteq S) \\ \text{subject to} \\ \sum(y_A:\ A \subseteq S,\ e \in A) \geq c_e,\ \text{for all } e \in S \\ y_A \geq 0,\ \text{for all } A \subseteq S. \end{gathered} \tag{2}$$

Now order S as $e_1 \dots, e_n$ such that $c_{e_1} \geq \cdots \geq c_{e_m} \geq 0 \geq c_{e_{m+1}} \geq \cdots \geq c_{e_n}$, and define T_i to be $\{e_1, \dots, e_i\}$ for $0 \leq i \leq n$.
The greedy algorithm is: Put $x_{e_i} = f(T_i) - f(T_{i-1})$ for $1 \leq i \leq m$ and $x_j = 0$ otherwise.
The dual greedy algorithm is: Put $y_{T_i} = c_{e_i} - c_{e_{i+1}}$ for $1 \leq i \leq m-1$, put $y_{T_m} = c_{e_m}$ and put all other $y_A = 0$.

The resulting solutions satisfy the LP optimality conditions for (1) and (2). It is also clear that if f is integral, then so is x, and if c is integral, then so is y. In particular, this proves a significant generalization of Theorem 4. As we shall see, it proves much more.

Polymatroid Intersection

Now here is the topper – Edmonds puts all three directions of generalization together.

Theorem 6 (Weighted Polymatroid Intersection). *Let f_1, f_2 be polymatroid functions on S, and let $c \in \mathrm{R}^S$. Consider the LP problem*

$$\begin{gathered} \max\ c^T x \\ x(A) \leq f_1(A),\ \text{for all } A \subseteq S \\ x(A) \leq f_2(A),\ \text{for all } A \subseteq S \\ x_e \geq 0,\ \text{for all } e \in S. \end{gathered} \tag{3}$$

and its dual problem

$$\begin{gathered} \min\ \sum(r_1(A)y_A^1 + r_2(A)y_A^2:\ A \subseteq S) \\ \text{subject to} \\ \sum(y_A^1 + y_A^2:\ A \subseteq S,\ e \in A) \geq c_e,\ \text{for all } e \in S \\ y_A^1, y_A^2 \geq 0,\ \text{for all } A \subseteq S. \end{gathered} \tag{4}$$

(a) *If f_1, f_2 are integer-valued, then (3) has an integral optimal solution.*

(b) *If c is integral, then (4) has an integral optimal solution.*

We will sketch Edmonds' ingenious proof. Consider an optimal solution $\hat{y}^1, \hat{y}^2$ of (4). The problem of optimizing over y^1 while keeping y^2 fixed at $\hat{y}^2$ is an LP problem of the form (2), which can be optimized by the dual greedy algorithm. Therefore, we can replace $\hat{y}^1$ by the output of that algorithm. Now we can fix y^1 and similarly replace $\hat{y}^2$.

We conclude that (4) has an optimal solution that is an optimal solution to a problem in which the constraint matrix has a very special structure. Namely, its columns split into two sets, each of which consists of the characteristic vectors of a telescoping family of subsets of S. Edmonds proved – it is a nice exercise – that such a matrix is totally unimodular. It follows that (4) has an optimal solution that is integral, assuming that c is integral, proving (b). Now with the benefits of hindsight, we can invoke the theory of total dual integrality, and (a) is proved. In fact, Edmonds did not have that tool. He used another argument, again a clever indirect use of total unimodularity, to prove (a).

There are several important consequences of the above theorem. For example, taking f_1, f_2 to be matroid rank functions, we get the Amazing Theorem. Taking each $c_j = 1$, we get the following important result.

Theorem 7 (Polymatroid Intersection Theorem). *Let f_1, f_2 be polymatroid functions on S. Then*

$$\max(x(S): \ x \in P(f_1) \cap P(f_2)) = \min(f_1(A) + f_2(S \backslash A): \ A \subseteq S).$$

Moreover, if f_1, f_2 are integer-valued, then x can be chosen integral.

Postlude

In the years since the sixties, much progress has been made, far too much to summarize here. I mention a few highlights, relating them to the work of the sixties. The books of Frank [6] and Schrijver [17] can be consulted for more detail.

Submodularity and Convexity

Let us call a function f *supermodular* if $-f$ is submodular, and call it *modular* if it is both submodular and supermodular. It is easy to see that a function f is modular if and only if it satisfies $f(A) = m(A) + k$ for some $m \in \mathrm{R}^S$ and $k \in \mathrm{R}$. Then we have the beautiful Discrete Separation Theorem of Frank [5].

Theorem 8. *Let f, g be functions defined on subsets of S such that f is submodular, g is supermodular, and $f \leq g$. Then there exists a modular function h such that $f \leq h \leq g$. Moreover, if f and g are integer-valued, then h may be chosen integer-valued.*

In fact, this theorem can be proved from the Polymatroid Intersection Theorem 7, and conversely. Its first part resembles a well-known result about the separation of convex and concave functions by an affine function. Actually, there is a connection. Lovász [11] defined the extension $\hat{f}$ to R_+^S of a set function f, using ideas suggested by the dual greedy algorithm. He then proved that $\hat{f}$ is convex if and only if f is submodular. Using this, one can derive the first part of Frank's theorem from the convexity result.

Submodular Function Minimization

The problem of minimizing a submodular function (given by an evaluation oracle) is fundamental. Its special cases include finding a minimum capacity s,t-cut in a directed graph, and (in view of the Matroid Intersection Theorem) finding the maximum size of a common independent set of two given matroids.

A good characterization of the minimum follows from the work of Edmonds [2]. One way to describe it is this. One can reduce the problem of minimizing a submodular function g to the problem of minimizing $f(A) + u(S \backslash A)$, where $u \geq 0$ and f is a polymatroid function. But

$$\max(x(S) : x \in P(f),\ x \leq u) = \min(f(A) + u(S \backslash A) :\ A \subseteq S).$$

This is a special case of the Polymatroid Intersection Theorem 7, but it can easily be proved directly. Now suppose we have A and x giving equality above. Then $x \in P(f)$ can be certified by expressing it as the convex combination of a small number of extreme points of $P(f)$, and each extreme point can be certified by the polymatroid greedy algorithm.

So much for characterizing the minimum. What about an algorithm to find the minimum? The first efficient algorithm was found by Grötschel, Lovász and Schrijver [7], based essentially on the equivalence, via the ellipsoid method, of separation and optimization. More recently, Iwata, Fleischer, and Fujishige [8] and Schrijver [16] gave combinatorial algorithms. Both use explicitly the method of certifying membership in $P(f)$ described above.

Weighted Polymatroid Intersection

The problem of finding an efficient algorithm for weighted polymatroid intersection, and other closely related models such as optimal submodular flows, was left open by Edmonds. (He, and also Lawler, did solve the special case of weighted matroid intersection.) Efficient combinatorial algorithms now exist. One may summarize their development as follows. Lawler and Martel and also Schönsleben gave efficient algorithms for the maximum component-sum problem. Cunningham and Frank combined this with a primal-dual approach to handle general weights. These algorithms need as a subroutine one of the above algorithms for submodular function minimization.

Matroid Intersection and Matching

Weighted versions of matroid intersection and matching have a common special case, optimal bipartite matching. In addition they share similar attractive results – polyhedral descriptions, and efficient solution algorithms. It is natural, therefore, to ask whether there exists a common generalization to which these results extend. Several candidates have been proposed. The most important one, proposed independently by Edmonds and Lawler, has several equivalent versions, one of which goes as follows. Given a graph G and a matroid M on its vertex-set, a *matroid matching* is a matching of G whose covered vertices form an independent set in M. It turned out that finding a maximum-weight matroid matching, even when the weights are all 1, is a hard problem. However, in the late seventies Lovász found an efficient algorithm and a min-max formula for the case where M arises from a given linear representation. Recently, Iwata and Pap independently have found efficient algorithms for the weighted version, answering a question that was open for more than thirty years.

References

[1] J. Copeland, *Colossus: The Secrets of Bletchley Park's Codebreaking Computers*, Oxford University Press, 2006.

[2] J. Edmonds, Submodular functions, matroids, and certain polyhedra in: R. Guy et al. (eds) *Combinatorial Structures and their Applications*, Gordon and Breach, New York, 1970, 69–87.

[3] Edmonds, Jack, A glimpse of heaven, in: J.K. Lenstra et al. (eds), *History of Mathematical Programming* North-Holland, Amsterdam, 1991. 32–54.

[4] J. Edmonds, Matroid partition, in: M. Juenger et al. (eds.) *Fifty Years of Integer Programming* Springer Verlag, Heidelberg, 2010, 199–217.

[5] A. Frank, An algorithm for submodular functions on graphs, Ann. Discrete Math 16 (1982), 97–210.

[6] A. Frank, *Connections in Combinatorial Optimization*, Oxford, U.K., 2011.

[7] M. Grötschel, L. Lovász and A. Schrijver, The ellipsoid method and its consequences in combinatorial optimization, Combinatorica 1 (1981), 169–197.

[8] S. Iwata, L. Fleischer, and S. Fujishige, A Combinatorial strongly polynomial algorithm for minimizing submodular functions, J. ACM 48 (2001), 761–777.

[9] D.G. Kelly and G.-C. Rota, Some problems in combinatorial theory, in: *A Survey of Combinatorial Theory* North-Holland, Amsterdam, 1973, pp. 309–312.

[10] A. Lehman, A Solution of the Shannon switching game, J. SIAM 12 (1964). 687–725.

[11] L. Lovász, Submodular functions and convexity in: Bachem et al. (eds.) *Mathematical Programming: The State of the Art*, Springer Verlag 1982.

[12] J. Oxley, *Matroid Theory*, Oxford University Press, Oxford, 2011.

[13] W. R. Pulleyblank, Edmonds, matching, and the birth of polyhedral combinatorics, this volume.

[14] R. Rado, A theorem on independence relations, Quarterly J. Math. Oxford (2) 13(1942), 83–89.

[15] R. Rado, Note on independence functions, Proc. London Math Soc (3) 7(1957), 300–320.

[16] A. Schrijver, A combinatorial algorithm minimizing submodular functions in polynomial time, J. Comb. Theory B 80(2000), 346–355.

[17] A. Schrijver, *Combinatorial Optimization: Polyhedra and Efficiency.* Springer-Verlag Berlin, 2003.

[18] W.T. Tutte, A homotopy theorem for matroids, I and II, Trans. AMS 88(1958), 153–184.

[19] W.T. Tutte, Matroids and graphs, Tran. AMS 89(1959), 527–552.

[20] W.T. Tutte, Lectures on matroids, J. Res. NBS 69B (1965), 1–47.

[21] W.T. Tutte, The coming of the matroids, Surveys in Combinatorics, LMS Lecture Note Series 267 (1999), 3–14.

[22] H. Whitney, The abstract properties of linear dependence, Am. J.Math. 57(1935), 509–533.

William H. Cunningham
Department of Combinatorics
& Optimization
University of Waterloo
Waterloo, ON
Canada, N2L 3G1
whcunnin@uwaterloo.ca