A Characterization of the Monge Property and its Connection to Statistics

Wolfgang W. Bein [1]
Department of Computer Science
University of New Mexico
Albuquerque, NM 87131

Pramod K. Pathak
Department of Mathematics and Statistics
University of New Mexico
Albuquerque, NM 87131

Written:	October 1990
Revised:	October 1991 and February 1992

[1] supported by Sandia National Laboratories grant, SURP No. 05-9858 task9

Abstract

We show that any Monge array $g(x,y)$ has a representation of the form $g(x,y) = u(x) + v(y) + F(x,y)$, where F is a distribution function $F(x,y) = \sum_{i \leq x, j \leq y} p_{ij}$ for nonnegative p_{ij}. This result as well as our applications show that there exists a close connection between statistics and the Monge property. Our result properly encompasses the earlier results on Monge arrays by Gilmore and Gomory. We furnish applications of these results to algorithms for the integration of surveys.

Keywords: Monge property, greedy algorithm, transportation problem, distribution function, integration of surveys, controlled selection

1 Introduction

Consider the two-dimensional array $g : N_n \times N_n \to R$, where $N_n := \{0, \ldots, n\}$. g is called a *Monge array* or *Monge* for short if

$$g(x_1, y_1) + g(x_2, y_2) \geq g(x_1, y_2) + g(x_2, y_1) \text{ for all } x_1 \leq x_2 \text{ and } y_1 \leq y_2 \quad (1.1)$$

Furthermore g is called *monotone* if $g(., y)$ and $g(x, .)$ are nondecreasing functions. Finally, if (1.1) holds with "$\leq$" in the first inequality we call g *reverse Monge*, where monotonicity means "nonincreasing" for the reverse case.

Define now

$$\square g(x,y) = g(x,y) - g(x, y-1) - g(x-1, y) + g(x-1, y-1),$$

where we assume $g(s,t)$ to vanish for any negative argument s or t.

We can equivalently define g to be Monge by

$$\square g(x,y) \geq 0 \text{ all } x, y \geq 1 \quad (1.2)$$

If

$$\Box g(x,y) \geq 0 \text{ all } (x,y) \in N_n \times N_n - (0,0) \tag{1.3}$$

then g is Monge and monotone.

The Monge property is named in honor of the French mathematician G. Monge who studied the property in the eighteenth century [Mon81]. It was rediscovered in 1961 by A.J. Hoffman [Hof61], when he showed that a transportation problem can be solved by a greedy method if the underlying costs are Monge.

Recently the Monge property has again been shown to be useful in diverse fields: in the speedup dynamic programs for the study of DNA as well as other problems [Yao80, LS90, EGG88], problems in computational geometry [AP88], statistics [FP89], and the theory of greedy algorithms. Finally the classical Monge property has been generalized to higher dimensions (see [BBP92]).

Notice that g is Monge if it is linearly separable, i.e $g(x,y) = u(x) + v(y)$. In this case (1.2) holds with equality. In fact this stronger property underlies some of the speed-up results in dynamic programming [Yao80]. In the study of integer programs, Gilmore and Gomory [GG64] have given a rather general way to generate Monge arrays that are not linearly generated via certain integrals.

In this paper we give a characterization of the Monge property that encompasses their result in a natural way. We will now give the characterization and show its validity in Section 2. Section 3 shows that our result is a natural generalization of the Gilmore-Gomory result.

A function $F : N_n \times N_n \to R^+$ is called a *distribution function* if it is of the form

$$F(x,y) = \sum_{i \le x, j \le y} p_{ij} \tag{1.4}$$

for nonnegative p_{ij}.

We are now ready to formulate our result:

Theorem 1 *Let $g : N_n \times N_n \to R$ be a Monge array. Then there exists a distribution function $F : N_n \times N_n \to R^+$ and functions $u, v : N_n \to R$ such that*

$$g(x,y) = u(x) + v(y) + F(x,y) \tag{1.5}$$

and

Theorem 2 *Let $g : N_n \times N_n \to R$ be a reverse Monge array. Then there exists a distribution function $F : N_n \times N_n \to R^+$ and functions $u, v : N_n \to R$ such that*

$$g(x,y) = u(x) + v(y) - F(x,y) \tag{1.6}$$

2 Proof of Characterization

We will only prove Theorem 1, the proof of Theorem 2 is entirely analogous, mutatis mutandis, to the proof of Theorem 1. We first show that a function that has the form (1.5) is Monge:

Remark 2.1 *Given a distribution function $F : N_n \times N_n \to R^+$ and functions $u, v : N_n \to R$ with*

$$g(x, y) = u(x) + v(y) + F(x, y)$$

then g is Monge.

Proof: First notice that $u(x)$, $u(x-1)$, $v(y)$, $v(y-1)$ cancel out in (1.2). Therefore

$$\Box g(x, y) = p_{x,y} \geq 0.$$

QED

Now we shall show that (1.5) is also necessary.

Lemma 1 *Let $g : N_n \times N_n \to R$ with*

$$\Box g(x, y) = 0 \ \text{all } N_n \times N_n - (0, 0). \tag{2.1}$$

Then there exist functions $u, v : N_n \to R$ such that

$$g(x, y) = u(x) + v(y).$$

Proof: Due to (2.1) we have

$$g(0, 0) = \sum_{0 \leq i \leq x, 0 \leq j \leq y} \Box g(i, j) = g(x, y) - g(x, 0) - g(0, y) + g(0, 0).$$

Therefore

$$g(x, y) = g(x, 0) + g(0, y) =: u(x) + v(y).$$

QED

Lemma 2 *Let $g : N_n \times N_n \to R$ then there exist functions $u, v : N_n \to R$ such that $g(x, y) = u(x) + v(y)$ is monotone.*

Proof: First consider the first two "columns" $g(0, y)$ and $g(1, y)$. If the two functions are not monotone add a constant $u(1)$. Then proceed with columns $1, 2;\ \ 2, 3$ and so forth. Repeat the process for rows to obtain $v(y)$. *QED*
Notice that if g is Monge, so is $g + u + v$.

Lemma 3 *Let $g : N_n \times N_n \to R$ be Monge and monotone. Furthermore let*

$$F(x, y) := \sum_{i \leq x, j \leq y} p_{ij} \text{ with } p_{ij} := \Box g(i, j) \quad i, j \in N_n. \tag{2.2}$$

Then F is a distribution function and

$$\Box(F - g) = 0.$$

Proof: All we have to show is that $p_{ij} \geq 0$ all $i, j \in N_n$. But this follows from the fact that g is Monge for nonzero pairs ij and else from the fact that g is monotone. *QED*

Now we are ready to prove our result: *Proof:*[of Theorem 1]: Due to Lemma 2 we can assume that g is indeed monotone. Then let F be as in Lemma 3 to obtain the result from Lemma 1. *QED*

3 The Gilmore-Gomory Results

We shall now show that our characterization is a natural generalization of the results of Gilmore and Gomory. We will consider the reverse Monge case where they consider arrays of the form:

$$w(i,j) = \begin{cases} \int_{\alpha_i}^{\beta_j} f(y)\,dy & \alpha_i \le \beta_j \\ \int_{\beta_j}^{\alpha_i} g(y)\,dy & \beta_j < \alpha_i \end{cases} \tag{3.1}$$

where f and g are given nonnegative integrable functions [2] and $\alpha_0 \le \ldots \alpha_n$, $\beta_0 \le \ldots \beta n$ are given real parameters.
Let F, G be the primitives of f and g respectively. Equation (3.1) can be written as

$$\begin{aligned} w(i,j) &= \max(0, F(\beta_j) - F(\alpha_i)) + \max(0, G(\alpha_i) - G(\beta_j)) && (3.2) \\ &= G(\beta_j) - F(\alpha_i) + \max(F(\beta_j), F(\alpha_i)) + \max(G(\alpha_i), G(\beta_j)) && (3.3) \end{aligned}$$

Then we have

$$\Box w(i,j) = \Box \max(F(\beta_j), F(\alpha_i)) + \Box \max(G(\alpha_i), G(\beta_j)) \tag{3.4}$$

which can be seen to be

$$\Box w(i,j) = \begin{cases} (F+G)(y_1) - (F+G)(y_2) & y_1 < y_2 \\ 0 & y_1 \ge y_2 \end{cases}, \tag{3.5}$$

where

$$y_1 = \max(\alpha_{i-1}, \beta_{j-1}) \text{ and } y_2 = \min(\alpha_i, \beta_j). \tag{3.6}$$

Thus to define the F in (1.6) we set

$$p_{i,j} = -\Box w(i,j). \tag{3.7}$$

[2] Gilmore and Gomory also make the assumption $f + g \ge 0$, our results hold for that case as well.

Equation (3.5) shows that characterization (1.6) is a proper generalization of the Gilmore-Gomory result (3.1), since $y_1 \geq y_2$ will always hold in (3.5) for some pairs i, j and thus some $p_{i,j}$ are 0. Thus while the Gilmore-Gomory result covers many important Monge arrays it does not characterize all Monge situations.

4 Applications to Statistics

Consider the problem of integration of two surveys. It is proposed to carry out two surveys (S, P) and (T, Q) on Z. Let $p_i,\ i = 1, \ldots, m$, denote the probability of selection associated with the sample s_i in S for the first survey, $\sum_i p_i = 1$. Similarly let $q_j,\ j = 1, \ldots, n$, denote the probability of selection associated with the sample t_j in T for the second survey.

An integrated survey is a joint probability distribution (sampling program) on $S \times T$ which assigns a probability x_{ij} to the pair $(s_i,\ t_j)$ and realizes the marginal probability distributions P and Q required for the two respective surveys. An optimal integrated survey is an integrated survey which given a distance (cost) function minimizes the expected distance between the two marginal surveys. (This is equivalent to maximizing the expected overlap between the two marginal surveys.) In terms of the transportation problem, the problem of the optimal integration of two surveys is as follows:

Integration of two surveys:

$$\text{Minimize} \quad E(d) = \sum_i \sum_j d_{ij} x_{ij}$$

$$\text{subject to} \quad \begin{aligned} \sum_j x_{ij} &= p_i \qquad (i = 1, \ldots, n) \\ \sum_i x_{ij} &= q_j \qquad (j = 1, \ldots, m) \\ x_{ij} &\geq 0 \text{ for all } i \text{ and } j \ . \end{aligned}$$

This problem is known as the Hitchcock transportation problem (see e.g. Chvátal [Chv83]). We note that the problem of controlled selection has the same formulation except for a a few changes in the terminology. In the integration of surveys, the objective function is to be minimized, whereas in controlled selection this function is to be maximized. The function d in the context of controlled selection is referred to as the distance function and is taken to represent a measure of preference.

Hoffman [Hof61] showed the now classical result that above transportation problem can be solved by the "north-west corner rule" if the underlying cost matrix satisfies the Monge condition in the case of maximization and reverse Monge property in the case of minimization. (see [Hof61] for details). Whereas the best known algorithms for the general transportation problem has complexity $O(mn^2 \lg n + n^2 \lg^2 n)$ in the Monge case an optimal solution can be found in $O(m + n)$.

We use our technique to establish the optimality of the Goodman-Kish approach [GK50] to the problem of controlled selection. There are two strata. Stratum 1 contains six PSU's A, B, C, D, E and F of which A, D and E are inland while B, C and F are coastal. Stratum 2 contains five PSU's a, b, c, d, and e of which only a is coastal and the others are inland. The object is to design a joint sampling scheme which maximizes the probability of selecting a pair of dissimilar units, i.e., an inland unit from one stratum and a coastal from the other according to the following probabilities:

Stratum 1							Stratum 2					
Units	A	B	C	D	E	F	Units	a	b	c	d	e
Prob	.1	.15	.1	.2	.25	.2	Prob	.2	.3	.1	.15	.25

Let the distance between PSU's from the two strata be zero if the two units are of the same type and one otherwise. The distance matrix is, therefore, as follows:

		Stratum 2				
		a	b	c	d	e
	A	1	0	0	0	0
	D	1	0	0	0	0
Stratum 1	E	1	0	0	0	0
	B	0	1	1	1	1
	C	0	1	1	1	1
	F	0	1	1	1	1

The distance matrix is a Mone matrix and therefore the north-west corner rule gives the following optimal solution.

		Stratum 2					
		a	b	c	e	d	
	A	.1	0	0	0	0	.1
	D	.1	.1	0	0	0	.2
Stratum 1	E	0	.2	.05	0	0	.25
	B	0	0	.05	.1	0	.15
	C	0	0	0	.1	0	.1
	F	0	0	0	.05	.15	.2
		.2	.3	.1	.25	.15	

In the next example, we establish the "optimality" of Lahiri's selection scheme [Lah54]. Under this scheme, PSU's from a given geographical area are listed on the sampling frame in a serpentine order. To minimize the travel cost, a variant of the Northwest Algorithm is used for selecting two sample units. For details, we refer the reader to Lahiri's paper [Lah54].

In this method, the cost of the survey is assumed to be proportional to the distance between the PSU's and is taken to be $d(i, j) = |i - j|$. The matrix d(i,j) is reverse Monge (this follows from the fact that d is a distribution function), which implies that Lahiri's selection scheme does indeed minimize the expectation of this distance function.

In fact, these techniques have enabled us to generalize the Lahiri result to the intergration of k surveys under a min-max distance function $d(i_1, \ldots, i_k) := \max_{j \leq k} i_j - \min_{j \leq k} i_j$, see [BBP92].

References

[AP88] A. Aggarwal and J. Park. Notes on searching in multidimensional monotone arrays. In *Proceedings 29th Annual Symposium on Foundations of Computer Science*, pages 497 – 511, 1988.

[BBP92] W.W. Bein, P. Brucker, and P.K. Pathak. A monge property for the d-dimensional transportation problem. Technical Report CS92-1, University of New Mexico, Department of Computer Science, 1992.

[Chv83] V. Chvátal. *Linear Programming.* Freeman, 1983.

[EGG88] D. Eppstein, Z. Galil, and R. Giancarlo. Speeding up dynamic programming. In *Proceedings 29th Annual Symposium on Foundations of Computer Science*, pages 488 – 495, 1988.

[FP89] M. Fahimi and P.K. Pathak. Applications of a size reduction technique for transportation problems in sampling. Technical report, University of New Mexico, Department of Mathematics and Statistics, 1989.

[GG64] P.C. Gilmore and R.E. Gomory. Sequencing a one state-variable machine: A solvable case of the traveling salesman problem. *Operations Research*, 12:655 – 679, 1964.

[GK50] R. Goodman and L. Kish. Controlled selection–a technique in probability sampling. *J. Amer. Statist. Assoc.*, 45:350 – 372, 1950.

[Hof61] A.J. Hoffman. On simple linear programming problems. In *Convexity, Proc. Symposia in Pure Mathematics*, volume 7, pages 317 – 327, Providence, RI, 1961. American Mathematical Society.

[Lah54] D.B. Lahiri. Technical paper on some aspects of development of the sample design. *SIAM Journal on Computing*, 15:246 – 316, 1954.

[LS90] L.L. Larmore and B. Schieber. On-line dynamic programming with applications to the prediction of rna secondary structure. In *Proceedings First Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 503 – 512, 1990.

[Mon81] G. Monge. *Déblai et remblai.* Mémoires de l' Académie des Sciences, Paris, 1781.

[Yao80] F.F. Yao. Speed-up in dynamic programming. *SIAM J. Alg. Disc. Methods*, 3:532 – 540, 1980.