Inequalities in Statistics and Probability
IMS Lecture Notes–Monograph Series Vol. 5 (1984), 84–103

LEAST ABSOLUTE VALUE AND MEDIAN POLISH

BY J. H. B. Kemperman
University of Rochester

We are interested in best L_p-approximations $\Sigma_r \beta_r g_r(x)$ to a given finite array of numbers $z^{(o)}(x)$, $(x \in X)$. For the case $p > 1$, a natural iterated polishing method is shown to converge to the unique optimal solution. Let $p = 1$. Several conditions are obtained, each of which is necessary and sufficient for a given array of residuals $z(x)$ $(x \in X)$ to be optimal. Detailed results are derived for the case of a two-way $m \times n$ layout, allowing several observations $z^{(o)}_{ijk}$ in cell (i, j). For instance, a set of residuals is optimal if and only if there exists a solution to an associated moment problem with given marginals, which depends only on the signs σ_{ijk} of the residuals z_{ijk}. This criterion leads to an elegant and efficient max-flow-min-cut type of algorithm for calculating a best L_1-approximation. For the case of a single observation in each cell, it is also determined precisely which pairs (m, n) are 'safe' for Tukey's median polish, in the sense that the endproduct of an $m \times n$ polish is necessarily a best L_1-approximation. The answer depends on the type of allowable medians.

1. Introduction. Let the $m \times n$ matrix $\mathbf{Z}^{(o)} = (z^{(o)}_{ij})$ represent a two-way table of observations. An elementary way of arriving at a reasonable additive approximation $\alpha_i + \beta_j$ is by means of median polish, as developed by Tukey (1977); see Section 4 for further details. An algorithm in APL and further comments can be found in Anscombe ((1981) p. 106, 382).

One motivating idea behind median polish is that it *might* minimize the L_1-norm of the matrix $\mathbf{Z} = (z_{ij})$ of residuals $z_{ij} = z^{(o)}_{ij} - \alpha_i - \beta_j$. However, this is not always true as follows already from the well-known fact that the norm of the final endproduct of a median polish or mid-median polish may not be the same when starting with a polishing of the rows as when starting with a polishing of the columns.

These endproducts will be called an EMP or EMMP, respectively. More generally, an $m \times n$ matrix $\mathbf{Z}$ will be called an EMP or EMMP, respectively, when 0 is a median or mid-median, respectively of each row and each column.

The matrix $\mathbf{Z}$ of residuals will be said to be *optimal* if its norm cannot be further reduced. For this it is necessary that $\mathbf{Z}$ be an EMP. It is shown (Theorem 6) that for each choice of (m, n) there exist non-optimal EMP's. There even exist non-optimal EMMP's, unless (m, n) is one of the special pairs (2, n); (3, 4); (4, 4); (4, 5) and (4, 6), (assuming that $2 \leq m \leq n$). Thus, if $m = 4$ and $n = 6$ then the endproduct of a convergent mid-median polish process is always optimal. This is false when $m = 3$ and $n = 3$ or 5.

The main purpose of the present paper is to derive efficient tests for optimality together with explicit procedures for improving a given non-optimal matrix. Many of our results lead to an explicit algorithm, usually safer and faster than median polish, though that algorithm may not be spelled out in any great detail. For, our principal goal is to achieve a good theoretical understanding of the main problem.

Most results are developed for the general regression problem, where one wants to minimize the L_p-norm (2.1) by a suitable choice of the free parameters β_r. Median polish carries over to this general problem in a natural way. We show in Theorem 1 that this

[1] Research supported in part by the National Science Foundation.
AMS 1980 subject classifications. Primary, 62J05; secondary, 41A50, 65D10.
Key words and phrases: Linear regression, least absolute value, best L_1 approximation, criterion for optimality, two-way layout, median polish, given marginals, max-flow-min-cut, algorithm.

generalized polish always converges to the unique optimal solution, provided $p > 1$.

From Section 5 on, it is assumed that $p = 1$. Then the optimality of $\mathbf{Z} = \{z(x); x \in X\}$ depends only on the associated sign pattern $\{\text{sgn } z(x); x \in X\}$ and one may speak of optimal sign patterns. It turns out that an optimal sign pattern remains optimal when one or more of the elements $+1$ or -1 are replaced by 0. In the different criteria for optimality, a central role is played by the set $D = \{x \in X: z(x) = 0\}$. For instance, in the case of an n-way layout, one necessary and sufficient condition for optimality requires the existence of a measure on D having preassigned marginals, see Section 7. The sections 8 and 9 are concerned with a two-way layout, allowing several observations per cell.

There is a large literature on explicit algorithms leading to an optimal L_1-approximation. Each of these algorithms amounts to a descent method of some type, often restricted to the finite set of basic solutions of the associated linear programming problem. See the surveys by Gentle (1977) and Kennedy and Gentle (1980) pp. 515–559.

A selected list of such papers has been included in the bibliography. Space did not allow us to give an adequate discussion of the many cross relations which exist between these papers and the present one.

2. Stating the problem. In this paper, X is a fixed finite index set and $Z^{(o)} = \{z^{(o)}(x): x \in X\}$ a given collection of real numbers or observations. Further, $g_r : X \to R$ $(r = 1, \ldots, M)$ is a given set of linearly independent functions on X. We will be interested in diferent aspects of the problem to

$$\text{minimize:} \quad S = \Sigma_{x\in X}\, \omega(x)|z^{(o)}(x) - \Sigma_{r=1}^{M}\beta_r g_r(x)|^p. \tag{2.1}$$

Here, $p \geqslant 1$ and only the β_r are unknown. The weights $\omega(x) > 0$ may indicate the multiplicity or importance of the corresponding observations.

This problem arises in many ways, for instance, as a maximum likelihood problem when

$$z^{(o)}(x) = \Sigma_{r=1}^{M}\beta_r g_r(z) + \eta(x), \qquad (x \in X)$$

where the errors $\eta(x)$ $(x \in X)$ are independent with a density type $c(\omega)\exp(-\omega(x)|y|^p)$, $(y \in R)$.

We will especially be interested in the case $p = 1$. One situation we have in mind, see Sections 8 and 9, is that of a two-way $m \times n$ layout with observations $z_{ijk}^{(o)}$ in cell (i, j), with $i \in Y_1 = \{1, \ldots, m\}$; $j \in Y_2 = \{1, \ldots, n\}$ and $k = 1, \ldots, k_{ij}$, where one wants to minimize the L_1-norm

$$S = \Sigma_{i,j,k}|z_{ijk}^{(o)} - \alpha_i - \beta_j|. \tag{2.2}$$

Here, X is taken as the set of triplets (i, j, k) with $i \in Y_1$, $j \in Y_2$, $k \in \{1, \ldots, k_{ij}\}$ while $\omega(x) = 1$. Further, $M = m + n$ and

$$\begin{aligned} g_r(i, j, k) &= \delta_i^r \qquad \text{for } r = 1, \ldots, m; \\ &= \delta_j^{r-m} \qquad \text{for } r = m + 1, \ldots, M; \end{aligned} \tag{2.3}$$

($\delta_q^r = 1$ or 0 if $q = r$ or $q \neq r$, respectively). The general layout problem is to minimize a norm of the form

$$S = \Sigma_{x\in X}\, \omega(x)|z^{(o)}(x) - \Sigma_{t\in T}\, \beta_t(\phi_t(x))|. \tag{2.4}$$

Here, for each $t \in T$, ϕ_t: $X \to Y_t$ is a given function, while the $\beta_t(y)$ are unknown parameters. Often X is a subset of $\mathcal{R}^n$ while $\phi_t(y)$ is expressed in terms of the coordinates x_h of x, e.g., $\phi_t(x) = x_1$ or $\phi_t(x) = (x_1, x_2)$.

In this illustration, the function g_r in (2.1) becomes

$$g_{t,y}(x) = \delta^y_{\phi_t(x)}, \quad (t \in T; y \in Y_t) \tag{2.5}$$

while $\beta_t(y)$ plays the role of β_r. The support of $g_{t,y}$ is

$$L_t(y) = \{x \in Z : g_{t,y}(x) \neq 0\} = \{x \in X : \phi_t(x) = y\} \tag{2.6}$$

and will also be referred to as a 'layer'. In the case (2.2) one would have $T = \{1, 2\}$, $\phi_1(i, j, k) = i$ and $\phi_2(i, j, k) = j$. Moreover, $L_1(i)$ and $L_2(j)$, respectively, would be the set of points $x = (i, j, k)$ having a fixed component i or j, respectively.

The general layout problem is equivalent to having, for each $t \in T$, a partition of X into disjoint sets $L_t(y)$ with associated additive parameters $\beta_t(y)$.

Remark. As will be shown in a subsequent paper, most results of the present paper carry over to the case where the exponent p in (2.1) is replaced by a function $p(x) \geq 1$ $(x \in X)$.

3. The p-Centre of a Mass Distribution. Be given a mass distribution on the reals having mass $q_i > 0$ at y_i $(i = 1, \ldots, n)$ and suppose one wants to minimize $\psi(s) = \Sigma_{i=1}^n q_i|s - y_i|^p$. This is somewhat comparable to least squares relative to the (variable) weights $q_i|y_i - s|^{p-2}$, as observed by Mosteller and Tukey (1977) p. 365. If $p < 2$ then relatively less weight is given to the very large observations.

If $p > 1$ then $\psi(s)$ is strictly convex and the minimum on hand is uniquely achieved at the so-called p-centre $s^\circ = \text{mean}_p\{y_i : q_i\}$ of the mass distribution. It is the unique solution of the equation

$$\Sigma_{i=1}^n \text{sgn}(s^\circ - y_i) q_i |s^\circ - y_i|^{p-1} = 0,$$

(where $\text{sgn}(u) = -1$, 0 or $+1$, depending on the sign of u). In particular, $\text{mean}_2\{y_i : q_i\}$ is nothing but the ordinary mean.

If $0 < p < 1$ then $\psi(s)$ would be strictly concave as long as s differs from the y_i and, thus, $\psi(s)$ takes its minimal value (only) at one of the points y_i. If $p = 1$ then $\psi(s)$ is piecewise linear and convex and, hence, attains its minimal value precisely at the values s where the nondecreasing derivative $\psi'(s) = \Sigma_{i=1}^n q_i \text{sgn}(s - y_i)$ changes sign from negative to positive; this always includes one of the points y_i. Such a 1-mean or median s° may also be defined by the inequalities

$$\Sigma_{y_i < s^\circ} q_i \leq Q/2 \leq \Sigma_{y_i \leq s^\circ} q_i, \quad \text{where } Q = \Sigma_{i=1}^n q_i. \tag{3.1}$$

This median is unique unless the second sum can take the value $Q/2$. The latter happens, for instance, when n is even and $q_i = 1$ for all i. The notation

$$s^\circ = \text{mean}_1\{y_i : q_i\} = \text{med}\{y_i : q_i\} \quad \text{or} \quad \text{mean}_1\{y_i : q_i\} = s^\circ$$

will simply indicate that s° is a median for the distribution on hand. The set of all medians is a finite closed interval $[s', s'']$. Its midpoint $(s' + s'')/2$ is called the mid-median and will be denoted as $\text{Med}\{y_i : q_i\}$.

LEMMA 1. *In order that* $\beta^* = (\beta^*_1, \ldots, \beta^*_M)$ *achieves the minimum in (2.1), it is necessary that the residuals*

$$z(x) = z^{(o)}(x) - \Sigma_{r=1}^M \beta^*_r g_r(x) \qquad (x \in X) \tag{3.2}$$

satisfy

$$\text{mean}_{p,r}\{z(x) = 0\} \quad \text{for } r = 1, \ldots, M. \tag{3.3}$$

The latter mean is defined as

$$\text{mean}_{p,r}\{z(x)\} = \text{mean}_p\{z(x)/g_r(x) : \omega(x)|g_r(x)|^p\}. \tag{3.4}$$

If $p > 1$ then the minimum in (2.1) is achieved at a unique point β^* and condition (3.3) is also sufficient.

Proof. Fix $1 \leq r \leq M$. A necessary condition is that

$$\Sigma_{x \in X} \omega(x)|z(x) - s g_r(x)|^p \qquad (s \in R) \tag{3.5}$$

takes its smallest value at $s = 0$. Since one may as well restrict $x \in X$ to the points with $g_r(x) \neq 0$, this is the same as saying that $s = 0$ minimizes the sum

$$\Sigma_{x \in X} \omega(x)|g_r(x)|^p |s - z(x)/g_r(x)|^p.$$

From the remarks preceding the Lemma, the latter in turn is equivalent to (3.3).

If $p \geq 1$ then the sum in (2.1) defines a nonnegative and convex function $F(\beta) = F(\beta_1, \ldots, \beta_M)$ on R^M. Condition (3.3) says precisely that F is minimal at β^* relative to changes in a single variable only. If $p > 1$ then F is of class C^1 and strictly convex (since the g_r are linearly independent). This easily yields the last assertion. □

Comments. Also note that the sum in (3.5) assumes its smallest value at s° if and only if $s^\circ = \text{mean}_{p,r}\{z(x)\}$, as defined in (3.4). In the case $p = 1$ this means that

$$\min_s \Sigma_{x \in X} \omega(x)|z(x) - s g_r(x)| \tag{3.6}$$

is achieved at s° if and only if s° is a median of the set of all numbers $z(x)/g_r(x)$ with $g_r(x) \neq 0$ having corresponding weights $\omega(x)|g_r(x)|$.

The last two assertions of Lemma 1 would be false when $p = 1$. Namely, medians and, in general, optimal L_1-approximations are often not unique. The following example shows that condition (3.3) is not sufficient for optimality when $p = 1$.

Choose $X = \{1, 2, 3\}$ and $\omega(x) = 1$. Further, $M = 2$; $g_1(1) = g_2(2) = 0$ while $g_r(x) = 1$, otherwise. Finally, $z(1) = -1$; $z(2) = +1$ and $z(3) = 0$. Then condition (3.3) holds with $\beta^*_1 = \beta^*_2 = 0$. Namely, it then says that 0 is a median of the set of numbers $z(2) = +1$ and $z(3) = 0$, with weight 1 each, and also that 0 is a median of the set of numbers $z(1) = -1$ and $z(3) = 0$, with weight 1 each. Which is true. Nevertheless, the minimum in (2.1) is not achieved at $\beta = (0, 0)$. For, the L_1-norm (equal to 2) of $\{z(x)\}$ can be reduced to 0 since $z(x) - g_1(x) + g_2(x) = 0$ for all $x \in X$. Essentially, one is confronted here with the simple function

$$f(u, v) = |u - 1| + |v + 1| + |u + v|.$$

It is convex and satisfies $f(0, 0) \leq f(u, 0)$, for all u, and $f(0, 0) \leq f(0, v)$, for all v. Nevertheless, $f(1, -1) = 0 < 2 = f(0, 0)$.

4. Median Polish. Tukey (1977) developed in detail the idea of calculating a reasonably good additive approximation to a given n-way layout of observations by so-called median polish. For a two-way layout as in (2.2), this additive approximatioon $\alpha_i + \beta_j$ to $z^{(o)}_{ijk}$ is derived as follows.

One starts with the 'matrix' $\mathbf{Z}^{(o)} = (z^{(o)}_{ijk})$ and applies a (median) row polish, yielding the matrix

$$\mathbf{Z}^{(1)}_{ijk} = (\mathbf{z}^{(1)}_{ijk} = z^{(o)}_{ijk} - \beta^{(1)}_i).$$

Here, the adjustment $\beta^{(1)}_i$ of the i-th row is taken as a fixed median (often the mid-median) of the set of numbers $z^{(o)}_{ijk}$ in i-th row (i fixed). Next, one applies a column polish to $\mathbf{Z}^{(1)}$ yielding

$$\mathbf{Z}^{(2)} = (z^{(2)}_{ijk} = z^{(1)}_{ijk} - \beta^{(2)}_j).$$

Here, $\beta^{(2)}_j$ is a fixed median of the set of numbers $z^{(1)}_{ijk}$ with j fixed. Polishing the rows of $\mathbf{Z}^{(2)}$ one obtains $\mathbf{Z}^{(3)}$ and so on. In general.

(4.1) $z_{ijk}^{(2h+1)} = z_{ijk}^{(2h)} - \beta_i^{(2h+1)}; \quad z_{ijk}^{(2h+2)} = z_{ijk}^{(2h+1)} - \beta_i^{(2h+2)}, (h = 0,1, \ldots).$

As observed by Gabriel (1983), a more efficient and self-correcting procedure would be to keep the original matrix $\mathbf{Z}^{(o)}$ and store not the $\mathbf{Z}^{(g)}$ but only the current cumulative sums $a_i^{(h)} = \Sigma_{g=1}^{h} \beta_i^{(2g-1)}$ and $b_j^{(h)} = \Sigma_{g=1}^{h} \beta_j^{(2g)}$ of the row and column adjustments. For, these can be recursively computed: (i) $a_i^{(h)}$ is the median of the numbers $z_{ijk}^{(o)} - b_i^{(h-1)}$, ($i$ fixed; $b_j^{(o)} = 0$); (ii) $b_j^{(h)}$ is the median of the set of numbers $z_{ijk}^{(o)} - a_i^{(h)}$, ($j$ fixed). Here, one typically uses the mid-median.

One continues (4.1) or the latter process until either convergencew is nearly obtained or else the norm of the matrix $Z^{(g)}$ seems to have reached its minimal value.

GENERALIZED POLISH. We now introduce the following analogous calculation for the general minimization problem (2.1). Here $p \geq 1$.

Start out by selecting a fixed infinite sequence of integers $\{r_n; n \geq 1\}$ taking values in $\{1, \ldots, M\}$, with the property that there exists a possibly large integer N, such that, for all $n \geq n_o$, each $r = 1, \ldots, M$ occurs in the finite subsequence $\{r_n, r_{n+1}, \ldots, r_{n+N}\}$. For instance, if $r_n \equiv n \pmod{M}$ then $N = M - 1$.

Now calculate recursively, for $n = 1, 2, \ldots,$

(4.2) $$z^{(n)}(x) = z^{(n-1)}(x) - s_n g_{r_n}(x), \qquad (x \in X)$$

where

(4.3) $$s_n = \text{mean}_{p,r}\{z^{(n-1)}(x)\},$$

(as defined in (3.4)). Therefore,

(4.4) $$\Sigma_{x \in X} \nu(x) |z^{(n)}(x)|^p = \min_s \Sigma_{x \in X} \nu(x) |z^{(n-1)}(x) - s g_{r_n}(x)|^p,$$

while s_n attains the latter minimum. If $p = 1$, this means that s_n is a median (to be specified) of the numbers $z^{(n-1)}(x)/g_{r_n}(x)$ with $g_{r_n} \neq 0$, relative to the weights $\omega(x)|g_{r_n}(x)|$.

As an illustration, in example (2.4) each index r_n is of the form $r_n = (t_n, y_n)$ with $t_n \in T$ and $y_n \in Y_{t_n}$. In the above generalization, the numbers $z^{(n)}(x)$ would be derived from the $z^{(n-1)}(x)$ by adjusting only the numbers with $x \in L_{t_n}(y_n)$. For instance, if $p = 1$ and $\omega(x) = 1$ then one substracts from the $z^{(n-1)}(x)$ with $x \in L_{t_n}(y_n)$ a fixed median of this same set of numbers.

In the remaining part of the present section, we consider the general minimization problem (2.1) with $p > 1$. Let $\beta^* = (\beta^*_1, \ldots, \beta^*_M)$ denote the unique vector attaining the minimum (2.1), see Lemma 1, and let

(4.5) $$\mathbf{Z}^* = \{z^*(x) = z^{(o)}(x) - \Sigma_{r=1}^{M} \beta^*_r g_r(x); x \in X\}$$

denote the corresponding optimal 'matrix' of residuals, the unique adjustment having the smallest possible L_p-norm.

THEOREM 1. *Suppose $p > 1$. Then the sequence $\mathbf{Z}^{(n)} = \{z^{(n)}(x); x \in X\}$ $(n = 0, 1, 2, \ldots)$ defined by the above generalized polish always converges to the optimal matrix $\mathbf{Z}^*$.*

Remark. The special case $p = 2$ can also be deduced from results due to Amemiya and Ando (1965) concerning projections in a Hilbert space. Smith, Solmon and Wagner ((1977) p. 1229) even established an exponential rate of convergence when $p = 2$ and $\{r_n\}$ is periodic.

Note from (4.2) that $\mathbf{Z}^{(n)} = \{z^{(n)}(x); x \in X\}$ is of the form

(4.6) $$z^{(n)}(x) = z^{(o)}(x) - \Sigma_{r=1}^{M} \beta_{n,r} g_r(x).$$

Here, the $\beta_{n,r}$ are unique since $g_1, \ldots, g_M$ are linearly independent. Moreover, from (4.4),

in going from the $z^{(n-1)}(x)$ to the $z^{(n)}(x)$, one takes $\beta_{n,r} = \beta_{n-1,r}$ for all $r \neq r_n$, while β_{n,r_n} is chosen so as to minimize the norm of $\mathbf{Z}^{(n)}$.

A GENERALIZATION. One can even replace the linear combination $\beta_1 g_1(x) + \ldots + \beta_M g_M(x)$ by a general continuous function $g(x; \beta_1, \beta_2, \ldots, \beta_M)$, thus the residuals take the form

$$d(x) = d(x; \beta_1, \ldots, \beta_M) = z^{(o)}(x) - g(x; \beta_1, \ldots, \beta_M).$$

The main goal would normally be to minimize some explicitly given expression in terms of these residuals, for instance, of the type $\Sigma_{x \in X} \omega(x)|d(x)|^p$. It measures the 'norm' or 'size' of the present set of residuals. Keeping the data $z^{(o)}(x)(x \in X)$ fixed, this expression becomes a known function $F(\beta_1, \ldots, \beta_M)$.

We will *assume* that this resulting function F is a *strictly convex* function on R^M of class C^1 and such that $F(\beta)$ tends to $+\infty$ as $\beta = (\beta_1, \ldots, \beta_M)$ tends to infinity. In the above special case with $p > 1$, the function F would be given by the right hand side of (2.1) and does indeed have the above properties.

The generalized polish starts again with $\mathbf{Z}^{(o)} = \{z^{(o)}(x); x \in X\}$. After $n-1$ steps, one obtains a set of residuals

$$z^{(n-1)}(x) = z^{(o)}(x) - g(x; \beta_{n-1,1}, \ldots, \beta_{n-1,M}) \tag{4.7}$$

$(x \in X)$ and having its 'size' equal to $F(\beta_{n-1,1}, \ldots, \beta_{n-1,M})$. We now derive that $z^{(n)}(x)$ from the $z^{(n-1)}(x)$ by choosing $\beta_{n,r} = \beta_{n-1,r}$ for all $r \neq r_n$ and choosing β_{n,r_n} such that $F(\beta_{n,1}, \ldots, \beta_{n,M})$ is as small as possible.

It follows from the above assumptions that $\beta_{n,r}$ is uniquely determined. We will assume as before that there exists a positive integer N such that, for all $n \geq n_o$, each $r = 1, \ldots, M$ occurs among $\{r_n, r_{n+1}, \ldots, r_{n+N}\}$. The following result generalizes Theorem 1. Here $b_n = (\beta_{n,1}, \ldots, \beta_{n,N})$, $(b_o = 0)$.

THEOREM 2. *The sequence $\{b_n\}$ converges to the unique point $\beta^* = (\beta^*_1, \ldots, \beta^*_M)$ where F assumes its smallest value. Hence, for each $x \in X$, $z^{(n)}(x)$ converges to*

$$z^*(x) = z^{(o)}(x) - g(x; \beta^*_1, \ldots, \beta^*_M). \tag{4.8}$$

Proof. It follows from the properties of F and the choice of β_{n,r_n} that, for all $n \geq 1$,

$$(\delta/\delta\beta_{r_n})F(\beta) = 0 \quad \text{at } \beta = b_n, \tag{4.9}$$

and further

$$F(b_n) \leq F((b_{n-1} + b_n)/2) \leq F(b_{n-1}). \tag{4.10}$$

Hence, $\lim_n F(b_n)$ exists and all the b_n belong to the compact set $K = \{\beta \in R^M: F(\beta) \leq F(b_o)\}$. We further claim that

$$\lim_n (b_n - b_{n-1}) = 0. \tag{4.11}$$

For, suppose not. Then there would exist integers $1 < n_1 < n_2 < \ldots$ such that $b_{n_k} \to u$ and $b_{n_k-1} \to v$ as $k \to \infty$, with $u, v \in K$ and $u \neq v$. Then $F(u) = \lim_n F(b_n) = F(v)$ while, from (4.10), $F(u) \leq F((u+v)/2) \leq F(v)$. Hence, all the equality signs hold here which contradicts the strict convexity of F.

Let $\{b_{n_k}\}$ be a convergent subsequence of $\{b_n\}$ with limit b^*. From (4.11), one also has for each fixed m that $\lim_k b_{nk+m} = b^*$. And we conclude from (4.9) that

(4.12) $$(\delta/\delta\beta_r)F(\beta) = 0 \quad (r = 1, \ldots, M) \quad \text{at } \beta = b^*.$$

After all, F is of class C^1 while, for each k, one has that each $r = 1, \ldots, M$ occurs among $r_{n_k+m}(m = 0, 1, \ldots, N)$. However, property (4.12) uniquely determines the single point β^* where F takes its smallest value, thus $b^* = \beta^*$. This proves that $\{b_n\}$ converges to β^*. The last assertion follows from (4.7) and (4.8). □

5. Optimal Sign Patterns. In the remaining part of the paper, we restrict ourselves to the case $p = 1$ of the general minimization problem (2.1). Let $\mathbf{Z} = \{z(x); x \in X\}$ be a fixed 'matrix'; it usually arises as a matrix of residuals. We associate to $\mathbf{Z}$ the function

(5.1) $$F(\beta) = \Sigma_{x\in X}\omega(x)|z(x) - \Sigma_{r=1}^{M}\beta_r g_r(x)|,$$

where $\beta = (\beta_1, \ldots, \beta_M)$. Clearly, F is a piecewise linear, continuous and convex function on $\mathcal{R}^M$ tending to $-\infty$ as β tends to infinity. Let K_Z denote the non-empty compact polyhedral convex set of points $\beta \in \mathcal{R}^M$ where F assumes its smallest value.

The matrix $\mathbf{Z}$ will be said to be *optimal* if its norm cannot be reduced by substracting from $z(x)$ a linear combination of the $g_r(x)$ $(r = 1, \ldots, M)$. Equivalently, $0 \in K_Z$. Similarly,

(5.2) $$Z_\beta = z_\beta(x) - \Sigma_r\beta_r g_r(x); x \in X$$

is optimal if and only if $\beta \in K_Z$. It is known, see Kennedy and Gentle (1980) p. 515, and can easily be proved by an induction on M, that at least one of these optimal matrices $\mathbf{Z}_\beta$ has M or more of its components $z_\beta(x)$ equal to 0.

We will derive several different criteria which are necessary and sufficient for $\mathbf{Z}$ to be optimal. An important role is played by the set

(5.4) $$D = \{x \in X: z(x) = 0\} = \{x \in X: \zeta(x) = 1\}$$

of zero locations. Here, $\zeta(x) = 1 - |\sigma(x)|$ with

(5.4) $$\sigma(x) = \operatorname{sgn} z(x), \qquad (x \in X).$$

We further associate to $\mathbf{Z}$ the set of M constants

(5.5) $$u_r = \Sigma_{x\in X}\theta(z(x))\omega(x)g_r(x), \qquad (r = 1, \ldots, M).$$

Here, $\theta(z(x)) = \sigma(x) + \zeta(x)$ equals -1 if $z(x)$ is negative and $+1$, otherwise. We always define $v_+ = \max(0, v)$, $v_- = \max(0, -v)$.

LEMMA 2. *In order that* $\mathbf{Z} = \{z(x); x \in X\}$ *be optimal, it is necessary and sufficient that*

(5.6) $$u_1\beta_1 + \ldots + u_M\beta_M \leq \Sigma_{x\in D}2\omega(x)[\beta_1 g_1(x) + \ldots + \beta_M g_M(x)]_+$$

holds for each choice of the real constants $\beta_1, \ldots, \beta_m$. *In fact, given* $\beta \in \mathcal{R}^M$, *(5.6) fails to hold if and only if* $F(\lambda\beta) < F(0)$ *for all sufficiently small scalars* $\lambda > 0$.

Proof. Optimality of $\mathbf{Z}$ means that the associated convex function F defined by (5.1) has the origin as a local minimum along *each* half line through the origin of $\mathcal{R}^M$; for, this implies global minimality. Equivalently, one must have for all β that

(5.7) $$\lim_{\lambda\downarrow 0}(F(\lambda\beta) - F(0))/\lambda \geq 0.$$

Given $\beta \in R^M$, we have from (5.1) that (5.7) is equivalent to

(5.8) $$\Sigma_{x\in D}\sigma(x)\omega(x)\,\Sigma_{r=1}^{M}\beta_r g_r(x) \leq \Sigma_{x\in D}\omega(x)|\Sigma_{r=1}^{M}\beta_r g_r(x)|.$$

Adding $\Sigma_{x\in D}\omega(x)\Sigma_{r=1}^{M}\beta_r g_r(x)$ to both sides of (5.8), one obtains condition (5.6). This proves Lemma 2. □

COROLLARY. Whether or not the matrix $\mathbf{Z} = \{z(x); x \in X\}$ is optimal depends only on the associated sign pattern $\{\sigma(x); x \in X\}$ and not on the values $z(x)$ themselves. Therefore, we will also speak of optimal and non-optimal sign patterns.

A sufficient condition for optimality is that $u_r = 0$ for all r. If $z(x) \neq 0$ for all $x \in X$, (that is, if D is empty) then the latter condition is also necessary.

It is convenient to introduce the set

$$B = B(\sigma) = \{\beta: \Sigma_{x\in D}\omega(x)|\ \Sigma_{r=1}^{M}\beta_r g_r(x)| < \Sigma_{x\notin D}\sigma(x)\omega(x)\Sigma_{r=1}^{M}\beta_r g_r(x)\}, \tag{5.9}$$

where $\beta = (\beta_1, \ldots, \beta_M)$. It is the set for which the (equivalent) conditions (5.6), (5.7), (5.8) fail to hold. That is, B is precisely the set of direction in which F is strictly decreasing when starting at the origin. Also note that B is an open convex cone; (naturally, $0 \notin B$).

The set B depends only on the associated sign pattern. A sign pattern σ is optimal or not depending on whether $B(\sigma)$ is empty or non-empty, respectively.

It is useful to introduce the following (quasi) partial ordering among sign patterns over X. Namely, let us say that the sign pattern $\sigma = \{\sigma(x); x \in X\}$ is smaller than the sign pattern $\tau = \{\tau(x); x \in X\}$ (and we write $\sigma \prec \tau$) if either $\sigma = \tau$ or else τ can be obtained from σ by replacing one or more elements $\sigma(x) = 0$ by either -1 or $+1$. Such an ordering is clearly transitive.

If $\sigma \prec \tau$ then the lower sign pattern σ has a larger set D of zero locations. Moreover, condition (5.8) for τ is easily seen to imply the analogous condition for σ. Therefore,

$$\text{if} \quad \sigma \prec \tau \quad \text{then} \; B(\sigma) \prec B(\tau). \tag{5.10}$$

THEOREM 3. *If a sign pattern $\sigma(x); x \in X$ is optimal then it remains optimal when one or more non-zero elements ($-$ or $+1$) are replaced by 0; ('the more zeros the better').*

Proof. What is asserted is that $\sigma \prec \tau$, together with the optimality of τ, (that is, $B(\tau)$ is empty), implies the optimality of σ, (that is, $B(\sigma)$ is empty). And this is evident from (5.10).

Note that, relative to the partial ordering on hand, the optimal sign patterns form a lower set while the non-optimal patterns form the (complementary) upper set. In order to be able to recognize a non-optimal pattern, it would be sufficient to have a list of all *minimal* non-optimal patterns σ. For each such non-optimal σ, the set $B(\sigma)$ is non-empty and it would be useful to list not only σ itself but also one or more members β of the associated set $B(\sigma)$. Namely, we know from (5.10) that β also belongs to each set $B(\tau)$ with $\sigma \prec \tau$ and supplies a method for improving any matrix whose ± 1 pattern contains that of σ (ignoring zeros).

Example. Consider the problem of minimizing $\Sigma_{i=1}^{3}\Sigma_{j=1}^{3}|z_{ij}^{(o)} - \alpha_i - \beta_j|$. It is not hard to show that a set of residuals $\mathbf{Z} = (z_{ij})$ which is invariant under median polish (a so-called EMP) is non-optimal if and only if $\sigma = (\sigma_{ij} = \operatorname{sgn} z_{ij})$ shows a subpattern of the type $\sigma_{11} = +1$; $\sigma_{22} = +1$; $\sigma_{33} = -1$. More precisely, Z is non-optimal if and only if, for some permutation (j_1, j_2, j_3) of $(1, 2, 3)$ the values $\eta_i = \sigma_{i,j_i}$ $(i = 1, 2, 3)$ are all non-zero but not all of the same sign. If for instance $\sigma_{13} = -1$; $\sigma_{21} = +1$; $\sigma_{32} = -1$ then the norm of $\mathbf{Z}$ can be reduced by adding a small positive number ϵ to the second and third column, and simultaneously substracting ϵ from the second row.

Comments. A typical algorithm for solving (2.1) (with $p = 1$), that is, for minimizing a function F as in (5.1), would first check whether or not the matrix $\mathbf{Z}$ on hand is optimal. If it is not then one locates somehow an element $\beta \in B$ (known to be non-empty), that is, a direction in which F is strictly decreasing when starting at 0. One may as well proceed

in that same direction until a minimum of $F(\lambda\beta)$ is reached. This dictates the choice

$$\lambda^\circ = \text{med}\{z(x)/h(x): \omega(x)|h(x)|)\}, \text{ where } h(x) = \Sigma_{r=1}^{M}\beta_r g_r(x) \tag{5.11}$$

and leads to the new matrix

$$\mathbf{Z}' = \{z'(x) = z(x) - \Sigma_{i=1}^{M}\lambda^\circ\beta_r g_r(x); x \in X\}.$$

It has a strictly smaller norm. One next checks whether $\mathbf{Z}'$ is optimal and so on.

Median polish as described by (4.2), (4.3) (with $p = 1$) is of a similar type except that one only allows motions parallel to one of the m coordinate axes. We will call $\mathbf{Z}$ an endproduct of median polish (EMP) if a single motion of that type does not improve the norm, (though two subsequent motions of that type might). In view of (5.11), this is equivalent to

$$\text{med}\{z(x)/g_r(x): \omega(x)|g_r(x)|\} = 0, \qquad (r = 1, \ldots, M). \tag{5.12}$$

We will call $\mathbf{Z}$ and endproduct of mid-median polish (EMMP) if (5.12) holds with 'median' replaced by 'mid-median'. Condition (5.12) means precisely that the (equivalent) conditions (5.6), (5.7), (5.8) hold on choosing $\beta_r = +1$ or -1 and $\beta_s = 0$, otherwise, ($r = 1, \ldots, M$). Thus, by (5.6), (5.12) is equivalent to

$$-\Sigma_{x\in D}2\omega(x)g_r(x)_- \leq u_r \leq \Sigma_{x\in D}2\omega(x)g_r(x)_+, (r = 1, \ldots, M). \tag{5.13}$$

6. Additional Criteria for Optimality.

THEOREM 4. *In order that* $\mathbf{Z} = z(x); x \in X$ *be optimal, it is necessary and sufficient that numbers* $w(x)$ $(x \in X)$ *exist with*

$$\begin{aligned} w(x) &= \sigma(x)\omega(x) \quad \text{if } x \notin D; \\ |w(x)| &\leq \omega(x) \quad \text{if } x \in D, \end{aligned} \tag{6.1}$$

and further

$$\Sigma_{x\in X}w(x)g_r(x) = 0 \quad \text{for all } r = 1, \ldots, M. \tag{6.2}$$

An equivalent condition is that the following moment problem has a solution. Namely, there must exist numbers $W(x)$ $(x \in D)$ *with*

$$0 \leq W(x) \leq 2\omega(x) \quad \text{for each } x \in D \tag{6.3}$$

and

$$\Sigma_{x\in D}W(x)g_r(x) = u_r \quad \text{for all } r = 1, \ldots, M. \tag{6.4}$$

Here, D is defined by (5.4) ***and*** u_r ***is defined by*** (5.5).

Proof. That the two conditions are equivalent is seen by letting $W(x) = \omega(x) - w(x)$, for $x \in D$. From (6.1) and the definitions of $\sigma(x)$ and $\zeta(x)$, (6.2) can be written as

$$\Sigma_{x\in X}\sigma(x)\omega(x)g_r(x) + \Sigma_{x\in X}\zeta(x)(\omega(x) - W(x))g_r(x) = 0.$$

In view of (5.5), this is equivalent to (6.4). □

As is obvious and well-known, see Wagner (1959), the minimization problem (2.1) (with $p = 1$) can be regarded as a linear programming problem to

(I) Minimize: $\Sigma_{x\in X}\omega(x)(p(x) + q(x))$,

subject to the conditions

$$z(x) - \Sigma_{r=1}^{M}\beta_r g_r(x) = p(x) - q(x); p(x) \geq 0; q(x) \geq 0, (x \in X).$$

The variables β_r are real-valued.

In order that $\mathbf{Z}$ be optimal, it is necessary and sufficient that the minimum on hand be

equal to $\Sigma_{x\in X}\omega(x)|z(x)|$, (corresponding to $p(x) = z(x)_+$; $q(x) = z(x)_-$ and $\beta_r = 0$ as an optimal solution).

The dual of problem (I) is to

(II) Maximize: $\Sigma_{x\in X}z(x)w(x)$, subject to the conditions (6.2) and $-\omega(x) \leq w(x) \leq \omega(x)$.

Each problem has feasible solutions. Thus, optimal solutions exist for each, and the minimum in (I) equals the maximum in (II). Let $w(x)$ $(x \in X)$ be a feasible solution of (II). Then

$$\Sigma_{x\in X}z(x)w(x) \leq \Sigma_{x\in X}|w(x)z(x)| \leq \Sigma_{x\in X}\omega(x)|z(x)|.$$

In order that $\mathbf{Z}$ be optimal and, simultaneously, $w(x)$ be optimal for (II), it is necessary and sufficient that the equality signs hold here. This means that $w(x) = \sigma(x)\omega(x)$ each time that $z(x) \neq 0$. Since optimal solutions for (II) always exist, this proves Theorem 4. □

Remark 1. Note that the conditions (6.1), (6.2) depend only on the sign pattern $\sigma = \{\sigma(x); x \in X\}$ and that it becomes weaker (less demanding) on replacing some elements $+1$ or -1 by 0. This yields a second proof of Theorem 3.

Remark 2. An other proof of Theorem 4 would be as follows. Consider the moment problem (6.3), (6.4). It requires the existence of a (nonnegative) measure μ on D satisfying

$$\int g_r(\xi)\,\mu(d\xi) = u_r; \quad \int \delta_\xi^x\,\mu(d\xi) \leq 2\omega(x);$$

$(r = 1, \ldots, M; x \in D)$. As is well-known, see Kemperman (1983), since D is finite such a measure exists if and only if $\rho(x) \geq 0$ $(x \in D)$ and

$$\Sigma_{r=1}^M \beta_r g_r(\xi) \leq \Sigma_{x\in D}\rho(x)\delta_\xi^x\ (\xi \in D) \text{ imply } \Sigma_{r=1}^M \beta_r u_r \leq \Sigma_{x\in D}2\rho(x)\omega(x).$$

One might as well choose $\rho(x) = [\beta_1 g_1(x) + \ldots + \beta_M g_m(x)]_+$ and then one obtains exactly condition (5.6) of Lemma 2, which is indeed equivalent to the optimality of $\mathbf{Z}$.

7. Optimality for a General Layout. Let us now apply the above results to the general layout problem as in (2.4). For simplicity, we assume that $\omega(x) = 1$ $(x \in X)$, thus, one is interested in minimizing

$$S = \Sigma_{x\in X}|z^{(o)}(x) - \Sigma_{t\in T}\beta_t(\phi_t(x))| \tag{7.1}$$

by a suitable choice of the regression parameters $\beta_t(y)$. Note that the role of the index $r = 1, \ldots, M$ is presently taken over by the pairs (t, y) with $t \in T$ and $y \in Y_t$. The total number M of such pairs is often large. The function $g_r = g_{t,y}$ on X is presently as in (2.5).

Let $\mathbf{Z} = \{z(x); x \in X\}$ be a fixed matrix, usually arising as a sequence of residuals

$$z(x) = z^{(o)}(x) - \Sigma_{t\in T}\beta_t(\phi_t(x)) \qquad (x \in X), \tag{7.2}$$

with $\mathbf{Z}^{(o)} = z^{(o)}(x); x \in X)$ as the original data. We like to know in how far $\mathbf{Z}$ is optimal.

The number of elements x in the layer $L_t(y) = \{x \in X: \phi_t(x) = y\}$ will be denoted as $n_t(y)$. Let further $n_t^+(y)$, $n_t^o(y)$ and $n_t^-(y)$, respectively, denote the number of elements $x \in L_t(y)$ such that $\sigma(x) = \operatorname{sgn} z(x) = +1, 0$ or -1 respectively. Put

$$u_t(y) = n_t^+(y) + n_t^o(y) - n_t^-(y) = n_t(y) - 2n_t^-(y), \tag{7.3}$$

$(t \in T; y \in Y_t)$. As usual, $D = \{x \in X: z(x) = 0\}$. Theorem 4 yields the following two criteria for optimality.

Criterion 1. In order that $\mathbf{Z}$ be optimal, it is necessary and sufficient that there exist numbers $w(x)$ $(x \in X)$ such that

$$\Sigma_{\phi_t(x)=y}\, w(x) = 0, \quad \text{for all } t \in T; \text{ all } y \in Y_t, \tag{7.4}$$

and

$$w(x) = +1 \quad \text{if } z(x) > 0;$$
$$w(x) = -1 \quad \text{if } z(x) < 0; \tag{7.5}$$
$$-1 \leqslant w(x) = \leqslant +1 \quad \text{if } z(x) = 0.$$

Remark. In certain cases, such as in the two-way layout, one may add the condition that $w(x) \in \{-1, 0, +1\}$ when $x \in D$. If moreover $z(x)$ is of the form (7.2) and the original data $z^{(o)}(x)$ are integers then it follows that $\mathbf{Z}$ has an integral L_1-norm as soon as it is optimal; (thus, a residual matrix with non-integral norm cannot be optimal).

Namely, as follows from the proof of Theorem 4, (7.4) and (7.5) imply that the norm of $\mathbf{Z}$ is equal to

$$\Sigma_{x\in X} w(x)z(x) = \Sigma_{x\in X} w(x)z^{(o)}(x).$$

Criterion II. In order that $\mathbf{Z}$ be optimal, it is necessary and sufficient that there exist numbers $W(x)$ $(x \in D)$ such that

$$\Sigma\{W(x);\ x \in D; \phi_t(x) = y\} = u_t(y) \quad \text{if } t \in T; y \in Y_t \tag{7.6}$$

and that further $0 \leqslant W(x) \leqslant 2$ for each $x \in D$.

One may describe condition II as requiring the existence of a measure μ on D, having at most a mass 2 at each point of D, and such that, for each $t \in T$, the ϕ_t-projection of μ onto Y_t is precisely equal to the signed measure μ_t on Y having a mass $u_t(y)$ at each $y \in Y_t$. Note that the total algebraic mass of μ_t equals

$$Q = \Sigma_{y\in Y_t} u_t(y) = N^+ + N^o - N^- = N - 2N^-, \tag{7.7}$$

which is independent of $t \in T$. Here, N denotes the number of elements in X while N^+, N^o, N^-, respectively, denotes the number of $x \in X$ with $\sigma(x) = \operatorname{sgn} z(x) = +1$, 0 or -1, respectively.

Obviously, the required measure μ can only exist when

$$0 \leqslant u_t(x) \leqslant 2n_t^o(y), \quad (t \in T; y \in Y_t). \tag{7.8}$$

This is equivalent to

$$n_t^-(y) \leqslant n_t(y)/2 \quad \text{and} \quad n_t^+(y) \leqslant n_t(y)/2, (t \in T; y \in Y_t). \tag{7.9}$$

In fact, (7.9) is precisely the condition the $\mathbf{Z}$ be an EMP. Equivalently, that for4 all $t \in T$ and $y \in Y_t$ the set of $n_t(y)$ numbers $z(x)$ with $x \in L_t(y)$ (each with weight 1) has 0 as a median. Which is exactly the condition which median polish tries to attain.

It is very easy to check condition (7.9). Thus, our main problem is to decide whether $\mathbf{Z}$ is optimal in a situation where (7.9) is true. In particular, the above given marginal measures μ_t are all nonnegative.

Lemma 2 easily yields the following criterion.

Criterion III. In order that $\mathbf{Z}$ be optimal, it is necessary and sufficient that

$$\Sigma_{t\in T}\Sigma_{y\in Y_t} u_t(y)\beta_t(y) \leqslant \Sigma_{x\in D} 2[\Sigma_{t\in T}\beta_t(\phi_t(x))]_+ \tag{7.10}$$

holds for each choice of M real numbers $\beta_t(y)$, $(t \in T; y \in Y_t)$.

The set B defined by (5.9) presently takes the form

$$B = \{\beta: \Sigma_{x\in D}|\Sigma_{t\in T}\beta_t(\phi_t(x))| < \Sigma_{x\notin D}\sigma(x)\,\Sigma_{t\in T}\beta_t(\phi_t(x))\}. \tag{7.11}$$

Here, β stands for the set of M numbers $\beta_t(y)$ $(t \in T; y \in Y_t)$. The matrix $\mathbf{Z}$ is optimal if and only if B is empty. If B is non-empty then each $\beta \in B$ supplies an explicit way of reducing the norm of $\mathbf{Z}$, see (5.11).

8. Optimality for a Two-Way Layout. Here, we only consider the case of a two-way

$m \times n$ layout ($m \geq 2$; $n \geq 2$) with a single observation $z^{(o)}(x) = z^{(o)}_{ij}$ in cell $x = (i, j)$ and weights $\omega(x) = 1$. One wants to minimize

$$S = \Sigma_{i,j}|z^{(o)}_{ij} - \alpha_i - \beta_j| \tag{8.1}$$

by a suitable choice of the $M = m + n$ numbers $\alpha_i = \beta_1(i)$ and $\beta_j = \beta_2(j)$.

Unspecified indices i and j run through $Y_1 = \{1, \ldots, m\}$ and $Y_2 = \{1, \ldots, n\}$, respectively. Presently, we have $X = Y_1 \times Y_2$ while $T = \{1, 2\}$ and $\phi_1(x) = i$; $\phi_2(x) = j$ when $x = (i, j)$.

Having chosen the numbers α_i and β_j (at a particular stage of the calculation), one is confronted with the problem whether or not the matrix $\mathbf{Z}$ of residuals is optimal, in the sense that its norm cannot be further reduced.

Let $\mathbf{Z} = (z_{ij})$ be a fixed $m \times n$ matrix. Whether or not $\mathbf{Z}$ is optimal depends only on the sign pattern $\sigma = (\sigma_{ij})$, where $\sigma_{ij} = \operatorname{sgn} z_{ij}$. The number of elements $\sigma_{ij} = -1, 0, +1$, respectively, in the i-th row of σ will be denoted as $n^-_1(i)$, $n^o_1(i)$, $n^+_1(i)$, respectively; similarly, $n^-_2(j)$, $n^o_2(j)$, $n^+_2(j)$ for the j-th column of σ.

Definition. The matrix $\mathbf{Z}$ is called an EMP (or EMMP) if 0 is a median (or mid-median, respectively) of each row and each column of $\mathbf{Z}$. Each EMMP is an EMP. In order that $\mathbf{Z}$ be an EMP it is clearly necessary and sufficient that

$$\max(n^-_1(i), n^+_1(i)) \leq n/2; \quad \max(n^-_2(j), n^+_2(j)) \leq m/2, \tag{8.2}$$

for all i and j.

An EMMP cannot have exactly one zero in a row unless n is odd. In fact, an EMMP has, for each fixed i, that $n^-_1(i) = n/2$ if and only if $n^+_1(i) = n/2$ and, similarly, for each fixed j, $n^+_2(j) = m/2$ if and only if $n^-_2(j) = m/2$. An EMP with the latter property will be called a weak EMMP.

A necessary condtion for $\mathbf{Z} = (z_{ij})$ to be optimal is that it be an EMP. Though numerical calculations suggest it, we are not asserting that a median polish (mid-median polish) always leads to an EMP (EMMP). For the case where either m or n is odd, it is not difficult to show that mid-median polish does create a convergent sequence $\mathbf{Z}^{(n)} = (z^{(n)}_{ij})$ of matrices whose limit is an EMMP.

Anyway, at the end of a median polish one is typically confronted with a matrix $\mathbf{Z} = (z_{ij})$ of residuals which already is an EMP or EMMP and then the question arises how one can recognize its optimality. And if this EMP is non-optimal (as is often true) then how should one proceed in determining an optimal matrix of residuals?

CONDITION (A, B). Let $A \subset Y_1$ and $B \subset Y_2$. Let $\mathbf{G}$ and $\mathbf{H}$ denote the submatrices of $\mathbf{Z}$ defined by

$$\mathbf{G} = (z_{ij}; i \in A, j \notin B); \quad \mathbf{H} = (z_{ij}; i \notin A, j \in B). \tag{8.3}$$

The number of positive, zero and negative elements in $\mathbf{G}$ will be denoted as $N^+_{\mathbf{G}}$, $N^o_{\mathbf{G}}$ and $N^-_{\mathbf{G}}$, respectively, while $N_{\mathbf{G}}$ denotes the total number of elements. Similarly, $N^+_{\mathbf{H}}$, $N^o_{\mathbf{H}}$ and $N^-_{\mathbf{H}}$ and $N_{\mathbf{H}}$ for $\mathbf{H}$. We will say that $\mathbf{Z}$ satisfies Condition (A, B) if

$$N^+_{\mathbf{G}} + N^-_{\mathbf{H}} \leq (N_{\mathbf{G}} + N_{\mathbf{H}})/2. \tag{8.4}$$

Note that Condition (A^c, B^c) requires that $N^-_{\mathbf{G}} + N^+_{\mathbf{H}} \leq (N_{\mathbf{G}} + N_{\mathbf{H}})/2$.

THEOREM 5. *In order that $\mathbf{Z} = (z_{ij})$ be optimal, it is necessary and sufficient that Condtion (A, B) holds for each choice of the subset A of Y_1 and subset B of Y_2.*

Proof. The sufficiency follows from Theorem 8 in Section 9. As to the necessity of (8.4), consider the modified matrix

$$\mathbf{Z}' = (z'_{ij} = z_{ij} - \alpha_i - \beta_j),$$

where $\alpha_i = +\lambda$ when $i \in A$; $\alpha_i = 0$ when $i \notin A$, while $\beta_j = -\lambda$ when $j \in B$; $\beta_j = 0$ when $j \notin B$. Here, λ denotes a sufficiently small positive constant. The effect of this transformation is that each element z_{ij} in $\mathbf{G}$ is decreased by λ, each element in $\mathbf{H}$ is increased by λ, while the remaining part of $\mathbf{Z}$ remains unchanged. This causes on the one hand a decrease in norm by $(N^+_{\mathbf{G}} + N^-_{\mathbf{G}})\lambda$ and on the other hand an increase in norm by

$$(N^-_{\mathbf{G}} + N^o_{\mathbf{G}} + N^+_{\mathbf{H}} + N^o_{\mathbf{H}})\lambda = (N_{\mathbf{G}} + N_{\mathbf{H}})\lambda - (N^+_{\mathbf{G}} + N^-_{\mathbf{H}})\lambda.$$

Hence, unless (8.4) holds the matrix $\mathbf{Z}'$ would have a strictly smaller norm than $\mathbf{Z}$ and $\mathbf{Z}$ would not be optimal. □

Remark 1. In order that $\mathbf{Z}$ be an EMP it is necessary and sufficient that condition (A, B) holds with one of the sets A, B empty and the other consisting either of a single element or else all but a single element. This in turn is equivalent to Condition (A, B) for the case where one of the two sets A, B is either empty or full. Thus, if $\mathbf{Z}$ is already known to be an EMP then one only needs to verify (8.4) for the case where neither $\mathbf{G}$ nor $\mathbf{H}$ is empty.

Remark 2. Theorem 5 suggests the following algorithm for determining an optimal set of residuals. After n steps one has a matrix $\mathbf{Z}^{(n)}$. If it satisfies all conditions (A, B) then it is optimal. If not then the above proof indicates how to arrive at a new matrix $\mathbf{Z}^{(n+1)}$ having a strictly smaller norm. It is best to choose $\lambda = \lambda_n$ in an optimal way, namely, as a median of the $N_{\mathbf{G}} + N_{\mathbf{H}}$ elements g_{ij} and $-h_{ij}$ in $\mathbf{G}$ and $-\mathbf{H}$. If the original data $z^{(o)}_{ij}$ are all integers then one can attain that, for all n, also $z^{(n)}_{ij}$ and λ_n and thus $z^{(n+1)}_{ij}$ are integers. But then the norm of the matrix decreases at each step by a positive integer, hence, the process must stop after finitely many steps.

Example. Armstrong, Elam and Hultz (1977) developed a quite different algorithm. Details were given for the following 4×5 matrix

$$\mathbf{Z}^{(o)} = \begin{pmatrix} 350 & 492 & 232 & 220 & 360 \\ 392 & 428 & 253 & 241 & 385 \\ 400 & 498 & 273 & 260 & 401 \\ 320 & 390 & 264 & 240 & 300 \end{pmatrix}.$$

Their method led to the following matrix of residuals

$$\mathbf{Z} = \begin{pmatrix} 0 & 72 & 0 & 0 & 0 \\ 21 & -13 & 0 & 0 & 4 \\ 9 & 37 & 0 & -1 & 0 \\ 0 & 0 & 62 & 50 & -30 \end{pmatrix}.$$

which has norm 299 and was claimed to be optimal. Note that $\mathbf{Z}$ is an EMP but not an EMMP. Actually, $\mathbf{Z}$ does not satisfy Condition (A, B) with $A = 1, 2, 3$ and $B = 3, 4, 5$. For, then $N_{\mathbf{G}} = 6$; $N^+_{\mathbf{G}} = 4$; $N_{\mathbf{H}} = 3$; $N^-_{\mathbf{H}} = 1$ so that (8.4) is violated. This allows an improvment as usual; it is best to choose $\lambda = 9$. In this way, substracting 9 from the first 3 rows and adding 9 to the last 3 columns, one arrives at the matrix

$$\mathbf{Z}' = \begin{pmatrix} -9 & +63 & 0 & 0 & 0 \\ +12 & -22 & 0 & 0 & +4 \\ 0 & +28 & 0 & -1 & 0 \\ 0 & 0 & +71 & +59 & -21 \end{pmatrix}.$$

which has norm 290. Using any of several criteria in Sections 8 and 9, it is easily seen that $\mathbf{Z}'$ is optimal. For instance, substracting 14 from each element in the second column, one obtains a matrix $\mathbf{Z}''$ which has the same norm is is an EMMP and, thus, optimal by Theorem 6 below. This implies that also $\mathbf{Z}'$ is optimal.

Definition. The pair (m, n) of integers ≥ 2 is safe for median polish if each EMP of dimension (m, n) is optimal. Similarly, (m, n) is safe for mid-median polish if each EMMP of dimension (m, n) is optimal.

THEOREM 6. *No pair (m, n) is safe for median polish. And further the only pairs which are safe for mid-median polish are the exceptional pairs* $(2, n)$; $(3, 4)$; $(4, 4)$; $(4, 5)$; $(4, 6)$ *and their reflections such as* $(n, 2)$.

In fact, for these exceptional pairs (m, n) it is even true that every weak EMMP of dimension (m, n) is optimal.

Remark. In particular, the pairs (3, 3); (3, 5); (6, 6); (4, 8) are all unsafe for mid-median polish. A weak EMMP may be described as a matrix whose sign pattern (having only elements -1, 0, $+1$) is exactly an EMMP. Thus, its sign pattern indicates an EMMP but the matrix may not have the property that 0 is exactly the mid-median of each row and each column. But indeed we know from Section 5 that the sign pattern alone already determines optimality or nonoptimality.

Proof. In order to prove the stated 'unsafety', it suffices to construct an EMP or EMMP which violates one of the conditions (8.4) and, hence, is not optimal. Consider an $m \times n$ matrix $\mathbf{Z}$ which after a suitable permutation of rows and columns takes the form

$$\mathbf{Z} = (z_{ij}) = \begin{pmatrix} \mathbf{G} & \mathbf{K} \\ \mathbf{L} & \mathbf{H} \end{pmatrix} \tag{8.5}$$

Here, $\mathbf{G}$ and $\mathbf{H}$ are of dimension $m_1 \times n_1$ and $m_2 \times n_2$, respectively, $(m_1 + m_2 = m;\ n_1 + n_2 = n)$, while $\mathbf{K}$ and $\mathbf{L}$ are *zero* matrices of dimension $m_1 \times n_2$ and $m_2 \times n_1$, respectively. From Theorem 5, the matrix $\mathbf{Z}$ is non-optimal as soon as (8.4) is false.

In fact, we will choose all the elements of $\mathbf{G}$ as (strictly) positive. Let further the elements of $\mathbf{H}$ be either negative or 0, in an alternating (checkerboard type) fashion, starting with a negative element in the left upper corner of $\mathbf{H}$. In this situation one has $N_{\mathbf{G}}^{+} = N_{\mathbf{G}}$ and $N_{\mathbf{H}}^{-} \geq N_{\mathbf{H}}/2$, hence, the difference between the left and right hand sides of (8.4) is at least $N_{\mathbf{G}}/2 = m_1 n_1/2$. Thus, $\mathbf{Z}$ is non-optimal as soon as m_1 and n_1 are positive. Moreover, $\mathbf{Z}$ is easily seen to be an EMP provided

$$1 \leq m_1 \leq m_2; \quad 1 \leq n_1 \leq n_2. \tag{8.6}$$

That is, $1 \leq m_1 \leq m/2$ and $1 \leq n_1 \leq n/2$. Since such a choice of m_1 and n_1 is always possible, this proves the first assertion of Theorem 6. By the way, the leeway $m_1 n_1/2$ above indicates that there are usually many mild modifications of $\mathbf{Z}$ which are also non-optimal EMP's.

The matrix $\mathbf{Z}$ in (8.5) is even an EMMP provided

$$m_1 < m_2; \quad m_1 \geq 2 \text{ if } m_2 \text{ odd}; \quad n_1 < n_2; \quad n_1 \geq 2 \text{ if } n_2 \text{ odd}, \tag{8.7}$$

and further all non-zero elements z_{ij} equal -1 or 1.

For, then 0 is a mid-median of each row and each column; (if m_2 is odd then some columns of $\mathbf{H}$ have an excess of negative elements; hence, if $m_1 = 1$ then that column would have 0 as a median but not as a mid-median; similarly if n_2 is odd). Therefore, a sufficient cond-

tion for (m, n) to be 'unsafe' for mid-median polish is that (8.6) can be strengthened to (8.7).

Choosing $m_1 = n_1 = 2$, this is true if both $m \geqslant 5$ and $n \geqslant 5$. Letting $m_1 = n_1 = 1$, it also holds whem m and n are odd, $m \geqslant 3$ and $n \geqslant 3$. Letting $m_1 = 1$ and $n_1 = 2$, this approach also covers the pairs $(3, n)$ with $n \geqslant 5$. Similarly for the pairs $(m, 3)$ with $m \geqslant 5$.

It only remains to consider the pairs $(4, n)$ with $n \geq 7$, (the pairs $(n, 4)$ having the same character). Since one can enlarge the matrix $\mathbf{Z}$ by adding pairs of columns of the type $\left(\begin{smallmatrix} +1 & -1 & +1 & -1 \\ -1 & +1 & -1 & +1 \end{smallmatrix}\right)^T$, ($T$ for transpose), one easily sees that it suffices to construct for the (m, n) pairs (4, 7) and (4, 8) an EMMP of the type (8.5) with $m_1 = 2$ and $n_2 = 4$ and such that (8.4) is false; thus, we do not require that $\mathbf{L}$ and $\mathbf{K}$ are zero matrices. Examples with $m = 4, n = 7$ are:

$$\left(\begin{array}{ccc|cccc} + & + & 0 & + & 0 & 0 & 0 \\ 0 & 0 & + & + & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & - & - & - & 0 \\ 0 & 0 & 0 & - & 0 & 0 & - \end{array}\right) \qquad \left(\begin{array}{ccc|cccc} + & + & - & + & - & 0 & 0 \\ - & - & + & + & + & 0 & 0 \\ \hline + & 0 & 0 & - & - & - & + \\ - & 0 & 0 & - & + & + & - \end{array}\right)$$

where + and − may, for instance be interpreted as +1 and −1, respectively. Examples with $m = 4$ and $n = 8$ are:

$$\left(\begin{array}{ccc|ccccc} + & + & + & 0 & 0 & 0 & 0 & 0 \\ + & 0 & 0 & + & 0 & 0 & 0 & 0 \\ \hline - & 0 & 0 & 0 & - & - & 0 & 0 \\ - & 0 & 0 & 0 & 0 & - & - \end{array}\right) \qquad \left(\begin{array}{ccc|ccccc} + & + & + & - & - & 0 & 0 & 0 \\ + & - & - & + & + & 0 & 0 & 0 \\ \hline - & 0 & 0 & 0 & - & - & + & + \\ - & 0 & 0 & 0 & + & + & - & - \end{array}\right).$$

Finally, let $\mathbf{Z}$ be a weak EMMP of dimension (m, n). It only remains to show that $\mathbf{Z}$ is optimal for the special dimensions $(2, n)$; (3, 4); (4, 4); (4, 5) and (4, 6). This will be done by verifying that in these cases $\mathbf{Z}$ satisfies all conditions (A, B), see Theorem 5. Rearranging rows and columns, one may assume that the associated matrices $\mathbf{G}$ and $\mathbf{H}$ as in (8.3) are located as in (8.5). In view of Remark 1 following Theorem 5, one may assume that the m_t and n_t $(t = 1, 2)$ are positive integers. For brevity, let $p = N^+_{\mathbf{G}}$ and $q = N^-_{\mathbf{H}}$. It must be shown that

$$p + q \leqslant (N_{\mathbf{G}} + N_{\mathbf{H}})/2 = (m_1 n_1 + m_2 n_2)/2, \tag{8.8}$$

whenever $\mathbf{Z}$ is a weak EMMP having one of the above dimensions. Thus, $\mathbf{Z}$ is an EMP such that each row containing exactly $n/2$ positive (negative) elements has the property that all the other $n/2$ elements in that row are negative (positive); equivalently such a row contains no zeros. Similarly for columns.

The case $(2, n)$ is particularly easy. Here $m_1 = m_2 = 1$ while $N_{\mathbf{G}} = n_1$ and $N_{\mathbf{H}} = n_2$ with $n_1 + n_2 = n$. It must be shown that $p + q \leqslant n/2$, with p as the number of positive elements among the $z_{1,j}$ with $1 \leqslant j \leqslant n_1$ and q as the number of negative elements among the z_{2j} with $n_1 < j \leqslant n$. But z_{2j} is negative if and only if z_{1j} is positive. Hence, $p + q$ equals the number of positive elements in the first row and therefore $p + q \leqslant n/2$.

The case (4, 4) can be reduced to the case (6, 4), while the case (3, 4) can be reduced to the case (5, 4). Namely, if $\mathbf{Z}$ is a weak EMMP of dimension $(m, 4)$ then adding two rows of the type $\left(\begin{smallmatrix} + & - & + & - \\ - & + & - & + \end{smallmatrix}\right)$, one obtains a weak EMMP of dimension $(m + 2, 4)$. And if this new matrix $\mathbf{Z}'$ is optimal then so is the original matrix $\mathbf{Z}$. For, it is easily seen that condition (8.8) for $\mathbf{Z}$ follows from the analogous condition for $\mathbf{Z}'$; (the implication would also be an easy consequence of Criterion I of Section 7). It remains to show (8.8) for the cases (4, 5); (4, 6).

Let us do the case where $\mathbf{Z}$ is of dimension (4, 6). Consider for instance the situation that $m_1 = 2$; $m_2 = 2$; $n_1 = 3$ and $n_2 = 3$. It must be shown that $p + q \leqslant 6$. Suppose z_{11},

z_{12}, z_{13} are all positive, thus $p \geq 3$ and z_{14}, z_{15}, z_{16} must be negative, therefore, $q \leq 3$. If a column has two negative elements the other two must be positive. Hence, at least q of the elements z_{24}, z_{25}, z_{26} are positive, thus, at most $3 - q$ of the elements z_{21}, z_{22}, z_{23} are positive, showing that $p \leq 6 - q$.

A similar reasoning applies when z_{21}, z_{22}, z_{23} are all positive. Thus, one may assume that at most two of z_{11}, z_{12}, z_{13} are positive and at most two of z_{21}, z_{22}, z_{23} are positive, hence, $p \leq 4$. Similarly, one may assume that at most two of z_{34}, z_{35}, z_{36} are negative and at most two of z_{44}, z_{45}, z_{46} are negative, hence $q \leq 4$. One is ready if $p \leq 3$ and $q \leq 3$.

Suppose instead that for example $p = 4$. Typically, interchanging the first two rows or first three columns if necessary, the elements z_{11}, z_{12}, z_{21} are positive and further one of z_{22} or z_{23}. This forces z_{31} and z_{41} to be negative. If z_{22} were positive then also z_{32} and z_{42} would be negative and then $q \leq 2$. Thus, suppose instead that z_{23} is positive.

One is ready if each row of **H** contains at most one negative element, (for, then $q \leq 2$). If not then typically z_{44} and z_{45} are negative. This forces z_{42}, z_{43} and z_{46} to be positive which in turn forces z_{32} and z_{33} to be negative which in turn forces z_{34}, z_{35}, z_{36} to be positive and therefore that $q = 2$.

The above takes care of the (actually most difficult) case that **G** (and thus **H**) has dimension (2, 3). The other cases follow by a quite similar reasoning. Analogously for the case where the weak EMMP **Z** is of dimension (4, 5). We omit the details. This completes the proof of Theorem 6. □

9. Optimality for the General Two-Way Layout. Here, we study the case of a two-way $m \times n$ layout with weights $\omega(x) = 1$, this time allowing for several values z_{ijk} ($k = 1, \ldots, k_{ij}$) in cell (i, j) where $k_{ij} = 0$ is possible. This more general case becomes important in applications where one has a large number of observations and one likes to keep m and n relatively small so as to simplify the calculations. Unspecified indices i, j and k will run through Y_1, Y_2 and $\{1, \ldots, k_{ij}\}$, respectively.

Our problem is to minimize the sum (2.2). Thus, one needs to determine whether a given matrix of residuals

$$\mathbf{Z} = (z_{ijk} = z^{(O)}_{ijk} - \alpha_i - \beta_j)$$

is optimal. A necessary condition for **Z** to be optimal is that it be an EMP. Equivalently, that for each of the $M = m + n$ layers $L_1(i)$ and $L_2(j)$ the associated set of numbers z_{ijk} (with i fixed or j fixed) has 0 as a median. This is equivalent to

$$\max(n_1^-(i), n_1^+(i)) \leq n_1(i)/2; \quad \max(n_2^-(j), n_2^+(j)) \leq n_2(j)/2. \tag{9.1}$$

Here,

$$n_1(i) = \Sigma_{j=1}^{n} k_{ij}; \quad n_2(j) = \Sigma_{i=1}^{m} k_{ij} \tag{9.2}$$

is the number of elements in $L_1(i)$ and $L_2(j)$, respectively. Further,

$$n_1^-(i) = \Sigma_{j=1}^{n} k_{ij}^- \tag{9.3}$$

is the number of negative elements $z(x) = z_{ijk}$ with $x = (i, j, k)$ in layer $L_1(i)$, thus i is fixed. Similarly for $n_1^o(i)$, $n_1^+(i)$ and $n_2^-(j)$, $n_2^o(j)$, $n_2^+(j)$. Moreover, k_{ij}^-, k_{ij}^o and k_{ij}^+, respectively, denote the number of negative, zero and positive elements z_{ijk}, respectively, located in cell (i, j).

Analogously to (7.3) we define

$$u_1(i) = n_1(i) - 2n_1^-(i); \quad u_2(j) = n_2(j) - 2n_2^-(j). \tag{9.4}$$

The EMP property (9.1) is equivalent to

(9.5) $$0 \leq u_1(i) \leq 2n_1^o(i) = \Sigma_{j=1}^n 2k_{ij}^o;\ 0 \leq u_2(j) \leq 2n_2^o(j) = \Sigma_{i=1}^m 2k_{ij}^o.$$

THEOREM 7. *In order that* $\mathbf{Z} = (z_{ijk})$ *be optimal, it is necessary and sufficient that there exist numbers* W_{ij} *satisfying*

(9.6) $$0 \leq W_{ij} \leq 2k_{ij}^o;$$

and

(9.7) $$\Sigma_{j=1}^n W_{ij} = u_1(i); \quad \Sigma_{i=1}^m W_{ij} = u_2(j),$$

for all $i \in Y_1$ *and* $j \in Y_2$.

Proof. This result is a direct consequence of Criterion II in Section 7, applied to the set X of triplets $x = (i, j, k)$ while $z(x) = z_{ijk}$. Further put $W_{ij} = \Sigma_k W(i, j, k)$, where $W(x) = 0$ when $z(x) \neq 0$. □

Remark 1. Since k_{ij}^o and $u_t(y)$ are all integers one may even require that the W_{ij} are integers. Namely, the moment problem (9.6), (9.7) corresponds to the usual transportation problem which has a totally unimodular matrix, see Garfinkel and Nemhauser (1972) p. 73 and Hu (1969) p. 123. If all $n_1(i)$, $n_2(j)$ and thus the $u_t(y)$ are even then one can even attain that W_{ij} are even. This additional information may simplify the problem of deciding whether $\mathbf{Z}$ is optimal.

In view of the Remark following (7.5), we may conclude that an optimal matrix $\mathbf{Z}$ of residues necessarily has an integral norm provided all the original data $z_{ijk}^{(o)}$ were integers.

Remark 2. If one allows not only additive adjustments of the form $\alpha_i + \beta_j$ but also one or more additive adjustments of the form $\gamma_r g_r(i, j)$ (with the g_r as given functions and the γ_r as free constants) then optimality of $\mathbf{Z} = (z_{ijk})$ is equivalent to the existence of numbers W_{ij} satisfying (9.6), (9.7) and, moreover, the additional 'moment' conditions

(9.8) $$\Sigma_{i,j} W_{ij} g_r(i, j) = \Sigma_{i,j}(k_{ij} - 2k_{ij}^-) g_r(i, j).$$

CONDITION (A, B). Let A and B be subsets of $Y_1 = \{1, \ldots, m\}$ and $Y_2 = \{1, \ldots, n\}$, respectively. Given the matrix $\mathbf{Z}$, consider the associated arrays

(9.9) $$\mathbf{G} = (z_{ijk};\ i \in A, j \notin B); \quad \mathbf{H} = (z_{ijk};\ i \notin A, j \in B).$$

We will say that $\mathbf{Z}$ satisfies Conditions (A, B) if

(9.10) $$N_{\mathbf{G}}^+ + N_{\mathbf{H}}^- \leq (N_{\mathbf{G}} + N_{\mathbf{H}})/2.$$

Here,

$$N_{\mathbf{G}} = \Sigma_{i \in A} \Sigma_{j \notin B} k_{ij}; \quad N_{\mathbf{G}}^+ = \Sigma_{i \in A} \Sigma_{j \notin B} k_{ij}^+$$

denote the number of elements in $\mathbf{G}$ and the number of positive elements in $\mathbf{G}$, respectively. Similarly for $N_{\mathbf{H}}$ and $N_{\mathbf{H}}^-$.

THEOREM 8. *In order that* $\mathbf{Z} = (z_{ijk})$ *be optimal, it is necessary and sufficient that* $\mathbf{Z}$ *satisfies Condition* (A, B) *for each choice of the subsets* A *of* Y_1 *and* B *of* Y_2.

THEOREM 9. *In order that* $\mathbf{Z} = (z_{ijk})$ *be optimal, it is necessary and sufficient that the inequality*

(9.11) $$\Sigma_{i \in A}\, u_1(i) \leq \Sigma_{j=1}^n \min[u_2(j), \Sigma_{i \in A}\, 2k_{ij}^o]$$

holds for each subset A *of* Y_1.

Moreover, if (9.11) fails for a given subset A *of* Y_1 *then Condition* (A, B) *fails for the associated pair defined by*

(9.12) $$B = \{j \in Y_2:\ u_2(j) < \Sigma_{i \in A} 2k_{ij}^o\}.$$

And in that case the matrix $\mathbf{Z}$ *admits an easy improvement.*

Remark. What is meant here is the improvement

$$\mathbf{Z}' = (z'_{ijk} = z_{ijk} - \alpha_i - \beta_j)$$

with the α_i and β_j as in the proof of Theorem 5. Namely, choose $\alpha_i = +\lambda$ when $i \in A$; $\beta_j = -\lambda$ when $j \in B$ and $\alpha_i = 0$, $\beta_j = 0$, otherwise. The best choice for λ is a median of the $N_{\mathbf{G}} + N_{\mathbf{H}}$ numbers z_{ijk} in $\mathbf{G}$ and $-z_{ijk}$ in $-\mathbf{H}$. Note that this choice of λ depends on the full matrix $\mathbf{Z}$, not only on the associated sign pattern or the numbers k^+_{ij} and k^-_{ij}.

Proof of Theorems 8 and 9. For $t \in T = \{1, 2\}$, let μ_t be the (possibly signed) measure on Y_t having a mass $u_t(y)$ at $y \in Y_t$. Let further $q(\ .\ )$ denote the (nonnegative) measure on $Y_1 \times Y_2$ having a mass $2k^o_{ij}$ at the point (i, j). The criterion for optimality stated in Theorem 7 requires precisely that there exists a (nonnegative) measure μ on $Y_1 \times Y_2$ having marginals μ_1 and μ_2 and such that $\mu(E) \leqslant q(E)$ for every subset E of $Y_1 \times Y_2$. A necessary condition for the existence of μ is that

$$\mu_1(A) \leqslant q(A \times B^c) + \mu_2(B), \tag{9.13}$$

for every subset A of Y_1 and every subset B of Y_2. After all, $A \times Y_2 \subset (A \times B^c) \cup (Y_1 \times B)$ and $\mu(A \times Y_2) = \mu_1(A)$; $\mu(Y_1 \times B) = \mu_2(B)$; (taking A empty, this requires that $\mu_2 \geqslant 0$; similarly $\mu_1 \geqslant 0$ since $\mu_1(Y_1) = \mu_2(Y_2)$, see (7.7)).

As was shown by Dall'Aglio (1961) and Kellerer (1961), condition (9.13) is also sufficient for the existence of μ, hence, it is equivalent to the optimality of $\mathbf{Z}$. See Strassen (1965) p. 423 for generalizations and further references. The sufficiency of (9.13) is also an immediate consequence of the Ford-Fulkerson max-flow-min-cut theorem, see Ford and Fulkerson (1962), Berge (1970) and Jacobs (1978) p. 539.

For each fixed pair A and B, (9.13) is equivalent to Condition (A, B), proving Theorem 8. After all, using (9.2), (9.3), (9.4), the inequality (9.13) can be written as

$$\Sigma_{i\in A}\Sigma_j(k^+_{ij} + k^o_{ij} - k^-_{ij}) \leqslant \Sigma_{i\in A}\Sigma_{j\notin B} 2k^o_{ij} + \Sigma_i\Sigma_{j\in B}(k^+_{ij} + k^o_{ij} - k^-_{ij}).$$

Equivalently,

$$0 \leqslant \Sigma_{i\in A}\Sigma_{j\notin B}(-k^+_{ij} + k^o_{ij} + k^-_{ij}) + \Sigma_{i\notin A}\Sigma_{j\in B}(k^+_{ij} + k^o_{ij} - k^-_{ij}).$$

In view of (9.9) this is equivalent to (9.10).

Given the subset A of Y_1, one might as well choose the subset B of Y_2 so as to make the right hand side of (9.13) as small as possible. For $j \in Y_2$, putting j in B yields a contribution $\mu_2(\{j\}) = u_2(j)$; putting j in B^c yields a contribution $q(A \times \{j\}) = \Sigma_{i\in A}\, 2k^o_{ij}$. Thus, the best choice for B would be as in (9.12), in which case (9.13) reduces to (9.11). Ths proves the first part of Theorem 9.

If (9.11) fails for a set A then (9.13) fails for the pair A, B with B as in (9.12), which in turn means precisely that Condition (A, B) fails. This in turn allows us to improve the matrix $\mathbf{Z}$ as explained in the above Remark. □

ALGORITHM. A good algorithm for minimizing the sum (2.2) would be to apply the Theorems 7 and 9, at each stage of the calculation, to the matrix $\mathbf{Z} = (z_{ijk})$ of residues on hand. One first tries to construct the set of numbers W_{ij} as in Theorem 7 by using the standard max-flow-min-cut algorithm. If this does not succeed then $\mathbf{Z}$ is optimal and we are ready.

If this attempt does not succeed then the calculation automatically leads to a 'cut' of small capacity which in turn corresponds to the failure of a well-defined Condition (A, B). Using the latter knowledge, one next improves the matrix $\mathbf{Z}$ of residuals as explained in the Remark following Theorem 9. Afterwards, one tests the optimality of the new matrix $\mathbf{Z}'$ of residuals by trying to construct the desired numbers W_{ij}. And so on. Provided the original data $z^{(o)}_{ijk}$ are all integers, one can arrange the calculation so that also all subsequent residual

matrices $\mathbf{Z}$ are integral in which case the norm decreases each time by a positive integer. Hence, the calculation will then lead in finitely many steps to an optimal matrix of residuals.

In more detail, in trying to construct (W_{ij}), one considers a *directed capacitated network* with vertex set $V = \{a\} \bigcup \{b\} \bigcup Y_1 \bigcup Y_2$ and with a as the only source, b as the only sink. One has the following directed edges (x, y) and associated capacities $k(x, y)$.

(i) The edges (a, i) with $i \in Y_1$ and capacity $u_1(i)$.

(ii) The edges (i, j) with $i \in Y_1, j \in Y_2$ and capacity $2k_{ij}^o$.

(iii) The edges (j, b) with $j \in Y_2$ and capacity $u_2(j)$. Note that the $u_1(i)$, $u_2(j)$ and k_{ij}^o are integers and that $\Sigma_i u_1(i) = \Sigma_j u_2(j) = Q$ (say), see (7.7).

One proceeds with determing an admissible flow f in this network (with $f(x, y)$ as the flow along the directed edge (x, y)) which maximizes the total flow from a to b. Admissibility means here that $0 \leqslant f(x, y) \leqslant k(x, y)$.

Relative to a given admissible flow f, an *unsaturated path* from the vertex a to the vertex x is defined as a sequence $x_o = a, x_1, \ldots, x_{n-1}, x_n = x$ of distinct vertices such that the flow from a to x along that path can be increased. More precisely, this requires that, for $i = 1, \ldots, n$, either (x_{i-1}, x_i) is an edge of the network and $f(x_{i-1}, x_i)$ is smaller than the capacity $k(x_{i-1}, x)$; or (x_i, x_{i-1}) is an edge of the network and $f(x_i, x_{i-1})$ is positive.

Let V_f denote the set of all vertices x such that some unsaturated path leads from a to x. During the construction of V_f, one marks each new member of V_f with a single label pointing to a previously constructed vertex in V_f (from which it 'originated') so as to allow for backtracking. As soon as V_f is found to contain the sink b, one obtains through backtracking an unsaturated path from a to b. One proceeds to increase the flow along that path in an obvious and maximal way. This new flow is again integer valued, provided one starts with an integer valued flow (such as the zero flow). After finitely many steps, no further increase of the total flow from a to b is possible and one has reached an integer flow f with the property that $b \notin V_f$. Let

$$F = \Sigma_{i=1}^m f(a, i) = \Sigma_{j=1}^n f(j, b)$$

be the resulting total flow from a to b. There are the following possibilities.

(I) $F = Q$. In this case, the set of edge flows $W_{ij} = f(i, j)$ $(i \in Y_1; j \in Y_2)$ satisfies (9.6) and (9.7), consequently, the present residual matrix $\mathbf{Z}$ is optimal.

(II) $F < Q$. The $\mathbf{Z}$ is not optimal. In fact, Condition (A, B) fails with

$$A = V_f \bigcap Y_1 \quad \text{and} \quad B = V_f \bigcap Y_2, \tag{9.14}$$

allowing us to improve the matrix $\mathbf{Z}$.

Proof. Let E denote the set of edges (x, y) with $x \in V_f$ and $y \notin V_f$. The sum of all the corresponding capacities $k(x, y)$ is called the capacity of E. The set E is known to define a 'cut' whose capacity is equal to the maximal flow F on hand and, thus, is smaller than Q.

In fact, E consists of the edges (a, i) with $i \in A^c$, further the edges (j, b) with $j \in B$ and finally the edges (i, j) with $i \in A$ and $j \in B^c$. As is easily seen, the capacity of this set E being smaller than Q means exactly that (9.13) is false and, hence, that Condition (A, B) fails. □

REFERENCES

ABDELMALEK, N. N. (1974). On the discrete L_1 approximation and L_1 solutions of overdetermined linear equations, *J. Approx. Theory 11* 38–53.

AMEMIYA, I. and ANDO, T. (1965). Convergence of random products of contractions in Hilbert space, *Acta Sci. Math.* (Szeged) *26* 239–244.

ANSCOMBE, F. (1981). *Computing in Statistical Science through APL.* Springer-Verlag, N.Y.

ARMSTRONG, R. D., ELAM, J. J. and HULTZ, J. W. (1977). Obtaining least absolute value estimates for a two-way classification model, *Commun. Statist. Simula. Computa. B6(4)* 365–381.

ARMSTRONG, R. D., FROME, E. L. and KUNG, D. S. (1979). A revised simplex algorithm for the absolute deviation curve fitting problem, *Commun. Statist. Simula. Computa. B8(2)* 175–190.

BARRODALE, I. and ROBERTS, F. D. K. (1973). An improved algorithm for discrete ℓ_1 approximation, *SIAM J. Numer. Anal. 10* 839–848.

BARRODALE, I. and ROBERTS, F. D. K. (1978). An efficient algorithm for discrete ℓ_1 approximations with linear constraints, *SIAM J. Numer. Anal. 15* 603–611.

BARTELS, R. H., CONN, A. R. and SINCLAIR, J. W. (1978). Minimization techniques for piecewise differentiable functions: The ℓ_1 solution of an overdetermined linear system, *SIAM J. Numer. Anal. 15* 224–241.

BERGE, C. (1973). *Graphs and Hypergraphs,* North-Holland, Amsterdam.

FORD, Jr., L. K. and FULKERSON, D. K. (1962). *Flows in Networks,* Princeton University Press, Princeton.

FORSYTHE, A. B. (1972). Robust estimation of straight line regression coefficients by minimizing p-th power deviations, *Technometrics 14* 159–166.

GABRIEL, K. R. (1983). An alternative computation of median fits in two-way tables, *Tech. Report* No. 83/01, Department of Statistics, University of Rochester, Rochester, N.Y.

GARFINKEL, R. S. and NEMHAUSER, G. L. (1972). *Integer Programming,* Wiley, New York.

GENTLE, J. E. (1977). Least absolute value estimation: an introduction, *Commun. Statist. Simula. Computa. B6(4)* 313–328.

HU, T. C. (1969). *Integer Programming and Network Flows,* Addison-Wesley, Reading, MA.

JACOBS, K. (1978). *Measure and Integral,* Academic Press, New York.

KELLERER, H. (1961). Funktionen auf Produkträumen mit vorgegebenen Marginalfunktionen, *Math. Ann. 144* 323–344.

KEMPERMAN, J. H. B. (1983). On the role of duality in the theory of moments, pp. 63–92 in FIACCO, A. V. and KORTANEK, K. O., eds., *Semi-Infinite Programming and Applications,* Lecture Notes in Economics and Mathematical Systems, vol. 215, Springer-Verlag, New York.

KENNEDY, Jr., W. J. and GENTLE, J. E. (1980). *Statistical Computing,* Marcel Dekker, New York.

MONEY, A. H., AFFLECK-GRAVES, J. F. and BARR, G. D. I. (1982). The linear regression model: L_p norm estimation and the choice of p, *Commun. Statist. Simula. Computa. 11(1)* 89–109.

MOSTELLER, F. and TUKEY, J. W. (1977). *Data Analysis and Regression,* Addison-Wesley, Reading, MA.

SMITH, K. T., SOLMON, D. C. and WAGNER, S. L. (1977). Practical and mathematical aspects of the problem of reconstructing objects from radiographs, *Bull. Amer. Math. Soc. 83* 1227–1270.

STRASSEN, V. (1965). The existence of probability measures with given marginals, *Annals Math. Statist. 36* 423–438.

TUKEY, J. W. (1977). *Exploratory Data Analysis,* Addison-Wesley, Reading, MA.

WAGNER, H. M. (1959). Linear programming techniques for regression analysis, *J. Amer. Statist. Assoc. 54* 206–212.