# THE DISTANCE TO THE ORIGIN OF A CERTAIN POINT SET IN $E^n$

W. KARUSH AND N. Z. WOLFSOHN

1. **The problem.** Let $(a_1, a_2, \cdots, a_{m+1})$ be a point in real euclidean $(m+1)$-space $E^{m+1}$ $(m=0, 1, 2, \cdots)$ subject to the conditions

$$(1)\quad \sum_{i=1}^{m+1} a_i = 1, \quad \sum_{i=1}^{m+1} (i-1)a_i = 0, \quad \cdots, \quad \sum_{i=1}^{m+1} (i-1)^q a_i = 0,$$
$$0 \leqq q < m.$$

(Equivalent conditions are $\sum_1^{m+1} i^k a_i = 1$, $k=0, 1, \cdots, q$.)

The purpose of this paper is to prove the following:

THEOREM. *The minimum value of*

$$f(a) = \sum_{i=1}^{m+1} a_i^2$$

*subject to the side conditions* (1) *is*

$$(2)\qquad f_{\min} = 1 - \frac{m(m-1)\cdots(m-q)}{(m+1)(m+2)\cdots(m+q+1)}.$$

*Hence*

$$(3)\qquad f_{\min} = \frac{(q+1)^2 P(m)}{(m+1)(m+2)\cdots(m+q+1)},$$

*where $P(m)$ is a polynomial of degree $q$ and leading coefficient* 1. *Thus*

$$\lim_{m\to\infty} \frac{f_{\min}}{(q+1)^2/m} = 1.$$

The restriction $0 \leqq q < m$ has been imposed on the side conditions (1) and this assumption will be maintained in making the proof. However it is a simple matter to verify that for $m=0, 1, 2, \cdots, q$ the only point satisfying the equations (1) is $(1, 0, \cdots, 0)$. Hence we see directly that $f_{\min}=1$. But this value is also given by formula (2). Thus, the theorem remains valid without the restriction $0 \leqq q < m$.

The minimization problem described above arose in connection with a problem of linear smoothing of statistical data. Let $\xi_1, \xi_2, \xi_3, \cdots$ represent an incoming sequence of values of independent random

Received by the editors October 19, 1953.

variables, referred to as the "raw" data. It is desired to "smooth" this data by a linear formula of the type

$$x_n = a_1\xi_n + a_2\xi_{n-1} + \cdots + a_{m+1}\xi_{n-m}. \tag{4}$$

The value $x_n$ is the "smoothed" value of $\xi_n$. For a given value of $m$ the problem is to select the coefficients $a_j$ in some optimal fashion.

Let $\bar{\xi}_n$, $\bar{x}_n$ denote the averages of the random variables $\xi_n$, $x_n$ respectively. Let it be assumed that the variance of $\xi_n$ is independent of $n$; denote the common variance by $\sigma^2$. Then from (4) and the independence of the $\xi$'s it follows that the variance $\tau^2$ of $x_n$ is independent of $n$ and is given by

$$\frac{\tau^2}{\sigma^2} = f(a) = \sum_{i=1}^{m+1} a_i^2.$$

Choosing the $a$'s to minimize the ratio of variances leads to the problem of minimizing $f(a)$. The side conditions (1) enter the problem as follows. From (4), the average values satisfy the relation

$$\bar{x}_n = a_1\bar{\xi}_n + a_2\bar{\xi}_{n-1} + \cdots + a_{m+1}\bar{\xi}_{n-m}.$$

While requiring that $\tau^2$ be a minimum we must at the same time ensure that there is no systematic deviation of $\bar{x}_n$ from $\bar{\xi}_n$. With this in mind we may impose the requirement that

$$\bar{x}_n = \xi_n$$

shall hold for all (sufficiently large) $n$ whenever $\bar{\xi}_n$ is given by a formula of the type

$$\bar{\xi}_n = \alpha_0 + \alpha_1 n + \cdots + \alpha_q n^q,$$

with the $\alpha$'s arbitrary. It is readily shown that this requirement is identical with the restrictions (1). In this way we are led to the problem treated herein.

2. **Evaluation of certain determinants.** The solution of our minimization problem will require the evaluation of certain determinants. These determinants have as elements terms of the form $S_k(m)$, where

$$S_k(m) = \sum_{i=1}^{m} i^k.$$

It is well known[1] that $S_k(m)$ is a polynomial with leading term given by

[1] Whittaker and Robinson, *The calculus of observations*, 4th ed., 1944, p. 138.

$$S_k(m) = \frac{m^{k+1}}{k+1} + \cdots \tag{5}$$

and with $m$ as a factor:

$$S_k(m) = mQ_k(m), \qquad Q_k(m) \text{ a polynomial.} \tag{6}$$

It will be convenient to employ the following terminology. By a $T$-matrix we shall mean a square matrix $V$ with 1's along the main diagonal and 0's above the main diagonal. I.e.,

$$V = \begin{bmatrix} 1 & & \\ & & 0 \\ & 1 & \\ - & & \\ & & 1 \end{bmatrix}.$$

Let $x$ denote a column vector. Then by a $T$-transformation is meant a transformation

$$y = Vx,$$

where $V$ is a $T$-matrix. A $T$-transformation acts on a column vector as follows: it adds to each component of the vector a linear combination of the preceding components. The following facts concerning $T$-matrices are obvious.

(i) $|V| = 1$ ($|V|$ = determinant of $V$).

(ii) The product of $T$-matrices is a $T$-matrix.

LEMMA 1. *Let $s$ be any positive integer. Let $t$ be any integer such that*

$$1 \leqq t \leqq s$$

*and let $r_1, r_2, \cdots, r_t$ be any $t$ numbers. Then there exists a $T$-matrix $V$ of order $s+1$ whose elements are polynomials in $m$ and which has the following property: For any integer $k \geqq 0$ and any $t$ numbers $a_1, a_2, \cdots a_t$ the vector*

$$x = \begin{bmatrix} a_1(m-r_1)^k + a_2(m-r_2)^k + \cdots + a_t(m-r_t)^k \\ a_1(m-r_1)^{k+1} + a_2(m-r_2)^{k+1} + \cdots + a_t(m-r_t)^{k+1} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_1(m-r_1)^{k+s} + a_2(m-r_2)^{k+s} + \cdots + a_t(m-r_t)^{k+s} \end{bmatrix} \tag{7}$$

*is transformed by $V$ into a vector of the form*

$$(8) \qquad Vx = \begin{pmatrix} v_1(m) \\ \cdot \\ \cdot \\ \cdot \\ v_t(m) \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix},$$

*where $v_j(m)$ is a polynomial in $m$.*

Proof. The proof will be made by induction on $s$. Consider first $s=1$. Then $t=1$ and

$$x = \begin{pmatrix} a_1(m - r_1)^k \\ a_1(m - r_1)^{k+1} \end{pmatrix}.$$

Then (8) holds if $V$ is chosen as

$$V = \begin{pmatrix} 1 & 0 \\ -(m - r_1) & 1 \end{pmatrix}.$$

Assume the theorem holds for $s-1$ (with $s>1$) and consider the value $s$. Suppose first that $t=1$. Then

$$x = \begin{pmatrix} a_1(m - r_1)^k \\ a_1(m - r_1)^{k+1} \\ \cdot \\ \cdot \\ \cdot \\ a_1(m - r_1)^{k+s} \end{pmatrix}.$$

This vector can be transformed to the form (8) by adding to each component (except the first) the product of $-(m-r)$ with the preceding component. I.e.,

$$V = \begin{pmatrix} 1 & & & \\ -(m - r_1) & 1 & 0 & \\ & -(m - r_1) & & \\ 0 & & 1 & \\ & & -(m - r_1) & 1 \end{pmatrix}.$$

This establishes the desired result for $t=1$. Suppose now that

$$1 < t \leqq s$$

and consider the vector $x$ of equation (7). Multiplying each component by $-(m-r_1)$ and adding to the next component we see that by a $T$-matrix $V_1$, the vector $x$ can be brought to

$$V_1x = \begin{pmatrix} a_1(m-r_1)^k + a_2(m-r_2)^k + \cdots + a_t(m-r_t)^k \\ a_2(r_1-r_2)(m-r_2)^k + \cdots + a_t(r_1-r_t)(m-r_t)^k \\ a_2(r_1-r_2)(m-r_2)^{k+1} + \cdots + a_t(r_1-r_t)(m-r_t)^{k+1} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_2(r-r_2)(m-r_2)^{k+s-1} + \cdots + a_t(r_1-r_t)(m-r_t)^{k+s-1} \end{pmatrix}.$$

Let $y'$ denote the vector with $s$ components which are those of the above vector with the first component deleted. Since $1 \leqq t-1 \leqq s-1$ and the theorem is assumed to hold for $s-1$, there exists a $T$-matrix $V_2'$ of order $s$ depending only upon $r_2, r_3, \cdots, r_t$ such that

$$V_2' y = \begin{pmatrix} v_2(m) \\ \cdot \\ \cdot \\ \cdot \\ v_t(m) \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix}.$$

Now augment the matrix $V_2'$ to a matrix of order $s+1$ as follows:

$$V_2 = \begin{pmatrix} 1 & 0 \\ 0 & V_2' \end{pmatrix}.$$

Then $V_2(V_1x)$, i.e., $(V_2V_1)x$ has the required form (8). This completes the proof.

LEMMA 2. *Let*[2]

$$A = \begin{pmatrix} S_0(m) & S_1(m) & \cdots S_q(m) \\ S_1(m) & S_2(m) & \cdots S_{q+1}(m) \\ \cdots & \cdots & \cdots \\ S_q(m) & S_{q+1}(m) & \cdots S_{2q}(m) \end{pmatrix}.$$

*Then*

(9) $$|A| = Hm^{q+1}(m-1)^q(m+1)^q \cdots (m-q)(m+q)$$

[2] The determinant $|A|$ belongs to a general class of determinants called Hankel (or recurrent) determiants. The general form of such a determinant is $|a_{i+j}|$, $i, j=0, 1, \cdots, q$, with $a_0, a_1, \cdots, a_{2q}$ a given set of numbers. For a discussion of these determinants see G. Kowalewski, *Determinantetheorie*, New York, Chelsea, 1948, p. 102 ff.

*where $H$ is the nonzero determinant:*[3]

$$H = \begin{vmatrix} 1/1 & 1/2 & \cdots & 1/(q+1) \\ 1/2 & 1/3 & \cdots & 1/(q+2) \\ \cdot & \cdot & \cdot & \cdot \\ 1/(q+1) & 1/(q+2) & \cdots & 1/(2q+1) \end{vmatrix}.$$

PROOF. Let us show first that the determinant is a polynomial of degree $(q+1)^2$ with leading coefficient $H$, as indicated by equation (9). By equation (5) the matrix of leading terms of the elements of $A$ is

$$C = \left(\frac{m^{i+j-1}}{i+j-1}\right), \qquad i, j = 1, 2, \cdots, q+1.$$

Consider the complete expansion of the determinant $C$. For each permutation $j(i)$ of $1, 2, \cdots, q+1$ there will be a term of this expansion of the form

$$\prod_{i=1}^{q+1} \frac{m^{i+j(i)-1}}{i+j(i)-1}.$$

The degree in $m$ of this term is $\sum_{i=1}^{q+1} (i+j(i)-1) = 2(\sum i) - (q+1) = (q+1)^2$. It follows that

$$|C| = Hm^{(q+1)^2}.$$

Thus,

$$|A| = Hg(m)$$

where $g(m)$ has degree $(q+1)^2$ and leading coefficient 1. The lemma will thereby be proved if $|A|$ can be shown to have each of the factors displayed in equation (9).

Consider first the factor $m$. From equation (6) each element of $A$ has $m$ as a factor; consequently $|A|$ has $m^{q+1}$ as a factor, as was to be shown. Next consider $m-r$ for any fixed $r=1, 2, \cdots, q$. For $k=0, 1, 2, \cdots$ we have

$$S_k(m) = S_k(m-r) + (m-r+1)^k + \cdots + m^k. \tag{10}$$

Thus

$$A = A_1 + A_2$$

[3] It is known that $H = (2!\ 3! \cdots q!)^4/2!\ 3! \cdots (2q+1)!$. See, e.g., Pólya and Szegö, *Aufgaben und Lehrsätze aus de Analysis*, vol. 2, New York, Dover, 1945, pp. 98–99 and 300.

where

$$A_1 = (S_{i+j}(m - r)), \qquad i, j = 0, 1, \cdots, q,.$$

and $A_2$ is a matrix with each column of the form (7), the various columns differing only in the value of $k$. By Lemma 1 there exists a $T$-matrix $V$ of order $q+1$ such that the matrix $VA_2$ is identically zero below the $r$th row. Also each element of $VA_1$ has the factor $(m-r)$ since this is true of each element of $A_1$ by equation (6). It follows that each element of $VA = VA_1 + VA_2$ below the $r$th row has $m-r$ as a factor. Thus $(m-r)^{q-r+1}$ is a factor of $|VA| = |V| \cdot |A| = |A|$, as was to be shown.

The factor $m+r$ may be treated in a similar manner. Writing

$$S_k(m) = S_k(m + r) - (m + r)^k - \cdots - (m + 1)^k \tag{11}$$

and eliminating $k$th powers on the right by pre-multiplication of $A$ by a $T$-matrix as above, we may show that $(m+r)^{q-r+1}$ is also a factor of $|A|$. This completes the proof.

Lemma 3. *Let*

$$B = \begin{vmatrix} S_0(m + 1) & S_1(m) & \cdots S_q(m) \\ S_1(m) & S_2(m) & \cdots S_{q+1}(m) \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot \\ S_q(m) & S_{q+1}(m) & \cdots S_{2q}(m) \end{vmatrix}.$$

*Then*

$$|B| = Hm^q(m-1)^{q-1}(m+1)^{q+1} \cdots (m-q)^0(m+q)^2(m+q+1). \tag{12}$$

Proof. Notice that $B$ differs from $A$ of Lemma 2 only in having the argument $m+1$ in place of $m$ in the first entry of the matrix. It is clear therefore that $|B|$ has the same leading coefficient and degree as $A$; this is indicated in equation (12). As before it is sufficient to show that the factors of $|B|$ are those displayed in (12).

The factor $m^q$ appears in $|B|$ because $m$ is a factor of each element of $B$ below the first row. Consider $m-r$ for any fixed $r=1, 2, \cdots, q-1$. Using (10) for $k=1, 2, 3, \cdots$ and leaving the first row of $B$ unaltered we may write

$$B = B_1 + B_2$$

where $B_1$ and $B_2$ are identical with $A_1$ and $A_2$ respectively below the first row. Let $x'$ be any column vector of $B_2$ with the first element deleted. Then by Lemma 1 there exists a $T$-matrix $V'$ of order $q$ such that the vector $V'x'$ has all its components beyond the first $r$ equal to 0; the same matrix $V'$ applies to each of the columns of $B_2$ with first element deleted. Augment $V'$ to

$$V = \begin{pmatrix} 1 & 0 \\ 0 & V' \end{pmatrix}.$$

It follows that $VB_2$ is identically zero below the $(r+1)$st row. Also, each element of $VB$ below the first row has $m-r$ as a factor. Thus each element of $VB = VB_1 + VB_2$ below the $(r+1)$st row has $m-r$ as a factor. It follows as above that $(m-r)^{q-r}$ is a factor of $|B|$, as was to be shown.

The factors $m+r$, $r=1, 2, \cdots, q+1$, are handled differently. First consider $m+1$. From $S_0(m+1) = m+1$ and from equation (11) with $r=1$, $k=1, 2, 3, \cdots$, we see that $m+1$ is a factor of each element of $B$. Therefore $(m+1)^{q+1}$ is a factor of $|B|$. Next consider $m+r$ for fixed $r=2, 3, \cdots, q+1$. From equation (11)

$$S_k(m) = R_k(m + r) - (m + r - 1)^k - \cdots - (m + 1)^k, \qquad k = 1, 2, 3, \cdots,$$

where $m+r$ is a factor of $R_k(m+r)$. For $k=0$ we write

$$S_0(m + 1) = R_0(m + r) - (m + r - 1)^0 - \cdots - (m + 1)^0, \qquad R_0(m + r) = m + r.$$

Thus

$$B = C_1 + C_2$$

with

$$C_1 = (R_{i+j}(m + r)), \qquad i, j = 0, 1, \cdots, q,$$

and every column of $C_2$ of the form (7) with $t = r-1$. Thus by Lemma 1 there exists a $T$-matrix $V$ such that $VC_2$ is identically zero below the $(r-1)$st row. By an argument like the one above it follows that $(m+r)^{q-r+2}$ is a factor of $|B|$, as was to be shown. The proof is complete.

3. **Proof of main theorem.** We turn now to the proof of the theorem stated in §1. With the Lagrange multiplier rule in mind, form

$$F = \sum_{i=1}^{m+1} a_1^2 - 2\lambda_0 \sum_{i=1}^{m+1} a_i - 2\lambda_1 \sum_{i=1}^{m+1} (i - 1)a_i - \cdots$$
$$- 2\lambda_q \sum_{i=1}^{m+1} (i - 1)^q a_i.$$

Set $\partial F/\partial a_i = 0$:

$$\lambda_0 + (i-1)\lambda_1 + \cdots + (i-1)^q\lambda_q = a_i, \quad i = 1, 2, \cdots, m+1. \tag{13}$$

The $m-q+2$ linear equations (1) and (13) are to be solved for the unknown $a$'s and $\lambda$'s. Let us transform this set of equations. Multiply each side of the $i$th equation (13) by $(i-1)^k$, $k=0, 1, \cdots, q$, and add. The result is the following set of equations

$$\begin{aligned} S_0(m+1)\lambda_0 + S_1(m)\lambda_1 + \cdots + S_q(m)\lambda_q &= 1, \\ S_1(m)\lambda_0 + S_2(m)\lambda_1 + \cdots + S_{q+1}(m)\lambda_q &= 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots &, \\ S_q(m)\lambda_0 + S_{q+1}(m)\lambda_1 + \cdots + S_{2q}(m)\lambda_q &= 0. \end{aligned} \tag{14}$$

The matrix of coefficients of the $m+q+2$ equations (13) and (14) is readily seen to have determinant $|B|$ (except possibly for sign) with $B$ as in Lemma 3. It follows from that lemma and the condition $q<m$ that the matrix of coefficients referred to is nonsingular. Therefore, equations (13) and (14) are equivalent to equations (1) and (13) and either set of equations has the same unique solution

$$(\bar{a}_1, \cdots, \bar{a}_{m+1}, \bar{\lambda}_0, \cdots, \bar{\lambda}_q).$$

Now let $(a_1, a_2, \cdots, a_{m+1})$ be a point which provides $f(a)$ with an absolute minimum relative to the side conditions (1). From the standard necessary condition for a relative minimum it follows that there exists a set of multipliers $\lambda_0, \lambda_1, \cdots, \lambda_q$ which together with the $a$'s satisfy equations (1) and (13), or equivalently, equations (13) and (14). Since equations (13) and (14) have exactly one solution, this solution corresponds to the absolute minimum of $f(a)$. We may now determine the minimum value $f_{\min}$. Multiply the $i$th equation (13) by $a_i$ and add. The result is

$$f_{\min} = \lambda_0.$$

From equations (14)

$$\lambda_0 = \frac{\begin{vmatrix} 1 & S_1(m) & \cdots & S_q(m) \\ 0 & S_2(m) & \cdots & S_{q+1}(m) \\ \cdot & \cdot & \cdots & \cdot \\ 0 & S_{q+1}(m) & \cdots & S_{2q}(m) \end{vmatrix}}{|B|}.$$

Thus

$$1 - \lambda_0 = \frac{|A|}{|B|}\,.$$

From Lemmas 2 and 3,

$$1 - \lambda_0 = \frac{m(m-1)\cdots(m-q)}{(m+1)(m+2)\cdots(m+q+1)}\,.$$

Hence

$$f_{\min} = \lambda_0 = 1 - \frac{m(m-1)\cdots(m-q)}{(m+1)(m+2)\cdots(m+q+1)},$$

which is equation (2) of the theorem to be proved. Equation (3) is immediately derivable from equation (2). The proof is complete.

University of Chicago and
Institute for Advanced Study