PSYCHOMETRIKA—VOL. 42, NO. 4
DECEMBER, 1977

LOWER BOUNDS FOR THE RELIABILITY OF THE TOTAL SCORE ON A TEST COMPOSED OF NON-HOMOGENEOUS ITEMS: II: A SEARCH PROCEDURE TO LOCATE THE GREATEST LOWER BOUND

BRIAN WOODHOUSE AND PAUL H. JACKSON

UNIVERSITY COLLEGE OF WALES

Finding the greatest lower bound for the reliability of the total score on a test comprising n non-homogenous items with dispersion matrix Σ_X is equivalent to maximizing the trace of a diagonal matrix Σ_E with elements θ_i, subject to Σ_E and $\Sigma_T = \Sigma_X - \Sigma_E$ being non-negative definite. The cases $n = 2$ and $n = 3$ are solved explicity. A computer search in the space of the θ_i is developed for the general case. When Guttman's λ_4 (maximum split-half coefficient alpha) is not the g.l.b., the maximizing set of θ_i makes the rank of Σ_T less than $n - 1$. Numerical examples of various bounds are given.

Key words: reliability bounds, non-homogeneous composites, search procedure, trace minimization with constraints.

1. Introduction

In the companion paper it was shown that the observed-score dispersion matrix Σ_X (assumed known or well-estimated) of the n items making up a test is the sum of the true-score and error-score dispersion matrices Σ_T and Σ_E, where Σ_E is a diagonal matrix with elements denoted by θ_i. Thus if Σ_X has elements σ_{ij}, then Σ_T has known off-diagonal elements σ_{ij} and unknown diagonal elements $t_i = \sigma_{ii} - \theta_i$. It was also shown that finding the greatest lower bound (g.l.b.) for the reliability ρ of the total score on the test is equivalent to maximizing the trace $\sum\theta_i$ of Σ_E, or minimizing the trace $\sum t_i$ of Σ_T, subject to the conditions

(A) Σ_T is non-negative definite, and

(B) $\theta_i \geq 0$ (or $t_i \leq \sigma_{ii}$) for all i:

the g.l.b. is the corresponding value of $1 - \sum \theta_i/\sigma_X{}^2$, where $\sigma_X{}^2 = \sum\sum \sigma_{ij}$. Continuity considerations show that the g.l.b. will be attainable (that is, will be

Present affiliation of the first author: St. Hild's College of Education, Durham City, England.
Requests for reprints should be sent to P. H. Jackson, Department of Statistics, University College of Wales, Aberystwyth, Dyfed, Wales.

a possible value of ρ), and that the corresponding matrix Σ_T will be positive semi-definite, with rank $n - 1$ or less, and with det $\Sigma_T = 0$.

The present paper develops a method for finding the g.l.b. by search in the space of vectors $\theta = (\theta_1, \theta_2, \cdots, \theta_n)'$, referred to as the θ-space, or in the space of vectors $t = (t_1, t_2, \cdots, t_n)'$, called the t-space. Since $t_i = \sigma_{ii} - \theta_i$, the t-space can be obtained from the θ-space by reflection in the origin and a known translation; it is therefore merely a matter of convenience which space we use. In Sections 2 and 4 the cases $n = 2$ and $n = 3$ are solved explicitly, using θ-space and t-space respectively. In Section 3 a general formulation of the search problem is given, and in Section 5 a computer search procedure is described which can, subject to constraints of machine size, handle any number of items.

Guttman's [1945] bounds are referred to at various points for comparison purposes: they were reviewed in the present framework in the companion paper.

In Section 6 we show that his λ_4 (sometimes) fails to be the g.l.b. only because the θ_i can be chosen so as to make the rank of Σ_T less than $n - 1$. (Here, as in the companion paper, we use λ_4 to denote the *largest* split-half coefficient alpha, which is the bound Guttman implicitly proposed, although in his paper he uses the symbol λ_4 to denote *any* split-half coefficient alpha.) Numerical examples of his bounds, together with the g.l.b. are given in Section 4.

The expression diag (a), where a is a vector, denotes a diagonal matrix whose elements are, in order, those of a; for example Σ_E = diag (θ). Also l denotes the vector $(1, 1, \cdots, 1)'$, and I_n the identity matrix of order n.

2. *The Case n = 2*

If we work in θ-space, the requirement that Σ_T and Σ_E be non-negative definite is equivalent to the inequalities

$$0 \leq \theta_1 \leq \sigma_{11}, \quad 0 \leq \theta_2 \leq \sigma_{22} \tag{2.1}$$

and

$$(\sigma_{11} - \theta_1)(\sigma_{22} - \theta_2) \geq \sigma_{12}^2. \tag{2.2}$$

Inequalities (2.1) confine θ to a rectangle OABC (see Fig. 1). If we take the equality sign in (2.2), we have a rectangular hyperbola with asymptotes $\theta_1 = \sigma_{11}$, $\theta_2 = \sigma_{22}$. Only one branch passes through the rectangle defined by (2.1), and the points satisfying (2.2) lie on the side of it nearer the origin. Thus (2.1) and (2.2) define an admissible region (shaded in Fig. 1) lying in the first quadrant between the axes and the hyperbola.

The equation $\theta_1 + \theta_2 = c$, a constant, represents a line of slope -1. If the tangent to the hyperbola with slope -1 meets it at a point S of the admissible region, as in Fig. 1, the co-ordinates of this point give the maximum of $\theta_1 + \theta_2$

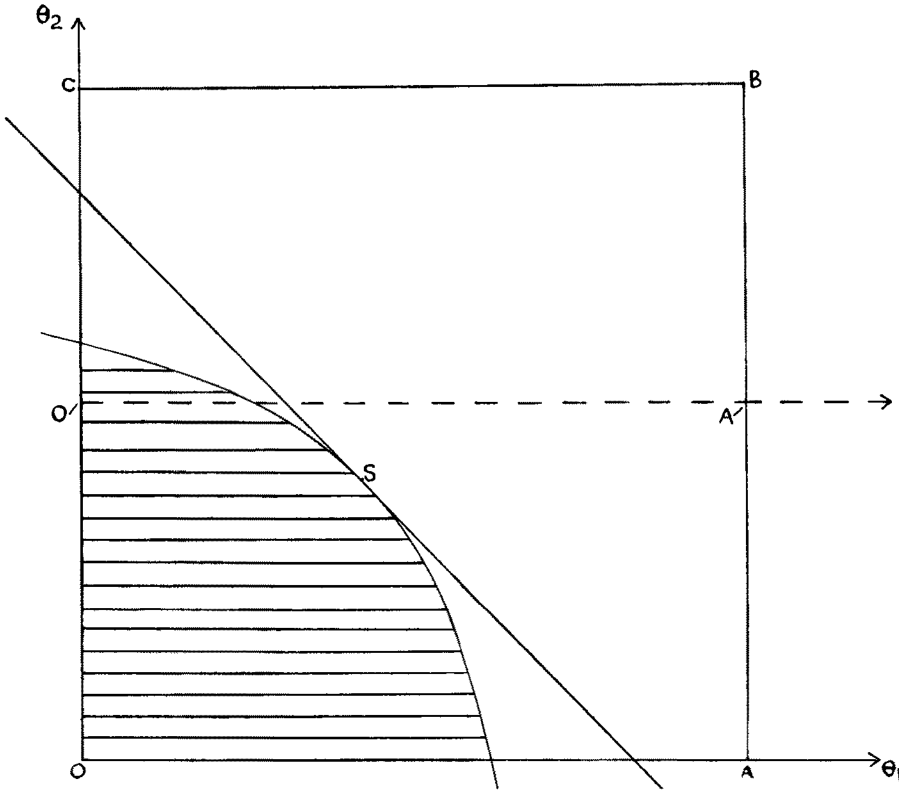

FIGURE 1
Geometrical representation of the case $n = 2$

subject to (2.1) and (2.2). Now the slope of the tangent at (θ_1, θ_2) is

$$m = -\frac{(\sigma_{22} - \theta_2)}{(\sigma_{11} - \theta_1)}$$

$$= -\frac{{\sigma_{12}}^2}{(\sigma_{11} - \theta_1)^2},$$

since (θ_1, θ_2) lies on the hyperbola. Thus the points at which the tangent has slope -1 are $(\sigma_{11} \pm \sigma_{12}, \sigma_{22} \pm \sigma_{12})$, and from (2.1) the point on the relevant branch has $\theta_1 = \sigma_{11} - |\sigma_{12}|, \theta_2 = \sigma_{22} - |\sigma_{12}|$, or $t_1 = t_2 = |\sigma_{12}|$. The corresponding greatest lower bound for ρ is

$$\frac{l'\Sigma_T l}{{\sigma_X}^2} = \frac{4\sigma_{12}}{{\sigma_X}^2} \quad \text{if} \quad \sigma_{12} > 0 \tag{2.3}$$

$$= 0 \qquad \text{if } \sigma_{12} \leq 0. \tag{2.4}$$

Equation (2.3) is just coefficient alpha (Guttman's λ_3) with $n = 2$. Equation

(2.4), which gives the same result as Guttman's λ_2, reflects the fact that if we add two negatively correlated scores in the absence of evidence about the relative sizes of the true-score variances, we have no way of knowing that the true scores do not cancel each other out, leaving a total score which is pure "noise".

Suppose, however, that the tangent meets the hyperbola at a point outside the admissible region. For the sake of definiteness, suppose the origin is at 0′ in Fig. 1 and the θ_1-axis is the dashed line. The coordinates of the point of contact of the tangent are therefore respectively positive and negative, implying that

$$\sigma_{11} > |\sigma_{12}| > \sigma_{22}. \tag{2.5}$$

Clearly the maximum possible value of $\theta_1 + \theta_2$ is now found by taking $\theta_2 = 0$, and $\theta_1 = \sigma_{11} - \sigma_{12}^2/\sigma_{22}$. (Analytically, this is "clear" because the admissible region is convex.) The corresponding g.l.b. for ρ is

$$1 - \frac{\theta_1}{\sigma_X^2} = \frac{(\sigma_{22} + \sigma_{12})^2}{\sigma_{22}\sigma_X^2}. \tag{2.6}$$

Note that this is positive even when σ_{12} is negative. This is because from (2.5) we have $t_2 \leq \sigma_{22} < |\sigma_{12}|$, and so, by (2.2), $t_1 > |\sigma_{12}|$. Thus we now *have* evidence that $t_1 \neq t_2$, and that the true-score components of the item scores cannot completely cancel each other out in the total score.

At first sight, a g.l.b. calculated on the assumption that one item error variance is zero may seem objectionable, since error-free scores are not encountered in practice. However, it must be remembered that this is still only a *lower* bound, not the actual value of ρ. If we were prepared to specify a minimum "practical" value, δ say, for θ_2, the g.l.b. would be obtained by taking $\theta_2 = \delta$, $\theta_1 = \sigma_{11} - \sigma_{12}^2/(\sigma_{22} - \delta)$, and would be even *larger* than (2.6).

3. *Formulation of the General Problem*

In θ-space, the constraints are

$$\theta_i \geq 0 \text{ all } i, \text{ and} \tag{3.1}$$

$$\Sigma_T = \Sigma_X - \text{diag}(\theta) \text{ must be non-negative definite.} \tag{3.2}$$

The admissible region has the following properties:

(i) It is convex. For if θ' and θ'' are any two vectors satisfying (3.1) and (3.2), with corresponding true-score dispersion matrices Σ_T' and Σ_T'' then the vector

$$\theta = s\theta' + (1 - s)\theta'', \qquad 0 \leq s \leq 1,$$

clearly has elements satisfying (3.1). Also if Σ_T is the matrix corresponding to θ, for any n-vector y we have

$$y'\Sigma_T y = sy'\Sigma_T' y + (1 - s)y'\Sigma_T'' y \geq 0,$$

showing that (3.2) is also satisfied.

(ii) It contains the origin. For Σ_X is positive definite.
(iii) It lies in the positive hyperoctant, bounded by the coordinate hyperplanes and the branch nearest to the origin of the surface

$$\det \Sigma_T = 0. \tag{3.3}$$

The first two parts of this statement follow from (3.1). To establish the third part, imagine a variable point θ, with associated true-score dispersion matrix Σ_T, moving outward from the origin along any line with non-negative direction cosines. At the origin $\Sigma_T = \Sigma_X$, and all its eigenvalues are positive. Condition (3.2) will first be violated when one of the eigenvalues becomes negative. Since θ varies continuously, so do the eigenvalues of Σ_T, and so before one becomes negative it must become zero, at which point (3.3) is satisfied (the determinant is the product of the eigenvalues). The important point to notice here is that all the other constraints implicit in (3.2)—all the principal determinants of any size must be non-negative—do not actually form part of the boundary of the admissible region: they simply determine which branch of (3.3) is the relevant one (compare the role of the constraints $\theta_1 \leq \sigma_{11}$, $\theta_2 \leq \sigma_{22}$ in the case $n = 2$). However, in the present problem, since we know that the origin is an admissible point, we have a much simpler definition of the relevant branch, namely the one which can be reached from the origin without passing through any other branch.

The equation $\sum \theta_i = c$ represents a hyperplane at distance $cn^{-1/2}$ from the origin, whose normal has direction cosines each $n^{-1/2}$. Maximizing $\sum \theta_i$ subject to (3.2) can be viewed as moving this plane away from the origin to the ultimate point at which it maintains contact with the relevant branch of (3.3); put another way, we need to find the point S on this branch which is furthest from the origin in the direction with cosines each $n^{-1/2}$ (a "north-easterly" direction). Provided this point also satisfies (3.1), it determines the greatest lower bound for ρ. If any θ_i is negative at this point, the convexity of the admissible region shows that it must be set to zero in determining the maximum admissible value of $\sum \theta_i$. The problem is then reduced to one of lower dimension. Specifically, if we renumber the items so that the first m of the θ_i are the ones to be set to zero, we can partition Σ_X and Σ_T as

$$\Sigma_X = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \quad \text{and} \quad \Sigma_T = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22}^* \end{pmatrix}$$

where Σ_{11} is a known $m \times m$ positive definite matrix. Then the standard result

$$\det \Sigma_T = \det \Sigma_{11} \cdot \det (\Sigma_{22}^* - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})$$

shows that the residual problem is identical with the $(n - m)$-dimensional problem which commences with an observed-score dispersion matrix $\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$.

4. *The Case* $n = 3$

The algebra will be simpler and more legible if we work in t-space and write $\sigma_{11} = a$, $\sigma_{22} = b$, $\sigma_{33} = c$, $\sigma_{23} = \sigma_{32} = f$, $\sigma_{31} = \sigma_{13} = g$, $\sigma_{12} = \sigma_{21} = h$. Without loss of generality we suppose that $f \geq g \geq h$.

Translating the formulation of the previous section into t-space, we find that the (convex) admissible region is bounded by the planes $t_1 = a$, $t_2 = b$, $t_3 = c$, and the surface

$$\det \Sigma_T = t_1 t_2 t_3 - f^2 t_1 - g^2 t_2 - h^2 t_3 + 2fgh = 0. \tag{4.1}$$

The point W with co-ordinates (a, b, c) is admissible. Consider planes $t_1 + t_2 + t_3 = d$, with d decreasing from its value $(a + b + c)$ at W until that plane is reached which just makes contact, at a point S, with the branch of (4.1) nearest to W. (The corresponding situation in two dimensions can be visualised by holding Fig. 1 upside down: B is the origin in t-space, and 0 is the point W). The co-ordinates of S give the minimum possible value of $\sum t_i$, provided that they satisfy $t_i \leq \sigma_{ii}$ for each i. We henceforth assume that the latter condition is satisfied at S, since we have seen that if it is not, we must take $\theta_i = 0$, $t_i = \sigma_{ii}$, and the problem reduces to the case $n = 2$ which was solved in Section 2.

To establish the geometrical properties of the surface (4.1), consider its sections by planes $t_3 = k > 0$. These are rectangular hyperbolas

$$(t_1 - g^2k^{-1})(t_2 - f^2k^{-1}) = (h - fgk^{-1})^2, \tag{4.2}$$

except in the case $k = fg/h$, when the section is the pair of straight lines

$$t_1 = \frac{gh}{f}, \quad t_2 = \frac{hf}{g}. \tag{4.3}$$

Note that (4.3) means that different branches of (4.1) meet at the point $(gh/f, hf/g, fg/h)$, and that the relevant branch has a singularity there. In algebraic terms, these values of t_i reduce the rank of Σ_T to 1, whereas at other points of (4.1) it is of rank 2. It is this slightly unexpected feature of the problem which results in Guttman's λ_4 not always being the g.l.b.

Where the *relevant* branch of (4.1) intersects the plane $t_3 = k > 0$, we have $kt_1 = t_3 t_1 \geq g^2$, and $kt_2 = t_3 t_2 \geq h^2$. Hence the two factors of the left member of (4.2) are positive, and the minimum value of $t_1 + t_2$ on the relevant curve is (using the inequality: arithmetic mean $\geq$ geometric mean)

$$k^{-1}(g^2 + f^2) + 2\,|h - fgk^{-1}|, \tag{4.4}$$

attained at the point given by

$$t_1 - g^2k^{-1} = t_2 - f^2k^{-1} = |h - fgk^{-1}|. \tag{4.5}$$

This statement is true also in the special case (4.3). There are three cases to consider, corresponding to three regions of k:

(i) if $h > fgk^{-1}$, the minimum value (4.4) is $(f - g)^2k^{-1} + 2h$;

(ii) if $h < fgk^{-1}$, the minimum value (4.4) is $(f + g)^2k^{-1} + 2h$;

(iii) if $h = fgk^{-1}$, the minimum value (4.4) is $h(f^2 + g^2)/fg$.

Adding k to each of these, and again using the fact that the arithmetic mean of two positive numbers is not less than the geometric mean, we find that the minimum value of $t_1 + t_2 + t_3$ on the relevant branch of (4.1) in each of the three regions is

(i) $2\{f - g + h\}$, attained with $t_3 = k = f - g$;

(ii) $2\{|f + g| - h\}$, attained with $t_3 = k = |f + g|$;

(iii) $(g^2h^2 + h^2f^2 + f^2g^2)/fgh$, attained with $t_3 = k = fg/h$;

provided that, in cases (i) and (ii), the point at which the minimum would be attained does indeed lie in the corresponding region of k; if not, the smallest value occurs on the boundary, $t_3 = k = fg/h$ as in case (iii). Taking account of the proviso, and the fact that $k > 0$, we find in each of the possible cases (see Table 4.1) that there is a single global minimum, not a choice between local minima. This must, of course, be so, since the admissible region is convex.

Examination of the last column of Table 4.1 shows that when $f \geq g \geq h$, there are only three possible values of the g.l.b. One of these arises from the singular point of (4.1). The remaining two are easily verified to be Guttman's λ_4 (under the stated conditions), based respectively on the split where one "half" comprises the first two items, and on the nugatory split in which one "half" comprises all three items—in the notation of the companion paper, corresponding to vectors $u = (1, 1, -1)$ and $u = (1, 1, 1)$ respectively. (The other splits would be relevant if the ordering of f, g, and h were different.) Thus

TABLE 4.1

Computation of the Greatest Lower Bound Assuming that $f \geq g \geq h$ and $t_i \leq \sigma_{ii}$

Conditions		Case	t_1	t_2	t_3	g.l.b.
$f \geq g \geq h \geq 0$,	$f+g \geq fg/h$	(iii)	gh/f	hf/g	fg/h	$(gh+hf+fg)^2/fgh\sigma_X^2$
$f \geq g \geq h \geq 0$,	$f+g < fg/h$	(ii)	$g-h$	$f-h$	$f+g$	$4(f+g)/\sigma_X^2$
$f \geq g \geq 0 \geq h$		(ii)	$g-h$	$f-h$	$f+g$	$4(f+g)/\sigma_X^2$
$f \geq 0 \geq g \geq h$,	$\|f+g\| \leq fg/h$	(iii)	gh/f	hf/g	fg/h	$(gh+hf+fg)^2/fgh\sigma_X^2$
$f \geq 0 \geq g \geq h$,	$f+g > fg/h$	(ii)	$g-h$	$f-h$	$f+g$	$4(f+g)/\sigma_X^2$
$f \geq 0 \geq g \geq h$,	$-(f+g) > fg/h$	(ii)	$-(g+h)$	$-(f+h)$	$-(f+g)$	0
$0 \geq f \geq g \geq h$		(ii)	$-(g+h)$	$-(f+h)$	$-(f+g)$	0

TABLE 4.2

Examples of the Various Bounds for Artificial Data Sets with n = 3

Ref	a	b	c	f	g	h	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	ρ_3	ρ_4	g.l.b.	θ_1	θ_2	θ_3
1	2	4	8	3	2	1	.462	.711	.692	.769	.739	.647	.711	.611	.381	.769	1	2	3
2	5	5	5	4	4	4	.615	.923	.923	.821	.905	.889	.923	.862	.885	.923	1	1	1
3	1	5	14	4	3	2	.474	.719	.711	.737	.737	.968	.793	.837	.965	.974	0	0	1

Guttman's λ_4 fails to be the g.l.b. in certain cases, only because it is possible to choose θ so that Σ_T has rank less than $n - 1 = 2$.

Table 4.2 gives artificial examples of the various bounds discussed in the companion paper, together with the g.l.b. and the co-ordinates in θ-space of the point at which it occurs (λ_3 is coefficient alpha, λ_4 the greatest split-half bound). It was shown in the companion paper that, assuming any bound which is negative to be replaced by zero, there is a partial ordering among the bounds, as shown in Table 4.3. The examples cited show that no further ordering is possible, since every pair of bounds not joined by a line on the ordering

TABLE 4.3

Partial Ordering of the Various Lower Bounds

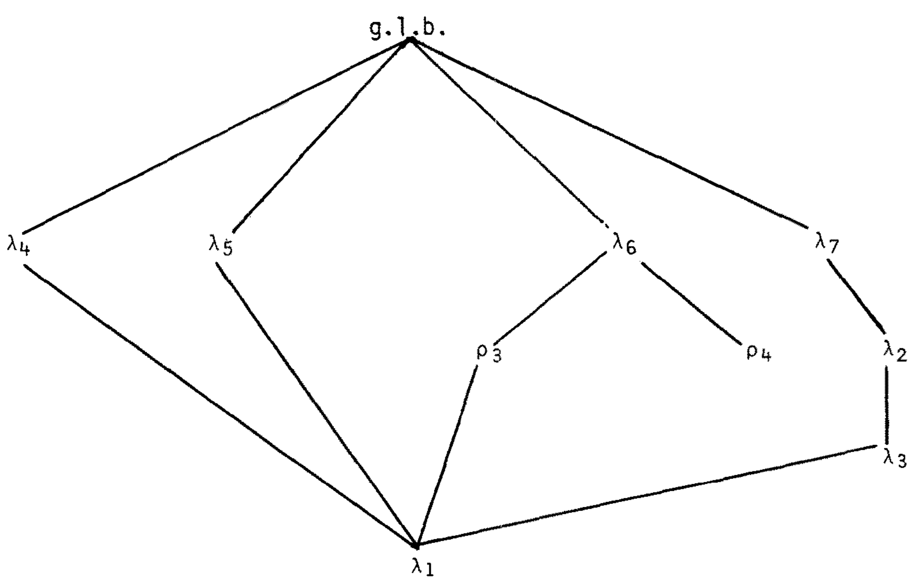

diagram occurs in the examples in both $a >$ and $a <$ relation. The examples also show that even in non-trivial cases the ordering must be understood as $\geq$ and not $>$. Specifically we have:

Example 1 g.l.b. $= \lambda_4 > \lambda_5 > \lambda_7 = \lambda_2 > \lambda_3 > \lambda_6 > \rho_3 > \lambda_1 > \rho_4$

2 g.l.b. $= \lambda_7 = \lambda_2 = \lambda_3 > \lambda_5 > \lambda_6 > \rho_4 > \rho_3 > \lambda_4 > \lambda_1$

3 g.l.b. $> \lambda_6 > \rho_4 > \rho_3 > \lambda_7 > \lambda_4 = \lambda_5 > \lambda_2 > \lambda_3 > \lambda_1$

5. *The Search Procedure*

In Section 3 we established that the greatest lower bound corresponds to that point of the (convex) admissible region of θ-space (bounded by the coordinate hyperplanes and the relevant branch of the surface det $\Sigma_T = 0$) which is farthest from the origin in the direction $l = (1, 1, \cdots, 1)'$, which we shall call the NE direction by analogy with the case $n = 2$. This point can be reached by a path lying nowhere outside the admissible region, and consisting of iteration of the following pair of steps:
(i) from any admissible point (initially we can use the origin) move in a NE direction to meet the surface;
(ii) from this point of the surface move in the plane normal to l to a point in the interior of the admissible region.
There are thus three main problems to solve:

(a) to determine the point at which a step of type (i) meets the surface;
(b) to devise a "good" method of taking steps of type (ii);
(c) to find satisfactory criteria for terminating the search.

Problem (a) is solved as follows. Let θ^* be any point of the admissible region and let $\Theta = \text{diag}(\theta^*)$. A general point on a step of type (i) from θ^δ is $\theta^* + kl$, $k > 0$. The step meets the boundary of the admissible region where

$$\text{(5.1)} \qquad \det \Sigma_T = \det (\Sigma_X - \Theta - kI_n) = 0$$

for the first time. But (5.1) determines the eigen values of $\Sigma_X - \Theta$; hence the value of k at the boundary is the smallest of these.

Problem (b) can be treated as follows. Let θ^* now denote a point on the boundary surface and let $\Theta = \text{diag}(\theta^*)$ as before. Let u_{gh} denote a vector with g-th element $+1$, h-th element -1, and remaining elements zero, and write $U = \text{diag}(u_{gh})$. Clearly u_{gh} is orthogonal to l, and so points $\theta = \theta^* + ku_{gh}$, for suitable k, are points of a step of type (ii). Such a point lies *on* the boundary surface if

$$\text{(5.2)} \qquad \det (\Sigma_X - \Theta - kU) = 0.$$

This is a quadratic equation for k, with roots k_1, k_2, of which one, say k_1, is known to be zero. The mid-point of the chord joining the points of the boundary surface corresponding to k_1 and k_2, $\theta = \theta^* + \frac{1}{2}k_2u_{gh}$, is a suitable interior point of the admissible region, provided that $\theta_g \geq 0$ and $\theta_h \geq 0$; if one of these conditions is violated, the point at which the chord leaves the admissible region is substituted (if $k_2 > 0$, take $k = \theta_h^*$; if $k_2 < 0$, take $k = -\theta_g^*$).

The following method can be used to solve (5.2) for k_2. Let

$$(5.3) \qquad \begin{aligned} D_i &= \det(\Sigma_X - \Theta - k_i U), \qquad i = 2, 3, 4, \\ &= ak_i^2 + bk_i \end{aligned}$$

(since (5.2) has one root zero). Eliminating a and b between the three equations (5.3) we obtain

$$(5.4) \qquad k_2 = \frac{(k_3^2 D_4 - k_4^2 D_3)}{(k_3 D_4 - k_4 D_3)},$$

where k_3 and k_4 can be chosen arbitrarily, and the corresponding determinants D_3 and D_4 evaluated. In practice, the choice can be exercised so as to reduce rounding error in the computation of the determinants and in the numerator and denominator of (5.4). It was found satisfactory to relate k_3 and k_4 to the chord length at the previous cycle, subject to a minimum value for each.

The vector u_{gh} represents a particular direction in the plane normal to l. The direction can be varied by choice of g and h. The search was initially programmed in cycles of n iterations, with g running through the values 1 to n, and $h = g + 1$ (modulo n). However, both speed and accuracy of convergence to the optimal point were found to be improved by using cycles comprising all $n(n - 1)/2$ possible pairs (g, h).

The solution of problem (*c*), the choice of convergence criterion, was facilitated by the observation that, using the improved procedure, the length of the step to the boundary from the interior point found by using a particular direction u_{gh} decreased from cycle to cycle, although step lengths were not necessarily decreasing within cycles. Since the step length is the increase in $\sum \theta_i$, this suggests that if we require accuracy to the m-th decimal place in the g.l.b., we can stop when the maximum step-length in the preceding $n(n - 1)/2$ iterations is $10^{-m}\sigma_X^2/n(n - 1)$; this ensures that the m-th decimal place will not change in the next $n(n - 1)/2$ iterations. In practice, this criterion may be unnecessarily stringent: many of the step lengths in a cycle will be very much smaller than the maximum—this was the reason for preferring the cycle of $n(n - 1)/2$ pairs (g, h) to the cycle of only n pairs—and stopping when the maximum step-length in the last cycle is less than $5.10^{-(m+1)}\sigma_X^2/n$ will usually be satisfactory.

The details of the implementation of this search procedure will depend upon the available computer hardware and software. A full description of an implementation in ALGOLW on an IBM 370/168 machine can be found in the first author's M.Sc. thesis [Note 1].

6. *The Greatest Lower Bound and Guttman's* λ_4

In this section we generalize a remark, made in Section 4 for the case $n = 3$, concerning the circumstances in which Guttman's λ_4 (maximum split-half coefficient alpha) is the greatest lower bound. We suppose that the conditions $\theta_i \geq 0$ are not active constraints. We then prove that if the greatest lower bound corresponds to a point of the boundary det $\Sigma_T = 0$ at which Σ_T has rank exactly $n - 1$, it is equal to λ_4. The possibility of finding better bounds than λ_4 in some situations thus arises from the fact that, when $n > 2$, θ can be chosen to make the rank of Σ_T less than $n - 1$.

As in the description of the search procedure, let θ^* be a point of the boundary det $\Sigma_T = 0$. Let y be any vector satisfying $y'1 = 0$, and let $Y = \text{diag}(y)$. Then $\theta^* + ky$ represents a point on a line through θ^* perpendicular to the (NE) direction of maximization. Because of the convexity of the admissible region, such a line will, in general, meet the boundary at one other point. If θ^* is the optimal point however, the line will meet the boundary at two coincident points with $k = 0$. Thus the equation in k,

$$\det(\Sigma_T - kY) = 0, \tag{6.1}$$

must have zero roots. The constant term in this equation is just det Σ_T, which we know must be zero at the optimal point.

The coefficient of k in (6.1) is $-\sum y_i T_{ii}$, where T_{ii} is the cofactor of the i-th diagonal element of Σ_T, and this coefficient must be zero at the optimal point for any y such that $y'1 = 0$. We can find $n - 1$ linearly independent vectors y satisfying the condition, giving $n - 1$ linearly independent equations, $\sum y_i T_{ii} = 0$, for the T_{ii}. These equations have two solution sets: (i) the "trivial" solution $T_{ii} = 0$ for all i, implying that Σ_T is of rank less than $n - 1$; and (ii) a single non-trivial solution, which must be

$$T_{11} = T_{22} = \cdots T_{nn} = c \neq 0, \tag{6.2}$$

since by definition of the linearly independent vectors y, the set $(1, 1, \cdots, 1)$ is a solution; in this case the rank of Σ_T is $n - 1$.

(The same conclusion can be reached using differential calculus, as follows. The partial derivative of $\det(\Sigma_X - \Theta)$ with respect to θ_i is T_{ii}. If the partial derivatives are not all zero, they are direction ratios of the tangent hyperplane to the boundary at θ, and so must be in the ratio $1:1:\cdots:1$ at the optimal point. If they are all zero, the point θ is a cusp, which could also be the optimal point.)

If adj Σ_T denotes the matrix adjugate to Σ_T, with (i, j)-th element the cofactor T_{ji} of the (j, i)-th element of Σ_T, we have the general standard result

$$\Sigma_T \operatorname{adj} \Sigma_T = (\det \Sigma_T) I_n. \tag{6.3}$$

It follows that when det $\Sigma_T = 0$, if u is any column of adj Σ_T, then

$$\Sigma_T u = 0. \tag{6.4}$$

But when Σ_T is of rank $n - 1$, the equations (6.4) have a unique non-trivial solution for u (apart from a scalar multiplier); thus all columns of adj Σ_T are constant multiples of each other. In particular, it follows that $T_{ii}/T_{ij} = T_{ji}/T_{jj}$, or, using the symmetry of Σ_T, and hence of adj Σ_T,

$$T_{ij}^2 = T_{ii}T_{jj} = c^2 \tag{6.5}$$

at the optimal point, by (6.2); that is, $T_{ij} = \pm c$, for all i, j.

Thus we have shown that if Σ_T is of rank $n - 1$ at the optimal point, then equation (6.4) holds with u a vector each of whose elements is ± 1. But in the companion paper (see particularly equations (2.9) and (2.10) and surrounding text) it was shown that λ_4 can be obtained by considering each of the 2^{n-1} essentially different vectors u of this form, finding the corresponding value of the reliability on the assumption that $u'\Sigma_T u = 0$, which value is necessarily a lower bound to the actual reliability, and selecting the greatest of these bounds. When Σ_T is of rank $n - 1$ at the optimal point, $u'\Sigma_T u = 0$ there for one of the set of vectors u considered; the corresponding value of the reliability is the greatest lower bound to the actual reliability, and therefore necessarily the greatest of the set of bounds considered; thus the g.l.b. is equal to λ_4.

7. *Relation to Earlier Work*

A referee has drawn our attention to an earlier approach to the problem by Bentler [1972], using a formulation similar to our own but without the constraints $\theta_i \geq 0$. The non-negative definiteness of Σ_T is ensured by writing it in the form FF', where F is an $n \times n$ matrix whose elements are regarded as variables, and the trace of FF' is minimized with respect to the elements of F and the θ_i, subject to the $n(n + 1)/2$ distinct constraints $FF' + \Sigma_E = \Sigma_X$. This leads to a set of equations in the n error variances θ_i and the $n(n + 1)/2$ Lagrange multipliers, which can be solved iteratively.

We have applied our search method to each of the examples in Bentler's paper, and verified that when the constraints $\theta_i \geq 0$ are not active in determining the greatest lower bound the two methods yield the same result. In his first

TABLE 8.1

Bounds Computed for Various Combinations of Subtests (Real Data)

Subtests included	λ_3	λ_2	λ_4	g.l.b.
1,3,4,5	.651	.687	.699	.730
1,2,5,6	.466	.529	.633	.633
1,2,3,4,5,6	.725	.774	.866	.870
1,2,3,6,8,10	.785	.820	.932	.936
1 through 18	.945	.954	.980	.989

example, however, where his method yields a "Heywood solution"—that is, a solution in which one or more of the θ_i is negative—his lower bound has the value .969 whereas the g.l.b. is .976.

Bentler's method can be extended to incorporate the conditions $\theta_i \geq 0$ by writing $\theta_i = u_i^2$ and differentiating with respect to u_i instead of u_i^2. However, this device leads only to the conclusion that *either* the solution ignoring the conditions is correct, *or* some of the θ_i must be equal to zero. Thus it seems that in order to use this approach to find the greatest lower bound, one must first solve the problem ignoring the constraints on the θ_i, and then use the procedure discussed at the end of Section 3 of this paper to reduce the dimensionality of the problem if any of the θ_i in the initial solution are negative.

8. Concluding Remarks

The methods of this paper were applied to data resulting from the administration of 20 subtests, using both the full 20 × 20 dispersion matrix and also selected submatrices (thus obtaining bounds for the reliability of the total score on a test comprising only the corresponding subtests). The data, the dispersion matrix, and examples of the bounds additional to those given in Table 8.1, are to be found in the first author's M.Sc. thesis [Note 1].

Table 8.1 gives examples which seem fairly typical for these particular data, in which λ_4 (maximum split-half coefficient alpha) is a considerable improvement on λ_3 (coefficient alpha) or λ_2, while the gain in passing to the g.l.b. is fairly modest. However, we have already seen in Table 4.2, Case 2, that this is not always so. With the computer and the programs used, computation of λ_4 required less time than computation of the g.l.b. for numbers of subtests up to 17 (this comparison of course depends on the degree of accuracy required for the g.l.b.).

Finally, it should be remarked that throughout these two papers it has been assumed that Σ_X is known (that is, well-estimated). This condition will easily be met, for example, in national test administrations by major testing corporations. However, before the bounds can safely be used with dispersion matrices computed from the scores of a modest number of candidates, we need to know more about their sampling distributions in samples from a population with known dispersion matrix Σ_X. It seems clear that this information will have to be obtained by simulation. Studies of this kind are in progress, and it is hoped that they will be reported later.

REFERENCE NOTE

1. Woodhouse, B. *Lower bounds for the reliability of a test.* University of Wales M.Sc. thesis, 1976.

REFERENCES

Bentler, P. M. A lower-bound method for the dimension-free measurement of internal consistency. *Social Science Research*, 1972, *1*, 343–357.

Guttman, L. A basis for analyzing test-retest reliability. *Psychometrika*, 1945, *10*, 255–282.

Manuscript received 9/2/76
Final version received 5/3/77