DISCUSSION

T. W. ANDERSON[1]

Stanford University

As stated by the author in Section 3, his emphasis in the study of

$$Dy = \Gamma(\theta) = \sum_{a=1}^{s} \theta_a A_a \tag{1}$$

is different from mine [Anderson (1969, 1970, 1973)]. A brief exposition of this other point of view might put the present paper into a larger perspective. I shall use the notation of the present paper as well as the assumption of $Ey = 0$ and corresponding modifications in statements.

Anderson considered N observations from $N(0, \Gamma(\theta))$; in the present paper $N = 1$. We suppose $A_1, \dots, A_s$ to be symmetric matrices and that for some vector $\theta = (\theta_1, \dots, \theta_s)'$ the matrix $\sum_{a=1}^s \theta_a A_a$ is positive definite. Then the derivative equations for the maximum likelihood estimate of θ are

$$\operatorname{tr}\left(\sum_{a=1}^{s} \hat{\theta}_a A_a\right)^{-1} A_b = y'\left(\sum_{a=1}^{s} \hat{\theta}_a A_a\right)^{-1} A_b \left(\sum_{a=1}^{s} \hat{\theta}_a A_a\right)^{-1} y, \qquad b = 1, \dots, s. \tag{2}$$

If there are several vectors $\hat{\theta}$ satisfying (2), the vector minimizing $|\sum_{a=1}^s \hat{\theta}_a A_a|$ is taken. Rewriting (2) as

$$\begin{aligned} &\sum_{c=1}^{s} \operatorname{tr}\left(\sum_{a=1}^{s} \hat{\theta}_a A_a\right)^{-1} A_b \left(\sum_{a=1}^{s} \hat{\theta}_a A_a\right)^{-1} A_c \hat{\theta}_c \\ &\quad = y'\left(\sum_{a=1}^{s} \hat{\theta}_a A_a\right)^{-1} A_b \left(\sum_{a=1}^{s} \hat{\theta}_a A_a\right)^{-1} y, \qquad b = 1, \dots, s, \end{aligned} \tag{3}$$

suggests an iterative procedure for solving (2). Let the vector $\hat{\theta}^{(0)}$ be an initial estimate, and define $\Gamma(\hat{\theta}^{(i)}) = \sum_{a=1}^s \hat{\theta}_a^{(i)} A_a$, $i = 0, 1, \dots$. The ith stage of the iteration is solving the linear equations

$$\begin{aligned} &\sum_{c=1}^{s} \operatorname{tr} \Gamma^{-1}(\hat{\theta}^{(i-1)}) A_b \Gamma^{-1}(\hat{\theta}^{(i-1)}) A_c \hat{\theta}_c^{(i)} \\ &\quad = y'\Gamma^{-1}(\hat{\theta}^{(i-1)}) A_b \Gamma^{-1}(\hat{\theta}^{(i-1)}) y, \qquad b = 1, \dots, s. \end{aligned} \tag{4}$$

This is the method of scoring.

If the matrices $\{A_a\}$ commute [the author's (c)], there exists an orthogonal matrix B such that $A_a = B\Lambda_a B'$ where Λ_a is diagonal with diagonal elements λ_{ta}, $t = 1, \dots, n$. Then

$$B'\Gamma(\theta)B = \sum_{a=1}^{s} \theta_a \Lambda_a \tag{5}$$

---

[1]Work sponsored by the Army Research Office Contract No. DAAG 29-85-0239 at Stanford University and written while the author was a National Research Council Resident Research Associate at the Naval Postgraduate School at Monterey, California.