PRACTICAL PROBLEMS OF MATRIX CALCULATION*

HAROLD HOTELLING

UNIVERSITY OF NORTH CAROLINA

1. Introduction

Statistical analysis involving any considerable number of variates leads usually to calculations which, if not well organized and conducted with the help of suitable machines, may be excessively laborious. The great possibilities opened up by the advances in the theory of multivariate statistical analysis will gain in accessibility as computational methods improve in this domain. The computations most needed involve matrices. This is true, for example, in the method of least squares and in calculating multiple correlation coefficients; it is true in the calculation of the generalized Student ratio and figurative distance that has become the modern substitute for Karl Pearson's coefficient of racial likeness; also in studying the relations between two or more sets of variates, and the principal components of one set.

The same computational problems arise also in many fields outside of statistics—if indeed we can speak of any field as being outside of statistics! Thus the study of vibrations in airplanes and other machines and structures; the analysis of stresses and strains in bridges, of electrical networks, of mass spectroscopy of petroleum products; and many other subjects of importance require calculations of the same kinds as multivariate analysis in statistics.

The calculations principally needed are of three main types:

a) Multiplication of matrices, that is, formation of sums of products.

b) Inversion of matrices and solution of systems of linear equations.

c) Determination of latent roots and latent vectors, also known as characteristic roots and vectors, as principal components, and by such mongrel terms as eigenvectors and eigenvalues. More generally, determination of values of $x_1, \cdots, x_p, \lambda$ satisfying

$$\sum_j (a_{ij} - \lambda b_{ij})x_j = 0\,, \qquad i, j = 1, \cdots, p.$$

In addition there are special matrix computational problems such as those encountered by Koopmans [**8**].[1] It might be thought that the calculation of determinants should have an important place here. However, in most problems of practical importance of which I am aware, if a determinant occurs in a formula it is associated with other determinants in such a way that all that is really needed is one or more of the ratios of determinants forming the inverse of a matrix. Hence these problems come under (*b*). For example, one of the

* A revision of a Symposium lecture, taking account of developments in the subject down to May, 1946, and, with an added note, to June, 1948.

[1] Boldface numbers in brackets refer to references at the end of the paper (p. 293).