ON THE APPLICATION OF HYPERSPACE GEOMETRY TO THE THEORY OF MULTIPLE CORRELATION.

BY RAJ CHANDRA BOSE, M.A.

[*Editorial Note.* In recent years, geometrical representation is being increasingly used in investigations in theoretical statistics. As these methods are not yet familiar to statistical workers in India, Mr. Raj Chandra Bose was requested to prepare a series of notes on this subject for presentation before study meetings of the Indian Statistical Institute.

The present note gives a summary of what may be called the classical portion of the subject. Prof. Karl Pearson in his paper on *Some Novel Properties of Partial and Multiple Correlation Coefficients in a Universe of Manifold Characteristics* (*Biometrika*, Vol. XI, May, 1916, pp. 231-238) had remarked (p. 237):—"It is greatly to be desired that the 'trigonometry' of higher dimensioned plane space should be fully worked out, for, all our relations between multiple correlation and partial correlation coefficients of n variates are properties of the 'angles,' 'edges' and 'perpendiculars' of sphero--polyhedra in multiple space. It would be a fine task for an adequately equipped pure mathematician to write a treatise on 'Spherical Polyhedrometry'; he need not fear that his results would be without practical application, for they embrace the whole range of problems from anatomy to medicine and from medicine to sociology, and ultimately to the doctrine of evolution."

Acting upon this suggestion, the late Professor James McMahon, of the Cornell University, gave a systematic treatment in an article on *Hyperspherical Goniometry; and its Application to Correlation Theory for N Variables* which was published in the *Biometrika*, Vol. XV, December, 1923, pp. 173-208.

Mr. Raj Chandra Bose, who was not acquainted at that time with Prof. McMahon's work, gives in this present note an independent discussion of the same subject. Although the ground covered is much the same, it will be noticed that there are interesting points on difference in his method of approach and treatment. He also gives some new results. As Prof. McMahon's article is not easily available, it is hoped that the present note will be found useful by workers in India.—*P. C. M.*]

1. Let there be m characters which have been measured for n individuals. We may then denote by x'_{ij} the ith character for the jth individual. Let us set

$$x_{ij} = x'_{ij} - S_j(x'_{ij})/n \qquad \dots \qquad \dots \qquad \dots \qquad \dots \quad (1\cdot1)$$

where S_j denotes summation from $j=1$ to $j=n$. That is, x_{ij} denotes the deviation of the ith character from its mean for the jth individual.

Now let us take a space of n dimensions and in it plot the points $X_1, X_2, \dots\dots X_m$, where the rectangular co-ordinates of X_i are $(x_{i1}, x_{i2} \dots\dots x_{in})$. The projection of the line OX_i along the jth axis is x_{ij}, and this gives the deviation of the ith character from its mean for the jth individual. The points $X_1, X_2 \ldots\ldots X_m$ may be called character points. The figure formed by these points is fundamental in our investigation.

2. Now let us consider an equation of the form

$$x_i = \lambda_{i1} x_1 + \lambda_{i2} x_2 + \ldots\ldots \lambda_{i(i-1)} x_{i-1} + \lambda_{i(i+1)} x_{i+1} \ldots\ldots + \lambda_{im} x_m \qquad (2\cdot1)$$

This may be regarded as an equation for determining the measure of the ith character for any individual when the measures of the remaining characters are known. The quantity

$$y_{ij} = x_{ij} - (\lambda_{i1} x_1 + \lambda_{i2} x_2 + \ldots \ldots\ldots \lambda_{im} x_m) \quad \ldots \quad \ldots \quad (2\cdot2)$$

is the residue or the error in estimating the ith character for the jth individual by means of equation (2·1).

It is our object to choose the constants $\lambda_{i1}, \lambda_{i2}$ .... in such a way as to make

$$S_j\,(y^2_{ij}) \quad \ldots \quad \ldots \quad \ldots \quad (2\cdot3)$$

as small as possible.

When so chosen, the constant λ_{ik} is the regression coefficient

$b_{ik\cdot123\ldots\ldots(i-1)(i+1)\ldots\ldots(k-1)(k+1)\ldots\ldots m}$ in Yule's notation,

and the equation (2·1) becomes

$$\begin{aligned} x_i = {} & b_{i1\cdot23\ldots(i-1)(i+1)\ldots\ldots m}\, x_1 + b_{i2\cdot13\ldots(i-1)(i+1)\ldots m}\, x_2 \\ & + \ldots b_{i(i-1)\cdot123\ldots(i-2)(i+1)\ldots m}\, x_{i-1} + b_{i(i+1)\cdot123\ldots(i-1)(i+2)\ldots m}\, x_{i+1} \\ & + \ldots b_{im\cdot12\ldots(i-1)(i+1)\ldots(m-1)}\, x_m \quad \ldots \quad \ldots \quad \ldots \quad (2\cdot4) \end{aligned}$$

This is the regression equation connecting the ith character with the remaining ones. The n values of y_{ij} may, in this case, be called the deviations of the nth order for the ith character, in relation to the characters 1, 2, 3,...... $(i-1)$, $(i+1)$, ... m. They may be denoted by $x_{i\cdot123\ldots(i-1)(i+1)\ldots m,j}$. If we put

$$\sigma^2_{i\cdot123\ldots\ldots(i-1)(i+1)\ldots\ldots m} = S_j(x^2_{i\cdot123\ldots(i-1)(i+1)\ldots m,j})/n \quad \ldots \quad (2\cdot5)$$

then $\sigma_{i\cdot123\ldots\ldots(i-1)(i+1)\ldots\ldots m}$ may be called the standard deviation of the mth order for the ith character in relation to the characters No. 1, 2, 3, ...$(i-1)(i+1)$, ...m.

3. It is now our object to interpret geometrically the regression coefficients $b_{ik\cdot123\ldots(i-1)(i+1)\ldots(k-1)(k+1)\ldots\ldots m}$, the deviations $x_{i\cdot123\ldots(i-1)(i+1)\ldots\ldots m,j}$, and the standard deviation $\sigma_{i\cdot123\ldots(i-1)(i+1)\ldots\ldots m}$.

For this purpose, we note that the co-ordinates of any point P lying in the hyperplane $OX_1 X_2 \ldots\ldots X_{i-1} X_{i+1} \ldots\ldots\ldots X_m$ are linear functions of the co-ordinates of the points $X_1, X_2 \ldots\ldots X_{i-1}, X_{i+1} \ldots\ldots X_m$. If the co-ordinates of P be

$$z_{ij} = \lambda_{i1}\, x_{1j} + \lambda_{i2}\, x_{2j} + \ldots\ldots \lambda_{i(i-1)}\, x_{(i-1)j} + \lambda_{i(i+1)}\, x_{(i+1)j} + \ldots\ldots\ldots \lambda_{im}\, x_{mj} \quad \ldots \quad (3\cdot1)$$

for $j = 1, 2, 3 \ldots\ldots\ldots\ldots\ldots m$, then

$$\overline{OP} = \lambda_{i1}\,\overline{OX}_1 + \lambda_{i2}\,\overline{OX}_2 + \ldots\ldots \lambda_{i(i-1)}\,\overline{OX}_{i-1} + \lambda_{i(i+1)}\,\overline{OX}_{i+1} + \ldots\ldots \lambda_{im}\,\overline{OX}_m \quad \ldots \quad (3\cdot2)$$

where a bar placed over any distance denotes that it is to be taken vectorially. Now

$$PX_i^2 = S_j(x_{ij} - z_{ij})^2 = S_j(y_{ij}^2) \quad \ldots \quad \ldots \quad \ldots \quad (3\cdot3)$$

We thus have to make PX_i^2 a minimum. P must then be the foot of the perpendicular from the point X_i on the hyperplane $OX_1X_2 \ldots\ldots\ldots X_{i-1}\ X_{i+1}\ldots\ldots\ldots X_m$.

We thus have the following result :—

If $X_{i\cdot 123\ldots(i-1)(i+1)\ldots m}$ denotes the foot of the perpendicular from the character point X_i on the hyperplane $OX_1\ X_2\ldots\ldots\ldots X_{i-1}\ X_{i+1}\ldots\ldots\ldots X_m$ formed by the origin and the remaining character points, then the regression coefficients $b_{ii\cdot 23\ldots\ldots(i-1)(i+1)\ldots\ldots m}$ are just those coefficients which are needed to express the co-ordinates of $X_{i\cdot 123\ldots(i-1)(i+1)\ldots m}$ linearly in terms of the co-ordinates of $X_1, X_2, \ldots\ldots X_{i-1}\ X_{i+1}, \ldots\ldots\ldots X_m$ (3·4)

Hence from (3·2) we have

$$\overline{OX}_{i\cdot 123\ldots(i-1)(i+1)\ldots m} = S_k(b_{ik\cdot 123\ldots(i-1)(i+1)\ldots(k-1)(k+1)\ldots m} \quad (3·5)$$

where the summation extends over all values of k from 1 to m except i.

Again the deviation $x_{i\cdot 23\ldots\ldots(i-1)(i+1)\ldots\ldots m\cdot j}$ is now the projection of the line $X_i\ X_{i\cdot 123\ldots(i-1)(i+1)\ldots m}$ along the axis of X_j ... ... (3·6)

While $X_i\ X_{i\cdot 123\ldots(i-1)(i+1)\ldots m} = \sqrt{m}.\sigma_{i\cdot 123\ldots(i-1)(i+1)\ldots m}$ ... (3·7)

4. Since the line $X_i\ X_{i\cdot 123\ldots(i-1)(i+1)\ldots m}$ is perpendicular to the hyperplane $OX_1\ X_2\ldots\ldots X_{i-1}\ X_{i+1}\ldots\ldots\ldots X_m$ it is perpendicular to the line OX_k $(i \neq k \leqq m)$ contained in the hyperplane. But the direction cosines of these lines are respectively proportional to $x_{i\cdot 123\ldots(i-1)(i+1)\ldots m\cdot j}$ and x_{kj} $(j=1, 2, 3, \ldots\ldots\ldots m)$.

Hence we have

$$S_j\ (x_{kj} x_{i\cdot 123\ldots(i-1)(i+1)\ldots m\cdot j}) = \quad \ldots \quad \ldots \quad \ldots \quad (4·1)$$

where S_j is a summation from $j=1$ to $j=n$, and $i \neq k$.

Thus the product moment of any deviation of the first order for any character, and a deviation of a higher order for any other character vanishes.

Next let us draw perpendiculars from X_1 and X_2 respectively on the hyperplane $OX_3\ X_4\ldots\ldots\ldots\ldots X_m$. According to our notation the feet of the perpendiculars are denoted by $X_{1\cdot 34\ldots\ldots m}$ and $X_{2\cdot 34\ldots\ldots m}$. Consider the projection of the line $X_1\ X_{1\cdot 34\ldots\ldots m}$ on the line $X_2\ X_{2\cdot 34\ldots\ldots m}$.

$$\text{Projec.}\ (X_1X_{1\cdot 34\ldots m}) = \text{Projec.}\ (X_1O) + \text{Projec.}\ (OX_{1\cdot 34\ldots m})$$
$$= \text{Projec.}\ (X_1O) \quad \ldots \quad \ldots \quad \ldots \quad (4·2)$$

since $OX_{1\cdot 34\ldots\ldots m}$ is perpendicular to $X_2\ X_{2\cdot 34\ldots\ldots m}$ being contained in a hyperplane to which $X_2\ X_{2\cdot 34\ldots\ldots m}$ is normal.

It follows from (3·6) that

$$S_j\ (x_{1\cdot 34\ldots\ldots m\cdot j})\ (x_{2\cdot 34\ldots\ldots m\cdot j})$$
$$= (X_2\ X_{2\cdot 34\ldots m}) \times \text{Projec.}\ (X_1\ X_{1\cdot 34\ldots m})\ \text{on}\ X_2\ X_{2\cdot 34\ldots m} \ldots (4·3)$$

In the same way

$$S_j\ (x_{1j})\ (x_{2\cdot 34\ldots m\cdot j}) = (X_2\ X_{2\cdot 34\ldots m}) \times \text{Projec.}\ (X_1O)\ \text{on}\ X_2\ X_{2\cdot 34\ldots m} \quad (4·4)$$

Hence from (4·2) we have

$$S_1 (x_{1.34\ldots m})(x_{2.34\ldots m}) = S_1 (x_0)(x_{2.34\ldots m}) \qquad \ldots \quad (4·5)$$

5. Now in conformity with our previous notation let $X_{1.234\ldots m}$ be the foot of the perpendicular from X_1 to the hyperplane $OX_2\, X_3\, X_4 \ldots\ldots X_m$ and let $X_{1.34\ldots m}$, $X_{2.34\ldots m}$ be the feet of the perpendiculars from X_1 and X_2 respectively on the hyperplane $OX_3\, X_4\, X_5 \ldots\ldots X_m$.

Consider the projection of the line $X_{1.34\ldots m}X_1$ on the line $X_{2.34\ldots m}X_2$. Now
Projec. $(X_{1.34\ldots m}X_1)$ = Projec. $(X_{1.34\ldots m}O)$ + Projec. $(OX_{1.234\ldots m})$ + Projec. $(X_{1.234\ldots m}X_1)$

But $X_{1.23\ldots m}X_1$ is perpendicular to a hyperplane which entirely contains $X_2\, X_{2.34\ldots m}$ and is therefore perpendicular to it, while $X_{1.34\ldots m}O$ lies in a hyperplane to which $X_2\, X_{2.34\ldots m}$ is perpendicular and is therefore perpendicular to it. Hence
Projec. $(X_{1.34\ldots m}X_1)$ = Projec. $(OX_{1.23\ldots m})$

$$= \text{Projec.}\,(b_{12.34\ldots m}OX_2) + \text{Projec.}\,(b_{13.245\ldots m}OX_3) + \ldots\ldots\ldots$$

$$+ \text{Projec.}\,(b_{1m.234\ldots(m-1)}OX_m) \qquad \text{from (3·5)}$$

$$= \text{Projec.}\,(b_{12.34\ldots m}OX_2)$$

$$= b_{12.34\ldots m}\, X_{2.34\ldots m}X_2$$

Thus if θ be the angle between the lines $X_{1.34\ldots m}X_1$ and $X_{2.34\ldots m}X_2$ we have

$$X_{1.34\ldots m}X_1 . \text{Cos}\,\theta = b_{12.34\ldots m}\, X_{2.34\ldots m}X_2 \qquad \ldots \qquad \ldots \quad (5·1)$$

It follows from (3·7) that

$$b_{12.34\ldots m} = (\sigma_{1.34\ldots m}\,\text{Cos}\,\theta)/\sigma_{2.34\ldots m} \qquad \ldots \qquad \ldots \quad (5·2)$$

It follows in the same way that

$$b_{21.34\ldots m} = (\sigma_{2.34\ldots m}\,\text{Cos}\,\theta)/\sigma_{1.34\ldots m} \qquad \ldots \qquad \ldots \quad (5·3)$$

If we define the partial coefficient of correlation $r_{12.34\ldots m}$ by the relation $r_{12.34\ldots m} = (b_{12.34\ldots m}.b_{21.34\ldots m})^{\frac{1}{2}}$, it follows from (5·2) and (5·3) that

$$r_{12.34\ldots m} = \text{Cos}\,\theta \qquad \ldots \qquad \ldots \qquad \ldots \qquad \ldots \quad (5·4)$$

But the angle between the lines $X_{1.34\ldots m}X_1$ and $X_{2.34\ldots m}X_2$ is the same as the angle between the hyperplanes $OX_1\, X_3\, X_4 \ldots\ldots X_m$ and $OX_2\, X_3 \ldots\ldots X_m$. We thus have the following geometrical interpretation for the correlation coefficient $r_{12.34\ldots m}$.

The correlation coefficient $r_{12.34\ldots m}$ is the cosine of the angle between the hyperplanes $OX_1\, X_3\, X_4 \ldots\ldots X_m$ and $OX_2\, X_3 \ldots\ldots X_m$. $\qquad \ldots \qquad \ldots \quad (5·5)$

6. The various identities and inequalities connecting the different correlation coefficients now flow from the result (5·5). We shall in the first instance consider the case of three variables only.

Let the lines OX_1, OX_2, OX_3 cut the unit sphere with centre O contained in the hyperplane $OX_1X_2X_3$ at the points A, B, C. These points form a spherical triangle A B C on the unit sphere whose elements we denote by the usual notation. Then from (5·5)

$$\left.\begin{array}{lll} r_{23} = \text{Cos } a, & r_{31} = \text{Cos } b, & r_{12} = \text{Cos } c \\ r_{23.1} = \text{Cos } A, & r_{31.2} = \text{Cos } B, & r_{12.3} = \text{Cos } C \end{array}\right\} \quad \ldots \quad (6.1)$$

But we know that

$$\text{Cos } A = (\text{Cos } b . \text{Cos } c - \text{Cos } a) / \text{Sin } b . \text{Sin } c$$

Hence,

$$r_{23.1} = (r_{31}\, r_{12} - r_{23}) / (1 - r^2_{31})^{\frac{1}{2}}\ (1 - r^2_{12})^{\frac{1}{2}} \quad \ldots \quad \ldots \quad (6.2)$$

Again from the reciprocal formula

$$\text{Cos } a = (\text{Cos } B . \text{Cos } C + \text{Cos } A) / \text{Sin } B . \text{Sin } C$$

we get

$$r_{23} = (r_{31.2}\, r_{12.3} - r_{23.1}) / (1 - r^2_{31.2})^{\frac{1}{2}}\ (1 - r^2_{12.3})^{\frac{1}{2}} \quad \ldots \quad \ldots \quad (6.3)$$

It is also known that

$$(\text{Sin } A)/(\text{Sin } a) = (\text{Sin } B)/(\text{Sin } b) = (\text{Sin } C)/(\text{Sin } c)$$

$$= \sqrt{(1 - \text{Cos}^2 a - \text{Cos}^2 b - \text{Cos}^2 c + 2 \text{ Cos } a \text{ Cos } b \text{ Cos } c)} / \text{Sin } a \text{ Sin } b \text{ Sin } c,$$

that is, $(1 - \text{Cos}^2 a - \text{Cos}^2 b - \text{Cos}^2 c + 2 \text{ Cos } a \text{ Cos } b \text{ Cos } c) < 1$

or $(\text{Cos}^2 a + \text{Cos}^2 b + \text{Cos}^2 c - 2 \text{ Cos } a \text{ Cos } b \text{ Cos } c) > 0$

or $(r^2_{23} + r^2_{31} + r^2_{12} - 2\, r_{23} . r_{31} . r_{12}) > 0$ ... ... (6.4)

From the reciprocal formula we get

$$(r^2_{23.1} + r^2_{31.2} + r^2_{12.3} + 2\, r_{23.1}\, r_{31.2}\, r_{12.3}) > 0 \quad \ldots \quad \ldots \quad (6.5)$$

In the same way, other well-known identities and inequalities, connecting the elements a, b, c, A, B, C, of the spherical triangle, will lead to corresponding identities and inequalities connecting the correlation coefficients.

Now coming to the general case, we note that the angles between the hyperplanes $OX_1\, X_4\, X_5 \ldots\ldots X_m$, $OX_2\, X_4\, X_5 \ldots\ldots X_m$, $OX_3\, X_4\, X_5 \ldots\ldots X_m$ are the same as the angles between the lines OY_1, OY_2, OY_3.........where OY_1, OY_2, OY_3.........are the sections of these hyperplanes by the hyperplane absolutely orthogonal to $OX_4\, X_5 \ldots\ldots X_m$. Again the angles between the hyperplanes $OX_2\, X_3\, X_4\, X_5 \ldots\ldots X_m$, $OX_3\, X_1\, X_4\, X_5 \ldots\ldots X_m$ $OX_1\, X_2\, X_4\, X_5 \ldots\ldots X_m$ are the same as the angles between the planes $OY_2\, Y_3$, $OY_3\, Y_1$ and $OY_1\, Y_2$. Consequently, if the lines OY_1, OY_2, OY_3 cut the unit sphere with centre O and lying in the hyperplane $OY_1\, Y_2\, Y_3$, at the points A, B, C, we have

$$\left.\begin{array}{lll} r_{23.45\ldots m} = \text{Cos } a, & r_{31.45\ldots m} = \text{Cos } b, & r_{12.45\ldots m} = \text{Cos } c \\ r_{23.145\ldots m} = \text{Cos } A, & r_{31.245\ldots m} = \text{Cos } B, & r_{12.345\ldots m} = \text{Cos } C \end{array}\right\} \quad (6.6)$$

Thus the results (6.2), (6.3), (6.4), (6.5) remain valid, if we add the same subscripts to every correlation coefficient. For example, corresponding to (6.2) we have

$$r_{23.145\ldots m} = \frac{r_{31.45\ldots m}\, r_{12.45\ldots m} - r_{23.45\ldots m}}{(1 - r^2_{31.45\ldots m})^{\frac{1}{2}}\ (1 - r^2_{12.45\ldots m})^{\frac{1}{2}}}$$

(*November, 1933.*)