On the Characteristic Values of the Correlation Tensor
and a New Measure of Correlation between Vector Quantities.

By Motosaburo Masuyama.

(Read Jan. 18, 1941).

Resume.

It is shown that none of characteristic values of the correlation tensor is negative. Utilising this fact, the author proposes the trace of the correlation tensor divided by $n$ as a measure of correlation between $n$-dimensional vector sets, which has the more favourable properties than the symmetric correlation coefficient.

§ 1. Introduction. We may postulate the following conditions for a measure of correlation, say $r(A; B)$, between vector quantities $\{A(t)\}$ and $\{B(t)\}$:

1°. $r(A; B)$ is scalar.

2°. $1 \equiv |r| \equiv 0$.

3°. If $A$ and $B$ are linearly related, then and only then $|r|=1$.

4°. If $A$ and $B$ are stochastically independent, then and only then $r=0$.

5°. $r$ is independent of the order of observation.

6°. $r(A; B)=r(B; A)$.

7°. $r$ is not affected by the choice of units of observed vector quantities.

8°. If $A(t)$ is a linear function of the stochastically independent
vector functions $B(t)$, $B_2(t)$, \ldots, $B_m(t)$, there exists a scalar function $f(r)$ which satisfies the equation

$$1 = \sum_{i=1}^{m} f(r(A; B_i)).$$

We have proved in our previous papers\(^{(1)}\) that the symmetric correlation coefficient (S.C.C.) almost satisfies the conditions $1^\circ-7^\circ$, but we can not yet find the explicit functional form of $f(r)$ for the general S.C.C. Investigating $f(r)$ for the S.C.C., we have obtained there the generalized Parseval's equation\(^{(2)}\) which shows that if $g(t)$ be a general function of $m$ $n$-dimensional totally orthogonal vector functions $u_1(t)$, $u_2(t)$, \ldots, $u_m(t)$, then

$$[gg] = \sum_{i=1}^{m} [u_i u_i]^{-1} [u_i g],$$

and that if the $n$-dimensional scattering of the set $\{g(t)\}$ be not equal to zero, we have from (1.2)

$$E = \sum_{i=1}^{m} R(g; u_i),$$

where

$$R(g; u_i) = [u_i u_i]^{-1} [u_i g] [gg]^{-1}.$$  

The formula (1.3) suggests us that the trace of $R(g; u_i)$ divided by $n$, say $\rho(g; u_i)$, may be used as a measure of correlation between vector quantities. The $\rho$ itself satisfies formally the condition $8^\circ$, viz.

$$1 = \sum_{i=1}^{m} \rho(g; u_i),$$

for the trace of $E$ is equal to $n$. It would be easy to see that $\rho (g; u_i)$ satisfies also the conditions $1^\circ, 5^\circ$ and $7^\circ$. Before verifying that $\rho(g; u_i)$ satisfies the conditions $2^\circ, 3^\circ$ and $6^\circ$, we shall investigate at first whether $\rho(g; u_i)$ satisfies the condition $4^\circ$ or not.

\section*{2. The Condition $4^\circ$.} As we consider only vectors and tensors of two valencies in the real Euclidean space $E_n$, we are able to use the matrix representation of tensors. The matrix in question will be called the correlation matrix and be expressed in the following form;

$$R = T'ATB,$$

where matrices $R$, $T'$, $A$, $T$ and $B$ correspond to the tensors $R$, $[g u_i]$.

\begin{itemize}
  \item[(1)] The author's papers on this line are issued in this Proceedings: I, 21 (1939), 638; II, 21 (1939), 647; III, 22 (1940), 579; IV, 22 (1940), 855; V, 22 (1940), 858; VI, 23 (1941), 194; VII, 23 (1941), 196.
  \item[(2)] That is to say, S.C.C. satisfies the condition $2^\circ$ with the exception of the sub-dependent cases and the condition $4^\circ$ under the normal law of frequency.
  \item[(3)] The paper III.
  \item[(4)] By slight modification we can extend our result for the bivector.
\end{itemize}
On the Characteristic Values of the Correlation Tensor.

$[u, u_i]^{-1}$, $[u^T g]$ and $[g g]^{-1}$ respectively. $T^*$ means the conjugate (or the transposed) of $T$. Both $A$ and $B$ are self-conjugate and positive definite.

As $A$ and $B$ are both self-conjugate and positive definite and their characteristic equations have no zero root, there exist self-conjugate and positive definite matrices $a$ and $b$ in such a relation as

$$(2.2) \quad a^2 = A, \quad b^2 = B.$$ 

Transform $R$ by the real non-singular matrix $b$, then we have

$$(2.3) \quad bRb^{-1} = (b^T a)(a^T b) = -(a^T b)'(a^T b).$$

for the matrices $a$ and $b$ are both self-conjugate.

The right side of this equation shows that the characteristic equation of $bRb^{-1}$, and accordingly that of $R$ itself, has only non-negative roots and therefore

$$(2.4) \quad Sp. R \geq 0.$$ 

The equality $Sp. R = 0$ takes place when and only when $R = 0$, viz. the condition $4^2$ is satisfied by $\rho(g; u_i)$ at least under the normal law of frequency.

§ 3. The Condition 2'. As

$$(3.1) \quad \rho(g; u_i) = \frac{1}{n} Sp. R(g; u_i) \geq 0 \quad (i = 1, 2, \ldots, m),$$

we have from (1.5) and (2.4)

$$(3.2) \quad 1 \geq \rho(g; u_i) \geq 0,$$

viz. the condition $2^2$ is satisfied by $\rho(g; u_i)$.

§ 4. The Condition 3'. The equality in the left side of the formula (3.2) occurs when and only when

$$(4.1) \quad \rho(g; u_i) = 0 \quad (k \neq i),$$

viz.

$$(4.2) \quad E = R(g; u_i) and R(g; u_k) = 0 \quad (k \neq i).$$

This means that $g$ is a linear function of $u_i$. The reverse however is not true. That is to say, even if there is a linear relation between $g$ and $u$, the equality

$$(4.3) \quad \rho(g; u) = 1$$

do not occur in some cases. Such a discrepancy takes place when

(5) The paper II.


(7) This inequality can be deduced by the more direct way:

$$Sp. R = Sp. Tu^0 T^0 b = Sp. 6^0 Tu^0 T^0 = Sp. a^0 Tu^0 T^0 a^0 \geq 0.$$ 


(8) The paper VII.
and only when \( g \) and \( u \) are linearly subdependent, viz. when we approximate \( g \) by a linear function of \( u \) in such a way that

\[
(4.4) \quad \sum_{r=1}^{n} |g(t) - Su(t)|^2 = \text{Min},
\]

the \( n \)-dimensional scattering of

\[
(4.5) \quad g(t) - Su(t) = r(t)
\]

is equal to zero, but \( r(t) \) itself is not equal to zero. Thus the condition 3 is not satisfied perfectly by \( \rho(g; u) \). This happens also for the S.C.C. \(^9\). But in such a case we cannot presume \( g \) with \( u \). Since our main intention to introduce the measure of correlation is to estimate the \( u \)'s contribution to \( g \) under the assumption of the presumability, it would be expedient to exclude the subdependent cases from our practical meaning of "the perfect linear relation" \(^10\).

To find the explicit functional form of linear dependency is equivalent to find the characteristic vectors of \( R(g; u) \) and \( R(u; g) \) (cf. \( \S5 \)). For example, if

\[
(4.6) \quad G'g + U'u = 0,
\]

then we have

\[
(4.7) \quad G - G'R(g; u) = 0, \quad \text{or} \quad R'G' = EG'.
\]

which means that the matrix \( G' \), and accordingly the matrix \( G \) itself, is composed of the characteristic vectors belonging to the characteristic value \( \lambda = 1 \).

\(5\). The Condition 6'. The remaining condition 6' is also satisfied by \( \rho(g; u) \). By definition

\[
(5.1) \quad R(u; g) = [u'] [g']^{-1} [u] [g]^{-1}.
\]

If \( \det[u g] \neq 0 \), we have

\[
R(g; u) = [g']^{-1} R(u; g) [u'] [g]^{-1} = ([g']^{-1} [u'] [g]^{-1}) R(u; g) ([g']^{-1} [u'] [g]^{-1})^{-1},
\]

that is to say, \( R(g; u) \) is equivalent to \( R(u; g) \). Thus \( R(u; g) \) and \( R(g; u) \) have the same characteristic values, and accordingly the same trace, viz.

\[
(5.3) \quad \text{Spec} R(g; u) = \text{Spec} R(u; g), \quad \text{or} \quad \rho(g; u) = \rho(u; g).
\]

Considering the limiting case for (5.3), we obtain also the equation of the form (5.3), even when \( \det[u g] = 0 \).

\( \S6\). A New Measure of Correlation \( \rho \). After our investigation in the above sections, we may take \( \rho(A; B) \) as a new measure of correlation.

\(^9\) The paper IV.

\(^{10}\) It is not impossible to construct a measure of correlation which satisfies the condition 3'. For example, we can construct such a theory on the basis of the equation (1.3) in our paper III.
On the Characteristic Values of the Correlation Tensor.

1941] On the Characteristic Values of the Correlation Tensor. 203

between vector sets \{A(t)\} and \{B(t)\}. \rho(A; B) may be called the additive correlation coefficient. If \(n=1\), the additive C.C. is reduced to the square of the Pearsonian correlation coefficient.

§ 7. A Historical Remark. In the course of preparation of our third paper on the measure of correlation between vector quantities, the author found unexpectedly the Sverdrup's paper\(^\text{(11)}\) as quoted in Hukui's "Climatology" (written in Japanese). And Sverdrup's paper teaches us the existence of the Dietzius' paper\(^\text{(12)}\) as written with the same intention as ours.

The Dietzius' paper gives simply the measure of correlation which we have called the primitive C.C.\(^\text{(13)}\). His paper shows us that he presumably intended to give the more general measure of correlation and further to determine the explicit form of an aggregate of coefficients which we have called the regression tensor. However he could not succeed in this task. Probably it would be due to the circumstance that he did not understand this aggregate of coefficients as a tensor and therefore did not endeavour to obtain its tensor form.

Although the Sverdrup's paper has proposed the more general measure of correlation than the Dietzius's one, the former is less general than our S.C.C. \(r^*\) or our additive C.C. \(\rho\). For example, the Sverdrup's correlation coefficient, say \(r\), can not distinguish the orientation of both sets, whereas our \(r^*\) (or the determinant of the regression tensor) can do. When the one set is a mirror image of the other set (a special case of the linear relation!), \(r\) is not necessarily equal to 1, whereas our \(r^*\) (or \(\rho\)) is necessarily equal to -1 (or \(\rho=-1\)).

In his paper the absolute meaning of \(r\) is evident when and only when \(r=1\), for the existence of the function \(f(r)\) conditioned by \(S^\text{c}\) (§ 1) is not yet proved. In ours the meaning of \(r^*\) or \(\rho\) is explained on the basis of the generalized Fourier series and the simple but rough explanation is also given\(^\text{(14)}\). The meaning of \(r=0\) is not clear, since it is not explained under what law of frequency \(r=0\) means the stochastic independency. Furthermore, there exists a removable mistake in his paper. That is to say, he does not use in his calculation of \(r\) the deviation from the mathematical expectation of the given vectors. Of course, he gives some explanation on this point. But if we consider the one dimensional case, his formula is reduced formally to the Pears-


\(\text{(13)}\) The paper I.

\(\text{(14)}\) The paper V. The similar explanation is possible for \(\rho\) viz. the square root of the additive C.C. is equal to \(\sqrt{(p/p+q)}\) under the same assumption as the case of \(r^*\).
and the Pearsonian theory is based upon the deviation, which gives exactly $r=0$ when two sets are stochastically independent under the normal law of frequency. This discrepancy would be more fatal than the one considered by Sverdrup. If we assume the equality of the standard deviations and the mean values of two sets, it is easy to prove that his uncorrected $r$, over-estimates the grade of correlation. In our theory $R=0$ or $\rho=0$ means at least under normal law of frequency the stochastical independency of the sets.

Furthermore it seems very difficult to extend his idea to the multiple correlation, whereas in our theory the multiple correlation can be discussed quite analogously as in the Pearsonian one. The sole difference between the Pearson's and ours would owe to the separation of the two roles of symbols—the one as quality and the other as quantity—in our theory.

By the way we have no excuse for missing in our first paper those paper's of respected precursors who studied more than twenty years ago on the same thema with the same intention as we have done.

The author desires to express his acknowledgement for the facilities given by Dr. T. Okada, the Director of our Observatory, and Prof. Dr. S. Fujiwhara.

Hygieno-Meteorological Section,
Central Meteorological Observatory,
Tokyo, Japan.

(Received Jan. 18, 1941).

(15) The paper VII. Strictly speaking, the complete stochastical independency of two vector sets would imply under the general law of frequency a good deal more than $r=0$ or $\rho=0$. 