HYPERMATRIX AND ITS APPLICATION

By KIN-ICHI YAMADA*

In this paper, the author will consider multidimensional matrices on a line different from that of R. Gouarné and I. Samuel.[1] In the first half, a singular situation shall be pointed out concerning multidimensional matrices and in the second half, an application of a specific kind of them shall be proposed.

§0 Hypermatrices

A setout of numbers on the lattice points $(i, j, \ldots, k, l)$ of an N-dimensional space will be called an N-dimensional matrix of order $I \times J \times \ldots \times K \times L$ or an $I \times J \times \ldots \times K \times L$ hypermatrix, where i ranges over 1, 2, ..., I, j over 1, 2, ..., J, ..., k over 1, 2, ..., K and l over 1, 2, ..., L.

The number $a_{ij.\,kl}$ set out on the point $(i, j, \ldots, k, l)$ will be called the $i \cdot j \cdot \ldots \cdot k \cdot l$ element of the hypermatrix, which as a whole will be denoted by A or by $[a_{ij\,\,kl}]$.

The elements shall be real numbers and the dimension be fixed hereafter.

Equality and inequality of two hypermatrices of the same order will be defined in the same way as usual. Thus they enjoy the usual fundamental properties: reflexivity, symmetricity or antisymmetricity and transitivity.

Addition of two hypermatrices of the same order will be defined also as usual, and it enjoys commutativity and associativity. There exists the additive identity and every hypermatrix has its additive reciprocal.

Scalar multiplication of a hypermatrix by a real number will be defined too as usual. And it enjoys commutativity, associativity and double distributivity.

Let $A = [a_{ij\,.\,kl}]$ be an $I \times J \times \ldots \times K \times L$ hypermatrix and $B = [b_{pq\,\,rs}]$ a $P \times Q \times \ldots \times R \times S$ hypermatrix with $L = P$. Then multiplication of A by B shall be defined by

$$AB = [\sum_{l, q, \ldots, r} a_{ij\,.kl} b_{lq\,.rs}] .$$

Thus it proves easily to be associative and doubly distributive over addition.

§1 Cubic Hypermatrices

An $L \times L \times \ldots \times L \times L$ hypermatrix will be called cubic. The set of all cubic hypermatrices with a specific L forms a ring with respect to addition and multiplication defined above, since it is closed under these operations which possess the fundamental properties as was shown

* Professor (*Kyōju*) in Mathematics.

[1] René Gouarné et Isaac Samuel: Introduction à l'étude des matrices multidimensionnelles. *Cahiers de Physique* 16 (1962), pp. 133–142.

Isaac Samuel et René Gouarné: Determinants des matrices à N dimension d'ordre n. *Cah. de Phys.* 16 (1962), pp. 143–152.

in the preceding section.

How about the multiplicative unit?

It might be natural to introduce the extended Kronecker's delta by

$$\delta_{ij\ \ kl}=\begin{cases}1 & \text{if } i=j=\ldots=k=l \\ 0 & \text{otherwise}\end{cases}$$

Let the hypermatrix $[\delta_{ij\ \ kl}]$ be denoted by E, then we have

(1) $$AE=A$$

for every A. Because

$$\sum_{l,q,\ldots,r} a_{ij\ \ kl}\delta_{lq\ .\ rs}=a_{ij\ldots ks}$$

holds for every N-ple $(i,j,\ldots,k,s)$, where $a_{ij\ldots kl}$ being the general element of A.

Though (1) is an identity in hypermatrix algebra,

(2) $$EA=A$$

is not. Because

$$\sum_{l,q,\ \ ,r} \delta_{ij\ \ kl}a_{lq\ \ rs}$$

vanishes unless $i=j=\ldots=k=l$.

It is a singular situation which does not occur in matrix algebra that (1) is an identity, but (2) is not. Thus E might be called justly a right unit.

How about the uniqueness of right unit?

For the sake of simplicity, two definitions shall be introduced.

Definition I: A cubic hypermatrix U is called a right unit, if

$$AU=A$$

holds for every cubic hypermatrix A.

Definition II: Let $A=[a_{ij\ .\ kl}]$ be a given cubic hypermatrix. The square matrix

$$[\sum_{j,\ \ ,k} a_{ij\ ..kl}]$$

shall be called the contracted matrix of A and denoted by Ctr A.

Then we can prove

[**Theorem 1**] A cubic hypermatrix is a right unit if and only if its contracted matrix is the unit matrix.

Proof: If $U=[u_{ij\ \ kl}]$ is a right unit, then the equality

(3) $$\sum_{l,q,.\ \ ,r} a_{ij\ \ kl}u_{lq\ \ .rs}=a_{ij..\ ks}$$

i.e.

(4) $$\sum_{l}(a_{ij\ \ kl}\sum_{q,\ \ ,r} u_{lq\ldots rs})=a_{ij..\ ks}$$

holds for every $a_{ij\ \ ks}$.

Therefore

(5) $$\sum_{q,\ldots,r} u_{lq\ \ rs}=\delta_{ls}$$

i.e.

(6) $$\text{Ctr } U=E^{(2)}\ ,$$

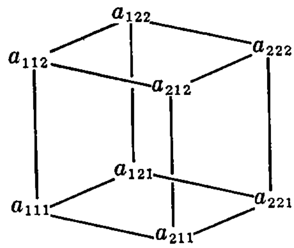
where $E^{(2)}$ is the unit matrix. Thus the condition is necessary.

Conversely, let (6) hold. Then we have elementwise (5), which implies (4) i.e. (3) for every $a_{ij\ldots ks}$. Thus the condition is sufficient, and the proof is completed.

This theorem tells us that right unit is not at all unique. Because L^2 equations (5) have L^N unknowns.

And this is another singular situation concerning hypermatrices.

A $2\times2\times2$ hypermatrix



will be denoted by

$$\left[\begin{array}{cc|cc} a_{111} & a_{121} & a_{112} & a_{122} \\ a_{211} & a_{221} & a_{212} & a_{222} \end{array}\right]$$

for the sake of typographical simplicity.

Example 1 Two right units shall be given for instance.

$$\left[\begin{array}{cc|cc} 3 & 4 & -5 & 4 \\ 2 & 3 & 9 & -7 \end{array}\right]\left[\begin{array}{cc|cc} 0.5 & 0.5 & 2 & -2 \\ -1 & 1 & -3 & 4 \end{array}\right]=\left[\begin{array}{cc|cc} 3 & 4 & -5 & 4 \\ 2 & 3 & 9 & -7 \end{array}\right],$$

$$\left[\begin{array}{cc|cc} 3 & 4 & -5 & 4 \\ 2 & 3 & 9 & -7 \end{array}\right]\left[\begin{array}{cc|cc} 4 & -3 & 1 & -1 \\ -2 & 2 & 0.5 & 0.5 \end{array}\right]=\left[\begin{array}{cc|cc} 3 & 4 & -5 & 4 \\ 2 & 3 & 9 & -7 \end{array}\right].$$

Concerning contracted matrices, we have

[Lemma] The contracted matrix of the product of two cubic hypermatrices is the product of their contracted matrices.

Proof: Let $A=[a_{ij...kl}]$ and $B=[b_{pq..rs}]$ be two given cubic hypermatrices.

Then by definition

$$\begin{aligned} \mathrm{Ctr}\,(AB) &= \mathrm{Ctr}\,[\sum_{l,q,\dots,r} a_{ij.kl}b_{lq..rs}] \\ &= [\sum_{j,\dots,k}(\sum_{l,q,\dots,r} a_{ij\ kl}b_{lq\dots rs})] \\ &= [\sum_{l}(\sum_{j,\dots,k} a_{ij\ kl})(\sum_{q,\dots,r} b_{lq.rs})] \\ &= [\sum_{j,\dots,k} a_{ij\dots kl}][\sum_{q,\dots,r} b_{lq\dots rs}]\,. \end{aligned}$$

The last member is just the product of the contracted matrix of A and that of B. And the lemma is proved.

How about the reciprocal?

For the sake of this problem, let us introduce

Definition III: Let A be a given cubic hypermatrix and U a right unit. If there exists a cubic hypermatrix B such that $AB=U$, then A is said to be nonsingular and B is called a right reciprocal of A.

And we have

[Theorem 2] A cubic hypermatrix is nonsingular if and only if its contracted matrix is nonsingular.

Proof: A cubic hypermatrix A is nonsingular if and only if there exists a cubic hypermatrix B such that

(7) $$AB=U\,,$$

where U is a right unit. On applying the lemma we have

$$\mathrm{Ctr}\, A \cdot \mathrm{Ctr}\, B = \mathrm{Ctr}\, U$$

which combines Theorem 1 to yield

$$\mathrm{Ctr}\, A \cdot \mathrm{Ctr}\, B = E^{(2)} . \tag{8}$$

Thus the nonsingularity of Ctr A is a necessary condition.

Conversely, let Ctr A be a nonsingular matrix. By definition there exists a matrix M such that

$$(\mathrm{Ctr}\, A)M = E^{(2)} .$$

Then we can construct easily a cubic hypermatrix B, whose contracted matrix is M. And we have (8). Application of the lemma on the left member yields

$$\mathrm{Ctr}\, (AB) = E^{(2)} .$$

which combines Theorem 1 to yield (7).

Thus the condition is sufficient and the proof is completed.

(Corollary) If AB is a right unit, then so is BA.

Proof: (8) and commutativity of reciprocal matrices together imply

$$\mathrm{Ctr}\, B \cdot \mathrm{Ctr}\, A = E^{(2)}$$

which is equivalent with

$$BA = U$$

by Theorem 1.

Example 2

$$\left[\begin{array}{cc|cc} 2 & 1 & 3 & 1 \\ -1 & 3 & 2 & 1 \end{array}\right]$$

is nonsingular because of

$$\left[\begin{array}{cc|cc} 2 & 1 & 3 & 1 \\ -1 & 3 & 2 & 1 \end{array}\right]\left[\begin{array}{cc|cc} 1 & 2 & -3 & -1 \\ 1 & -3 & 2 & 1 \end{array}\right]=\left[\begin{array}{cc|cc} 0 & 1 & 1 & -1 \\ -7 & 7 & 10 & -9 \end{array}\right]$$

and

$$\mathrm{Ctr}\left[\begin{array}{cc|cc} 0 & 1 & 1 & -1 \\ -7 & 7 & 10 & -9 \end{array}\right]=\left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right].$$

$$\left[\begin{array}{cc|cc} 1 & 2 & -3 & -1 \\ 1 & -3 & 2 & 1 \end{array}\right]\left[\begin{array}{cc|cc} 2 & 1 & 3 & 1 \\ -1 & 3 & 2 & 1 \end{array}\right]=\left[\begin{array}{cc|cc} -3 & 4 & -5 & 5 \\ 7 & -7 & 10 & -9 \end{array}\right]$$

and

$$\mathrm{Ctr}\left[\begin{array}{cc|cc} -3 & 4 & -5 & 5 \\ 7 & -7 & 10 & -9 \end{array}\right]=\left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right]$$

combine the first half to verify the lemma.

Theorem 2 and the corollary are very close to those of matrix algebra except the uniqueness of reciprocal.

§ 2 Transition Hypermatrices

Let us consider a specific kind of hypermatrices in this section.

Definition IV: A cubic hypermatrix $A=[a_{ij..kl}]$ will be called a transition hypermatrix if $A \geqq O$ and

$$\sum_l a_{ij..kl} = 1 \qquad \text{for all} \quad (i, j, \ldots, k),$$

where O is a cubic hypermatrix with zero elements exclusively.

In order to apply this kind of hypermatrices in somewhat practical situation, the following definition shall be introduced.

Definition V: Let $A=[a_{ij\dots kl}]$ and $B=[b_{ij\dots kl}]$ be transition hypermatrices of order L^N. Then the hypermatrix

$$\left[\frac{1}{L^{N-2}}\sum_{l,q,\dots,r} a_{ij\dots kl}b_{lq\dots rs}\right],$$

will be called the transition product of A and B and denoted by $A*B$.

[Theorem 3] The transition product of two transition hypermatrices is a transition hypermatrix.

Proof: Let A and B be given transition hypermatrices of order L^N.

Due to nonnegativity of the elements and Definition V, we have

(1) $$A*B\geqq O,$$

$$\sum_l a_{ij\dots kl}=1 \quad \text{for all} \quad (i,j,\dots,k)$$

and

$$\sum b_{lq\dots rs}=1 \quad \text{for all} \quad (l,q,\dots,r)$$

together imply that

$$\sum_s\left(\frac{1}{L^{N-2}}\sum_{l,q,\dots,r} a_{ij\dots kl}b_{lq\dots rs}\right)$$
$$=\frac{1}{L^{N-2}}\sum_{q,\dots,r}\left(\sum_l a_{ij\dots kl}\sum_s b_{lq\dots rs}\right)$$
$$=\frac{1}{L^{N-2}}\sum_{q,\dots,r}1.$$

Since each of the $N-2$ numbers $q,\dots,r$ ranges over $1,2,\dots,L$,

$$\sum_{q,\dots,r}1=L^{N-2}.$$

Thus we have

(2) $$\sum_s\left(\frac{1}{L^{N-2}}\sum_{l,q,\dots,r} a_{ij\dots kl}b_{lq\dots rs}\right)=1.$$

(1) and (2) combine to prove that $A*B$ is also a transition hypermatrix.

Again for the sake of application, let us introduce a specific kind of transition hypermatrices by a usual

Definition VI: A transition hypermatrix A will be said to be regular, if there exists a natural number H such that

$$^{H}A>O,$$

where presuperscript denotes a transition power of a transition hypermatrix.

And we can prove the

[Principal Theorem 1] If A is a regular transition hypermatrix, then

$$^{n}A\rightarrow W \quad \text{as} \quad n\rightarrow\infty,$$

where $W=[w_{ij\dots kl}]$ is a transition hypermatrix with such elements as

$$w_{ij\dots kl}=w_l$$

irrespective of $(i,j,\dots,k)$.

Proof: We can assume without loss of generality that A is positive. Because otherwise, there exists a natural number H with $^{H}A>O$, and we can prove that

$${}^{m}({}^{H}A) \to W \quad \text{as} \quad m \to \infty$$

i.e.

$${}^{n}A \to W \quad \text{as} \quad n \to \infty .$$

Let μ be defined by

$$\mu = \min_{i,j,\ldots,k,l} a_{ij\ldots kl} ,$$

then due to the assumption and Definition IV the inequalities

(3) $$0 < \mu < 1$$

hold.

s shall be specified hereafter and M_t, m_t and $a^{(t)}_{ij\ldots ks}$ defined by induction with respect to t.

We will begin with

(4) $$M_0 = \max_{i,j,\ldots,k} a_{ij\ldots ks}$$

(5) $$m_0 = \min_{i,j,\ldots,k} a_{ij\ldots ks} ,$$

(6) $$a^{(1)}_{ij\ldots ks} = \frac{1}{L^{N-2}} \sum_{l,q,\ldots,r} a_{ij\ldots kl} a_{lq\ldots rs} ,$$

(7) $$M_1 = \max_{i,j,\ldots,k} a^{(1)}_{ij\ldots ks} ,$$

and

(8) $$m_1 = \min_{i,j,\ldots,k} a^{(1)}_{ij\ldots ks} .$$

We have

(9) $$m_0 \leqq M_0 \quad \text{and} \quad m_1 \leqq M_1$$

obviously by definitions.

On substituting M_0 for $a_{lq\ldots rs}$ in (6) and simplifying by $\sum_l a_{ij\ldots kl} = 1$, we obtain

$$a^{(1)}_{ij\ldots ks} \leqq M_0$$

which combines (7) to prove

(10) $$M_1 \leqq M_0 .$$

Similarly, substitution of m_0 for $a_{lq\ldots rs}$ in (6) and simplification by $\sum_l a_{ij\ldots kl} = 1$ yield

$$a^{(1)}_{ij\ldots ks} \geqq m_0$$

which is combined with (8) to have

(11) $$m_1 \geqq m_0 .$$

Due to (9), (10) and (11), inequalities

(12) $$m_0 \leqq m_1 \leqq M_1 \leqq M_0$$

hold.

Now we will define $a^{(t)}_{ij\ldots ks}$, M_t and m_t by induction

(13) $$a^{(t)}_{ij\ldots ks} = \frac{1}{L^{N-2}} \sum_{l,q,\ldots,r} a_{ij\ldots kl} a^{(t-1)}_{lq\ldots rs} ,$$

(14) $$M_t = \max_{i,j,\ldots,k} a^{(t)}_{ij\ldots ks} ,$$

(15) $$m_t = \min_{i,j,\ldots,k} a^{(t)}_{ij\ldots ks} .$$

The inequalities (12) mean that inequalities

(16) $$m_0 \leqq m_1 \ldots \leqq m_{t-1} \leqq M_{t-1} \leqq \ldots \leqq M_1 \leqq M_0$$

hold for $t=2$.

By substituting M_{t-1} for $a^{(t-1)}_{lq\ldots rs}$ in (13) and simplifying by $\sum_l a_{ij\ldots kl} = 1$, we have

$$a^{(t)}_{ij..ks} \leqq M_{t-1}$$

which combines (14) to yield

(17) $$M_t \leqq M_{t-1}.$$

Similarly, on substituting m_{t-1} for $a^{(t-1)}_{lq..rs}$ in (13) and applying $\sum_l a_{ij...kl}=1$, it is observed that

$$a^{(t)}_{ij..ks} \geqq m_{t-1}.$$

And we have

(18) $$m_t \geqq m_{t-1}$$

by definition (15).

(16), (17) and (18) together imply that

(19) $$m_0 \leqq m_1 \leqq \dots \leqq m_t \leqq M_t \leqq \dots \leqq M_1 \leqq M_0.$$

Thus (19) has proved to be true for all t.

By definitions, two classes of inequalities

$$a_{ij..kl} - \mu \geqq 0 \qquad \text{for all } (i, j, \dots, k, l)$$

and

$$M_0 - a_{lq..rs} \geqq 0 \qquad \text{for all } (l, q, \dots, r)$$

hold. Therefore we have

$$(a_{ij..kl} - \mu)(M_0 - a_{lq..rs}) \geqq 0$$

i.e.

$$a_{ij..kl}M_0 - \mu M_0 + \mu a_{lq..rs} \geqq a_{ij..kl}a_{lq..rs}$$

for all $(i, j, \dots, k, l)$ and all $(l, q, \dots, r)$.

Summation over $(l, q, \dots, r)$ and division by L^{N-2} yield

$$\frac{M_0}{L^{N-2}} \sum_{l,q,..,r} a_{ij..kl} - L\mu M_0 + \frac{\mu}{L^{N-2}} \sum_{l,q,\dots,r} a_{lq...rs} \geqq \frac{1}{L^{N-2}} \sum_{l,q,\dots,r} a_{ij..kl}a_{lq..rs}.$$

On applying $\sum_l a_{ij...kl}=1$ and (6), we obtain

$$M_0 - L\mu M_0 + \frac{\mu}{L^{N-2}} \sum_{l,q,..,r} a_{lq..rs} \geqq a^{(1)}_{ij..ks} \qquad \text{for all } (i, j, \dots, k).$$

Thus by (7), the inequality

(20) $$(1 - L\mu)M_0 + \frac{\mu}{L^{N-2}} \sum_{l,q,..,r} a_{lq..rs} \geqq M_1$$

holds.

Similarly by definitions, we have two classes of inequalities

$$a_{ij..kl} - \mu \geqq 0 \qquad \text{for all } (i, j, \dots, k, l)$$

and

$$a_{lq..rs} - m_0 \geqq 0 \qquad \text{for all } (l, q, \dots, r),$$

whence we infer that

$$(a_{ij..kl} - \mu)(a_{lq...rs} - m_0) \geqq 0,$$

i.e.

$$a_{ij..kl}m_0 - \mu m_0 + \mu a_{lq..rs} \leqq a_{ij..kl}a_{lq...rs} \qquad \text{for all } (i, j, \dots, k, l) \quad \text{and all } (l, q, \dots, r).$$

We sum up these inequalities over $(l, q, \dots, r)$ and divide by L^{N-2} to obtain

$$\frac{m_0}{L^{N-2}} \sum_{l,q,..,r} a_{ij..kl} - L\mu m_0 + \frac{\mu}{L^{N-2}} \sum_{l,q,..,r} a_{lq...rs} \leqq \frac{1}{L^{N-2}} \sum_{l,b,..,r} a_{ij...kl}a_{lq..rs}.$$

Application of $\sum_l a_{ij...kl}=1$ and (6) yield

$$m_0 - L\mu m_0 + \frac{\mu}{L^{N-2}} \sum_{l,q,\ldots,r} a_{lq\ldots rs} \leqq a^{(1)}_{ij\ldots ks} \quad \text{for all} \quad (i, j \ldots, k).$$

Thus by (8) the inequality

$$(1 - L\mu)m_0 + \frac{\mu}{L^{N-2}} \sum_{l,q,\ldots,r} a_{lq\ldots rs} \leqq m_1 \tag{21}$$

holds.

Subtracting (21) from (20), we obtain

$$(1 - L\mu)(M_0 - m_0) \geqq M_1 - m_1$$

which concludes

$$1 - L\mu \geqq 0$$

and that

$$(1 - L\mu)^t (M_0 - m_0) \geqq M_t - m_t, \tag{22}$$

is true for $t = 1$.

Let us assume that (22) is true for $t = u - 1$ i.e.

$$(1 - L\mu)^{u-1}(M_0 - m_0) \geqq M_{u-1} - m_{u-1}. \tag{23}$$

By definitions, we have two classes of inequalities

$$a_{ij\ldots kl} - \mu \geqq 0 \quad \text{for all} \quad (i, j, \ldots, k, l)$$

and

$$M_{u-1} - a^{(u-1)}_{lq\ldots rs} \geqq 0 \quad \text{for all} \quad (l, q, \ldots, r),$$

whence we obtain

$$(a_{ij\ldots kl} - \mu)(M_{u-1} - a^{(u-1)}_{lq\ldots rs}) \geqq 0$$

i.e.

$$a_{ij\ldots kl} M_{u-1} - \mu M_{u-1} + \mu a^{(u-1)}_{lq\ldots rs} \geqq a_{ij\ldots kl} a^{(u-1)}_{lq\ldots rs} \quad \text{for all} \quad (i, j, \ldots, k, l) \text{ and all } (l, q, \ldots, r).$$

On summing up these inequalities over $(l, q, \ldots, r)$ and dividing by L^{N-2}, we have

$$\frac{M_{u-1}}{L^{N-2}} \sum_{l,q,\ldots,r} a_{ij\ldots kl} - L\mu M_{u-1} + \frac{\mu}{L^{N-2}} \sum_{l,q,\ldots,r} a^{(u-1)}_{lq\ldots rs} \geqq \frac{1}{L^{N-2}} \sum_{l,q,\ldots,r} a_{ij\ldots kl} a^{(u-1)}_{lq\ldots rs},$$

which is combined with $\sum_l a_{ij\ldots kl} = 1$, (13) and (14) to yield

$$(1 - L\mu) M_{u-1} + \frac{\mu}{L^{N-2}} \sum_{l,q,\ldots,r} a^{(u-1)}_{lq\ldots rs} \geqq M_u. \tag{24}$$

Similarly, we have two classes of inequalities

$$a_{ij\ldots kl} - \mu \geqq 0 \quad \text{for all} \quad (i, j, \ldots, k, l)$$

and

$$a^{(u-1)}_{lq\ldots rs} - m_{u-1} \geqq 0 \quad \text{for all} \quad (l, q, \ldots, r)$$

by definitions.

We combine them to obtain

$$(a_{ij\ldots kl} - \mu)(a^{(u-1)}_{lq\ldots rs} - m_{u-1}) \geqq 0$$

i.e.

$$a_{ij\ldots kl} m_{u-1} - \mu m_{u-1} + \mu a^{(u-1)}_{lq\ldots rs} \leqq a_{ij\ldots kl} a^{(u-1)}_{lq\ldots rs} \quad \text{for all } (i, j, \ldots, k, l) \text{ and all } (l, q, \ldots, r).$$

Summation over $(l, q, \ldots, r)$ and division by L^{N-2} yield

$$\frac{m_{u-1}}{L^{N-2}} \sum_{l,q,\ldots,r} a_{ij\ldots kl} - L\mu m_{u-1} + \frac{\mu}{L^{N-2}} \sum_{l,q,\ldots,r} a^{(u-1)}_{lq\ldots rs} \leqq \frac{1}{L^{N-2}} \sum_{l,q,\ldots,r} a_{ij\ldots kl} a^{(u-1)}_{lq\ldots rs}$$

which we simplify into

(25) $$(1-L\mu)m_{u-1}+\frac{\mu}{L^{N-2}}\sum_{l,q,\cdots,r} a_{lq\cdots rs}^{(u-1)} \leqq m_u$$

by $\sum_l a_{ij\ldots kl}=1$, (13) and (15).

Subtraction of (25) from (24) yields

(26) $$(1-L\mu)(M_{u-1}-m_{u-1}) \geqq M_u - m_u .$$

(23) and (26) together with $(1-L\mu)>0$ imply that

$$(1-L\mu)^u(M_0-m_0) \geqq M_u-m_u$$

i.e. (22) is true for $t=u$.

Thus the general validity of (22) is proved by induction.

From (1) and (22) we can infer that

(27) $$M_t-m_t \to 0 \quad \text{as} \quad t\to\infty ,$$

since $0 \leqq 1-L\mu < 1$.

(19) and (27) together conclude that $a_{ij\cdot\cdot kl}^{(t)}$ is the $i\cdot j\cdot\ldots\cdot k\cdot l$ element of ${}^tA$, and there exists a number w_l for each l such that

$$a_{ij\cdot\cdot kl}^{(t)} \to w_l \quad \text{as} \quad t\to\infty ,$$

independent of $(i, j, \ldots, k)$.

And this is what is to be proved.

§ 3 An Application

Let us assume that there are L states given and from every $N-1$ ple combination of L states one of the L states may result. If the probability that the state l may result from a state-combination $(i, j, \ldots, k)$ is $a_{ij\ldots kl}$ with $\sum_l a_{ij\cdot\cdot kl}=1$, then this situation might be realised compactly by the L^N transition hypermatrix $[a_{ij\ldots kl}]$.

Example 3 The following table gives the percentage of children with bright-coloured hair.

FATHER \ MOTHER	BRIGHT	DARK
BRIGHT	80	50
DARK	30	60

This is realised by 2^3 transition hypermatrix

$$\left[\begin{array}{cc|cc} 0.8 & 0.5 & 0.2 & 0.5 \\ 0.3 & 0.6 & 0.7 & 0.4 \end{array}\right].$$

The following theorem is very important in this connection.

[Theorem 4] The $i, j, \ldots, k, l$ element of ${}^nA$ gives the probability that the process which has started in a state-combination $(i, j, \ldots, k)$ will be in a state l after n steps provided that every state-combination is equally likely.

Proof: The theorem is true for $n=1$ by the above-given interpretation of a transition hypermatrix.

Assume that the theorem is true for $n=m-1$, and let ${}^rA$ be denoted by $[a_{ij\cdot\cdot kl}^{(r)}]$.

The $i\cdot j\cdot\ldots\cdot k\cdot l$ element of ${}^mA$ is

(1) $$a^{(m)}_{ij\ \ kl}=\frac{1}{L^{N-2}}\sum_{p,q,\ \ ,r} a^{(m-1)}_{ij\ .\ kp}a_{pq\ \ rl}\ ,$$

Since by assumption $a^{(m-1)}_{ij\ .kp}$ is the probability that the process starting in a state-combination $(i,j,...,k)$ will be in a state p after $m-1$ steps, $1/L^{N-2}$ is that of p to be combined with $(q,...,r)$ and $a_{pq\ \ rl}$ is that of $(p,q,...,r)$ to be followed by a state l on the m-th step, each summand of (1) is the probability that the process which started in a state-combination $(i,j,...,k)$ will be in state l through a state-combination $(p,q,...,r)$ after m steps. Such summands being summed up over $(p,q,...,r)$ give the probability that the process which has started in a state-combination $(i,j,...,k)$ will be in a state l after m steps. Thus the theorem is true for $n=m$.

And the proof is completed by induction.

Example 4 (Continued)

$$^{2}\left[\begin{array}{cc|cc}0.8 & 0.5 & 0.2 & 0.5\\ 0.3 & 0.6 & 0.7 & 0.4\end{array}\right]=\left[\begin{array}{cc|cc}0.61 & 0.55 & 0.39 & 0.45\\ 0.51 & 0.57 & 0.49 & 0.43\end{array}\right],$$

tells that 55% of the grandchildren of bright-haired fathers and dark-haired mothers may be bright-haired.

Implication of Principal Theorem 1 is now very important, since it tells that after a large number of steps, the probability of the process to be in a specific state l will be nearly w_l no matter what the initial state-combination may be.

This situation shall be explained by

Example 5 (Continued)

$$^{3}\left[\begin{array}{cc|cc}0.8 & 0.5 & 0.2 & 0.5\\ 0.3 & 0.6 & 0.7 & 0.4\end{array}\right]=\left[\begin{array}{cc|cc}0.572 & 0.56 & 0.428 & 0.44\\ 0.552 & 0.564 & 0.448 & 0.436\end{array}\right],$$

$$^{4}\left[\begin{array}{cc|cc}0.8 & 0.5 & 0.2 & 0.5\\ 0.3 & 0.6 & 0.7 & 0.4\end{array}\right]=\left[\begin{array}{cc|cc}0.5644 & 0.562 & 0.4356 & 0.438\\ 0.5604 & 0.5628 & 0.4396 & 0.4372\end{array}\right],$$

$$^{5}\left[\begin{array}{cc|cc}0.8 & 0.5 & 0.2 & 0.5\\ 0.3 & 0.6 & 0.7 & 0.4\end{array}\right]=\left[\begin{array}{cc|cc}0.56288 & 0.5624 & 0.43712 & 0.4376\\ 0.56208 & 0.5626 & 0.43792 & 0.4374\end{array}\right].$$

The next problem to be attacked is

how to get W in Principal Theorem 1.

Definition VII Let A be a given transition hypermatrix. A transition hypermatrix F for which

$$F*A=F,$$

holds will be called a fixed transition hypermatrix of A.

Now we can prove

[Principal Theorem 2] W in Principal Theorem 1 is the unique fixed transition hypermatrix of A.

Proof: By Principal Theorem 1, we have

(2) $$^{n}A\to W \quad \text{as} \quad n\to\infty.$$

Whence we obtain

(3) $$^{n+1}A\to W*A \quad \text{as} \quad n\to\infty.$$

(2) combined with (3) yields

$$W*A-W\to 0.$$

Since two hypermatrices in the left member are constant hypermatrices, we have

$$W*A=W,$$

which means that W is a fixed transition hypermatrix of A.

Let F be a fixed transition hypermatrix, then we have

$$F*A=F.$$

By successive postmultiplication of A we obtain

$$F*^nA=F,$$

which combines (2) to yield

$$F\to F*W.$$

Since both members are constant, we infer that

$$F=F*W$$

i.e.

$$f_{ij\ldots kl}=\frac{1}{L^{N-2}}\sum_{p,q,\ldots,r} f_{ij\ldots kq}w_{pq\ldots rl} \qquad \text{for all} \quad (i,j,\ldots,k,l).$$

On applying $w_{pq\ldots rl}=w_l$, it is observed that

$$\frac{1}{L^{N-2}}\sum_{p,q,\ldots,r} f_{ij\ldots kp}w_{pq\ldots rl}=\frac{w_l}{L^{N-2}}\sum_{p,q,\ldots,r} f_{ij\ldots kp}.$$

Simplification by $\sum_p f_{ij\ldots kp}=1$ yields

$$\frac{w_l}{L^{N-2}}\sum_{p,q,\ldots,r} f_{ij\ldots kp}=w_l.$$

These three equalities combine to prove

$$f_{ij\ldots kl}=w_l \qquad \text{for all} \quad (i,j,\ldots,k,l)$$

which is equivalent with

$$F=W.$$

Thus the uniqueness is proved to complete the proof.

By Principal Theorem 2, we can evaluate the ultimate distribution of states, as shown in Example 6 (Continued)

$$\left[\begin{array}{cc|cc} w_1 & w_1 & w_2 & w_2 \\ w_1 & w_1 & w_2 & w_2 \end{array}\right] * \left[\begin{array}{cc|cc} 0.8 & 0.5 & 0.2 & 0.5 \\ 0.3 & 0.6 & 0.7 & 0.4 \end{array}\right] = \left[\begin{array}{cc|cc} w_1 & w_1 & w_2 & w_2 \\ w_1 & w_1 & w_2 & w_2 \end{array}\right]$$

gives

$$\begin{cases} 0.65w_1+0.45w_2=w_1 \\ 0.35w_1+0.55w_2=w_2 \\ \quad w_1+ \quad w_2=1 \end{cases}.$$

And we have

$$w_1=0.5625, \qquad w_2=0.4375$$

which means that

$${}^n\left[\begin{array}{cc|cc} 0.8 & 0.5 & 0.2 & 0.5 \\ 0.3 & 0.6 & 0.7 & 0.4 \end{array}\right] \to \left[\begin{array}{cc|cc} 0.5625 & 0.5625 & 0.4375 & 0.4375 \\ 0.5625 & 0.5625 & 0.4375 & 0.4375 \end{array}\right] \qquad \text{as} \quad n\to\infty.$$

§ 4 Concluding Remarks

Thus far, we have considered the concept "hypermatrix" from applicative point of view on Markov chain. And some relevant results have been reached as far as regular Markov chain concerns. In § 1, however, some singular situations were referred to concerning unit and reciprocal hypermatrices. Therefore we are little ready for absorbing Markov chain. The next problems are to elaborate definitions and to arrange theorems thereof.

1965-5-10