<table>
<thead>
<tr>
<th>Title</th>
<th>Hypermatrix and its Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Yamada, Kin-ichi</td>
</tr>
<tr>
<td>Citation</td>
<td>Hitotsubashi journal of arts and sciences, 6(1): 34-44</td>
</tr>
<tr>
<td>Issue Date</td>
<td>1965-09</td>
</tr>
<tr>
<td>Type</td>
<td>Departmental Bulletin Paper</td>
</tr>
<tr>
<td>Text Version</td>
<td>publisher</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://doi.org/10.15057/4453">http://doi.org/10.15057/4453</a></td>
</tr>
</tbody>
</table>
In this paper, the author will consider multidimensional matrices on a line different from that of R. Gourné and I. Samuel. 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, \ldots, \(I\), \(j\) over 1, 2, \ldots, \(J\), \ldots, \(k\) over 1, 2, \ldots, \(K\) and \(l\) over 1, 2, \ldots, \(L\).

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

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

\[
AB = \sum_{i, q, \ldots, r} a_{ij\ldots kl} b_{pq\ldots rs}.
\]

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

§ 1 Cubic Hypermatrices

An \(L \times L \times \ldots \times L\) 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.


in the preceding section.

How about the multiplicative unit?

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

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

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

(1) $$AE = A$$

for every $A$. Because

$$\sum_{i,j,k} a_{ij} \delta_{il} = a_{ij,k}$$

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

Though (1) is an identity in hypermatrix algebra,

(2) $$EA = A$$

is not. Because

$$\sum_{i,j,k} \delta_{ij} \equiv k a_{il} r_s$$

ever 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 $A=[a_{ij} \equiv k]$ be a given cubic hypermatrix. The square matrix $[ \sum_{j,k} a_{ij} \equiv k ]$

shall be called the contracted matrix of $A$ and denoted by $\text{Ctr } A$.

Then we can prove

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

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

(3) $$\sum_{i,j,k} a_{ij} \equiv k u_{il} r_s = a_{ij,k}$$

i.e.

(4) $$\sum_{i,j,k} (a_{ij} \equiv k \sum_{q,r} u_{q,r} r_s) = a_{ij,k}$$

holds for every $a_{ij,k}$.

Therefore

(5) $$\sum_{q,r} u_{q,r} r_s = \delta_{il}$$

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 $a_{ij,k}$. 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 $2 \times 2 \times 2$ hypermatrix

$$
\begin{bmatrix}
  a_{111} & a_{121} & a_{122} \\
  a_{211} & a_{221} & a_{222}
\end{bmatrix}
$$

will be denoted by

$$
\begin{bmatrix}
  a_{111} & a_{121} & a_{122} & a_{123} \\
  a_{211} & a_{221} & a_{222} & a_{223}
\end{bmatrix}
$$

for the sake of typographical simplicity.

Example 1 Two right units shall be given for instance.

$$
\begin{bmatrix}
  3 & 4 & -5 & 4 & 0.5 & 0.5 & 2 & -2 \\
  2 & 3 & 9 & -7 & -1 & 1 & -3 & 4
\end{bmatrix}
= \begin{bmatrix}
  3 & 4 & -5 & 4 & 0.5 & 0.5 & 2 & -2 \\
  2 & 3 & 9 & -7 & -1 & 1 & -3 & 4
\end{bmatrix}
$$

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 $A=\begin{bmatrix} a_{ij} \ldots \end{bmatrix}$ and $B=\begin{bmatrix} b_{pq} \ldots \end{bmatrix}$ be two given cubic hypermatrices. Then by definition

$$
\text{Ctr}(AB) = \text{Ctr}\left[ \sum_{i,j} a_{ij} \sum_{k,l} b_{kl} \right]
$$

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
\[
\text{Ctr } A \cdot \text{Ctr } B = \text{Ctr } U
\]
which combines Theorem 1 to yield
\[
\text{Ctr } A \cdot \text{Ctr } B = E^{(2)}.
\]
Thus the nonsingularity of \( \text{Ctr } A \) is a necessary condition.

Conversely, let \( \text{Ctr } A \) be a nonsingular matrix. By definition there exists a matrix \( M \) such that
\[
(C\text{tr } A)M = E^{(2)}.
\]
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
\[
\text{Ctr } (AB) = E^{(2)}
\]
which combines Theorem 1 to yield (7).

Therefore 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
\[
\text{Ctr } B \cdot \text{Ctr } A = E^{(2)}
\]
which is equivalent with
\[
BA = U
\]
by Theorem 1.

Example 2
\[
\begin{bmatrix}
  2 & 1 & 3 & 1 \\
-1 & 3 & 2 & 1 \\
\end{bmatrix}
\]
is nonsingular because of
\[
\begin{bmatrix}
  2 & 1 & 3 & 1 \\
-1 & 3 & 2 & 1 \\
\end{bmatrix}
\begin{bmatrix}
  1 & 2 & -3 & -1 \\
  1 & -3 & -2 & -1 \\
\end{bmatrix} =
\begin{bmatrix}
  1 & 0 \\
  10 & 9 \\
\end{bmatrix}
\]
and
\[
\text{Ctr }
\begin{bmatrix}
  0 & 1 & 1 & -1 \\
-7 & 7 & 10 & -9 \\
\end{bmatrix} =
\begin{bmatrix}
  1 & 0 \\
0 & 1 \\
\end{bmatrix}
\]
and
\[
\text{Ctr }
\begin{bmatrix}
  2 & 1 & 3 & 1 \\
  1 & 2 & -3 & -1 \\
-1 & 3 & 2 & 1 \\
\end{bmatrix} =
\begin{bmatrix}
  -3 & 4 & -5 & 5 \\
  7 & -7 & 10 & -9 \\
\end{bmatrix}
\]
and
\[
\text{Ctr }
\begin{bmatrix}
  -3 & 4 & -5 & 5 \\
  7 & -7 & 10 & -9 \\
\end{bmatrix} =
\begin{bmatrix}
  1 & 0 \\
0 & 1 \\
\end{bmatrix}
\]
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.

\section{2 Transition Hypermatrices}

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

Definition IV: A cubic hypermatrix \( A = [a_{ij..k}] \) will be called a transition hypermatrix if
\[
a_{ij..k} \geq 0
\]
and
\[
\sum_{i} a_{ij..k} = 1 \quad \text{for all } (i, j, ..., k),
\]
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 \( A = [a_{ij\ldots k}] \) and \( B = [b_{ij\ldots k}] \) be transition hypermatrices of order \( L^N \). Then the hypermatrix

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

will be called the transition product of \( A \) and \( B \) and denoted by \( A \ast B \).

[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

\[
\sum_i a_{ij\ldots kl} = 1 \quad \text{for all } (i, j, \ldots, k)
\]

and

\[
\sum_{l} b_{lq\ldots rs} = 1 \quad \text{for all } (l, q, \ldots, r)
\]

together imply that

\[
\sum_s \left( \frac{1}{L^{N-2}} \sum_{i, q, \ldots, r} a_{ij\ldots kl} b_{lq\ldots rs} \right) = \frac{1}{L^{N-2}} \sum_{q, \ldots, r} 1.
\]

Since each of the \( N-2 \) numbers \( q, \ldots, r \) ranges over \( 1, 2, \ldots, L \),

\[
\sum_{q, \ldots, r} 1 = L^{N-2}.
\]

Thus we have

\[
\sum_s \left( \frac{1}{L^{N-2}} \sum_{i, q, \ldots, r} a_{ij\ldots kl} b_{lq\ldots rs} \right) = 1.
\]

(1) and (2) combine to prove that \( A \ast B \) 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

\[
^H A > 0,
\]

where superscript denotes a transition power of a transition hypermatrix.

And we can prove the

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

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

where \( W = [w_{ij\ldots kl}] \) is a transition hypermatrix with such elements as

\[
w_{ij\ldots kl} = w_i
\]

irrespective of \( (i, j, \ldots, k) \).

**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
\[ n^{(HA)} \to W \quad \text{as} \quad m \to \infty \]

\[ nA \to W \quad \text{as} \quad n \to \infty . \]

Let \( \mu \) be defined by

\[
\mu = \min_{i,j,k} a_{ij,kl},
\]

then due to the assumption and Definition IV the inequalities

\[
0 < \mu < 1
\]

hold.

We will begin with

\[
M_0 = \max_{i,j,k} a_{ij,kl}
\]

\[
m_0 = \min_{i,j,k} a_{ij,kl},
\]

\[
a_{ij,kl}^{(0)} = \frac{1}{L^{n-2}} \sum_{q,r,s} a_{ij,kl} a_{qr,ls},
\]

\[
M_1 = \max_{i,j,k} a_{ij,kl},
\]

\[
m_1 = \min_{i,j,k} a_{ij,kl},
\]

and

\[
M_0 \leq M_1 \quad \text{and} \quad m_0 \leq m_1.
\]

We have

\[
m_0 \leq M_0 \quad \text{and} \quad m_1 \leq M_1
\]

obviously by definitions.

On substituting \( M_0 \) for \( a_{ij,kl} \) in (6) and simplifying by \( \sum a_{ij,kl} = 1 \), we obtain

\[
a_{ij,kl}^{(0)} \leq M_0
\]

which combines (7) to prove

\[
M_1 \leq M_0.
\]

Similarly, substitution of \( m_0 \) for \( a_{ij,kl} \) in (6) and simplification by \( \sum a_{ij,kl} = 1 \) yield

\[
a_{ij,kl}^{(0)} \geq m_0
\]

which is combined with (8) to have

\[
m_1 \leq m_0.
\]

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

\[
m_0 \leq m_1 \leq M_1 \leq M_0
\]

hold.

Now we will define \( a_{ij,kl}^{(t)}, M_t \) and \( m_t \) by induction

\[
a_{ij,kl}^{(t)} = \frac{1}{L^{n-2}} \sum_{q,r,s} a_{ij,kl} a_{qr,ls},
\]

\[
M_t = \max_{i,j,k} a_{ij,kl},
\]

\[
m_t = \min_{i,j,k} a_{ij,kl},
\]

The inequalities (12) mean that inequalities

\[
m_0 \leq m_1 \leq \cdots \leq m_{t-1} \leq M_{t-1} \leq \cdots \leq M_1 \leq M_0
\]

hold for \( t = 2 \).

By substituting \( M_{t-1} \) for \( a_{ij,kl}^{(t-1)} \) in (13) and simplifying by \( \sum a_{ij,kl} = 1 \), we have
which combines (14) to yield
(17)  \[ M_t \leq M_{t-1}. \]
Similarly, on substituting \( m_{t-1} \) for \( a(t-1) \) in (13) and applying \( \sum_{t} a_{ij...kl} = 1 \), it is observed that
\[ a_{ij...k}\equiv m_{t-1}. \]
And we have
(18)  \[ m_t \geq m_{t-1}. \]
by definition (15).

(16), (17) and (18) together imply that
(19)  \[ m_0 \leq m_1 \leq \ldots \leq m_t \leq M_1 \leq \ldots \leq M_t \leq M_0. \]
Thus (19) has proved to be true for all \( t \).

By definitions, two classes of inequalities
\[ a_{ij...k} - \mu \geq 0 \quad \text{for all (} i, j, \ldots, k, l \text{)} \]
and
\[ M_0 - a_{lq...rs} \geq 0 \quad \text{for all (} l, q, \ldots, r \text{)} \]
hold. Therefore we have
\[ (a_{ij...k} - \mu)(M_0 - a_{lq...rs}) \geq 0 \]
i.e.
\[ a_{ij...k} M_0 - \mu M_0 + \mu a_{lq...rs} \geq a_{ij...k} a_{lq...rs} \]
for all (\( i, j, \ldots, k, l \)) and all (\( l, q, \ldots, r \)).

Summation over (\( l, q, \ldots, r \)) and division by \( L^{N-2} \) yield
\[ \frac{M_0}{L^{N-2}} \sum_{l, q, \ldots, r} a_{ij...k} - L \mu M_0 + \mu \frac{L^{N-2}}{L^{N-2}} \sum_{l, q, \ldots, r} a_{lq...rs} \geq \frac{1}{L^{N-2}} \sum_{l, q, \ldots, r} a_{ij...k} a_{lq...rs}. \]

On applying \( \sum_{t} a_{ij...k} = 1 \) and (6), we obtain
\[ M_0 - L \mu M_0 + \mu \frac{L^{N-2}}{L^{N-2}} \sum_{l, q, \ldots, r} a_{lq...rs} \geq a_{ij...k} \quad \text{for all (} i, j, \ldots, k \text{)} . \]
Thus by (7), the inequality
(20)  \[ (1 - L \mu) M_0 + \frac{\mu}{L^{N-2}} \sum_{l, q, \ldots, r} a_{lq...rs} \geq M_t \]
holds.

Similarly by definitions, we have two classes of inequalities
\[ a_{ij...k} - \mu \geq 0 \quad \text{for all (} i, j, \ldots, k, l \text{)} \]
and
\[ a_{lq...rs} - m_0 \geq 0 \quad \text{for all (} l, q, \ldots, r \text{)} , \]
whence we infer that
\[ (a_{ij...k} - \mu)(a_{lq...rs} - m_0) \geq 0 , \]
i.e.
\[ a_{ij...k} m_0 - \mu m_0 + \mu a_{lq...rs} \geq a_{ij...k} a_{lq...rs} \quad \text{for all (} i, j, \ldots, k, l \text{)} \quad \text{and all (} l, q, \ldots, r \text{)} . \]

We sum up these inequalities over (\( l, q, \ldots, r \)) and divide by \( L^{N-2} \) to obtain
\[ \frac{m_0}{L^{N-2}} \sum_{l, q, \ldots, r} a_{ij...k} - L \mu m_0 + \mu \frac{L^{N-2}}{L^{N-2}} \sum_{l, q, \ldots, r} a_{lq...rs} \geq \frac{1}{L^{N-2}} \sum_{l, q, \ldots, r} a_{ij...k} a_{lq...rs} . \]
Application of \( \sum_{t} a_{ij...k} = 1 \) and (6) yield
Thus by (8) the inequality

\[(1-L\mu)m_o+\frac{\mu}{L^{N-2}} \sum \sigma_{q...r} \leq m_1 \]

holds.

Subtracting (21) from (20), we obtain

\[(1-L\mu)(m_o-m_o) \leq m_1 - m_1 \]

which concludes

\[1-L\mu \geq 0 \]

and that

\[(1-L\mu)(m_o-m_o) \leq m_1 - m_1 \]

is true for \(t=1\).

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

\[(1-L\mu)^{u-1}(m_o-m_o) \leq m_{u-1} - m_{u-1} \]

By definitions, we have two classes of inequalities

\[a_{ij, ki} - \mu \geq 0 \quad \text{for all } (i, j, k, l) \]

and

\[a_{ij, ki} - \mu \geq 0 \quad \text{for all } (i, j, k, l) \]

whence we obtain

\[(a_{ij, ki} - \mu)(m_{u-1} - (a_{ij, ki})^{u-1}) \geq 0 \]

i.e.

\[a_{ij, ki}m_{u-1} - \mu m_{u-1} + \mu a_{ij, ki}^{(u-1)} \geq a_{ij, ki} a_{(i, j, k, l)}^{(u-1)} \quad \text{for all } (i, j, k, l) \]

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

\[\frac{M_{u-1}}{L^{N-2}} \sum \sigma_{q...r} a_{ij, ki} - L\mu M_{u-1} + \frac{\mu}{L^{N-2}} \sum \sigma_{q...r} a_{ij, ki}^{(u-1)} \geq 1 \sum \sigma_{q...r} a_{ij, ki} a_{(i, j, k, l)}^{(u-1)}, \]

which is combined with \(a_{ij, ki} = 1, (13) \) and \(14) \) to yield

\[(1-L\mu)M_{u-1} + \frac{\mu}{L^{N-2}} \sum \sigma_{q...r} a_{ij, ki}^{(u-1)} \geq 0 \]

Similarly, we have two classes of inequalities

\[a_{ij, ki} - \mu \geq 0 \quad \text{for all } (i, j, k, l) \]

and

\[d_{ij, ki}^{(u-1)} - m_{u-1} \geq 0 \quad \text{for all } (l, q, ..., r) \]

by definitions.

We combine them to obtain

\[(a_{ij, ki} - \mu)(d_{ij, ki}^{(u-1)} - m_{u-1}) \geq 0 \]

i.e.

\[a_{ij, ki}m_{u-1} - \mu m_{u-1} + \mu d_{ij, ki}^{(u-1)} \geq a_{ij, ki} d_{ij, ki}^{(u-1)} \quad \text{for all } (i, j, k, l) \]

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

\[\frac{m_{u-1}}{L^{N-2}} \sum \sigma_{q...r} a_{ij, ki} - L\mu m_{u-1} + \frac{\mu}{L^{N-2}} \sum \sigma_{q...r} d_{ij, ki}^{(u-1)} \geq 1 \sum \sigma_{q...r} a_{ij, ki} d_{ij, ki}^{(u-1)}, \]

which we simplify into
Subtraction of (25) from (24) yields 

\[(1-L\mu)(M_{u-1}-m_{u-1}) \geq M_u - m_u.\]  

(23) and (26) together with 

\[(1-L\mu)>0 \implies (1-L\mu)^u(M_u - m_u) \geq M_u - m_u\]  

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

\[M_t - m_t \to 0 \quad \text{as} \quad t \to \infty,\]

since \(0 \leq 1-L\mu < 1\).

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

\[a_{ij,...,k}^{(t)} \to w_t \quad \text{as} \quad t \to \infty,\]

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

And this is what is to be proved.

\textbf{§ 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 \((i,j,...,k)\) is \(a_{ij,...,k}^{(t)}\) with \(\sum a_{ij,...,k} = 1\), then this situation might be realised compactly by the \(L^N\) transition hypermatrix \([a_{ij,...,k}^{(t)}]\).

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

<table>
<thead>
<tr>
<th>MOTHER</th>
<th>BRIGHT</th>
<th>DARK</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRIGHT</td>
<td>80</td>
<td>50</td>
</tr>
<tr>
<td>DARK</td>
<td>30</td>
<td>60</td>
</tr>
</tbody>
</table>

This is realised by \(2^3\) transition hypermatrix

\[
\begin{bmatrix}
0.8 & 0.5 & 0.2 & 0.5 \\
0.3 & 0.6 & 0.7 & 0.4
\end{bmatrix}.
\]

The following theorem is very important in this connection.

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

\textbf{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 \(^mA\) be denoted by \([a_{ij,...,k}^{(m)}]\).

The \(i\cdot j\cdot ...\cdot k\cdot l\) element of \(^mA\) is
Since by assumption \( a^{(m-1)k}_{ipq} \) is the probability that the process starting in a state-combination \((i,j,\ldots,k)\) will be in a state \(p\) after \(m-1\) steps, \(1/L^{m-2}\) is that of \(p\) to be combined with \((q,\ldots,r)\) and \(a_{pq}^{r}\) is that of \((p,q,\ldots,r)\) 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 \((i,j,\ldots,k)\) will be in state \(l\) through a state-combination \((p,q,\ldots,r)\) after \(m\) steps. Such summands being summed up over \((p,q,\ldots,r)\) give the probability that the process which has started in a state-combination \((i,j,\ldots,k)\) 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)

\[
\begin{bmatrix}
0.8 & 0.5 & 0.2 & 0.5 \\
0.3 & 0.6 & 0.7 & 0.4
\end{bmatrix}
\begin{bmatrix}
0.61 & 0.55 & 0.39 & 0.45 \\
0.51 & 0.57 & 0.49 & 0.43
\end{bmatrix}
\]
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)

\[
\begin{bmatrix}
0.8 & 0.5 & 0.2 & 0.5 \\
0.3 & 0.6 & 0.7 & 0.4
\end{bmatrix}
\begin{bmatrix}
0.572 & 0.56 & 0.428 & 0.44 \\
0.552 & 0.564 & 0.448 & 0.436
\end{bmatrix}
\]

\[
\begin{bmatrix}
0.8 & 0.5 & 0.2 & 0.5 \\
0.3 & 0.6 & 0.7 & 0.4
\end{bmatrix}
\begin{bmatrix}
0.5644 & 0.562 & 0.4356 & 0.438 \\
0.5604 & 0.5628 & 0.4396 & 0.4372
\end{bmatrix}
\]

\[
\begin{bmatrix}
0.8 & 0.5 & 0.2 & 0.5 \\
0.3 & 0.6 & 0.7 & 0.4
\end{bmatrix}
\begin{bmatrix}
0.56288 & 0.5624 & 0.43712 & 0.4376 \\
0.56208 & 0.5626 & 0.43792 & 0.4374
\end{bmatrix}
\]

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) \(A^nW\) as \(n\to\infty\).

Whence we obtain

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

(2) combined with (3) yields

\[
W*A=W,
\]

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

$$F^n A = F,$$

which combines (2) to yield

$$F = F*W.$$ 

Since both members are constant, we infer that

$$F = F*W$$

i.e.

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

for all $(i,j, ..., k, l)$.

On applying $w_{pq...r} = w_1$, it is observed that

$$\frac{1}{L^{N-2}} \sum_{p,q,...,r} f_{ij...kp} w_{pq...r} = \frac{w_1}{L^{N-2}} \sum_{p,q,...,r} f_{ij...kp}.$$

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

$$\frac{w_1}{L^{N-2}} \sum_{p,q,...,r} f_{ij...kp} = w_1.$$

These three equalities combine to prove

$$f_{ij...kl} = w_1$$

for all $(i,j, ..., k, l)$

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)

$$\begin{bmatrix} w_1 & w_1 & w_2 & w_2 \\ w_1 & w_1 & w_2 & w_2 \end{bmatrix} \begin{bmatrix} 0.8 & 0.5 & 0.2 & 0.5 \\ 0.3 & 0.6 & 0.7 & 0.4 \end{bmatrix} = \begin{bmatrix} w_1 & w_1 & w_2 & w_2 \end{bmatrix}$$

gives

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

And we have

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

which means that

$$n \begin{bmatrix} 0.8 & 0.5 & 0.2 & 0.5 \\ 0.3 & 0.6 & 0.7 & 0.4 \end{bmatrix} \rightarrow \begin{bmatrix} 0.5625 & 0.5625 & 0.4375 & 0.4375 \\ 0.5625 & 0.5625 & 0.4375 & 0.4375 \end{bmatrix}$$

as $n \rightarrow \infty$.

§ 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