Consecutive 1's propertyについて

1. C性とは

標題の英語をいったいどう訳したら良いのか，これが大変にむずかしいのである．もし consecutive を辞書にあるとおりに連続と訳すならば，それは数学によく出てくる continuous とまぎらわしくなり，以下読んでいただくとすぐわかるように，この二つは全然違うものである．そこで，同じことを人によっては consecutive retrieval property ともいい，これを略して単に CR 性と訳される方もあるので，それにならって以下では上記をC性と称することにする．ちなみに，CR 性というのは情報検索に応用されることからつけられた言葉である．

さてC性の内容をまず例によって説明する．**図1**および**図2**は，いわゆる01行列である．行列の要素が0か1であるから普通そうよんでいる．各図左側の小文字は行をあらわし，また上部の大文字は列を示している．この二つの行列を見比べると，それは行の順序が異なっているだけである．すなわち，図2の行列は，図1の行列を横に切って，*caebd* の順に並べかえたものである．ところが，図1では1のあらわれ方が，縦に見るとばらばらであるのに対して，図2ではすべての列において1が引き続いて，かたまって並んでいる．まさに consecutive である．このように，ある01行列の行を入れかえることによって，各列とも1がつづいて並ぶようにできるならば，その行列はC性をもつという．したがって図1また当然に図2の各行列はC性を有している．しかし，**図3**の行列はどうやっても1が各列で続くようにはできないことが容易にわかるであろう．すなわち，この行列はC性をもたないのである．

$$\begin{array}{c|ccccc|}
 & A & B & C & D & E \\
\hline
a & 1 & 0 & 1 & 1 & 0 \\
b & 0 & 1 & 0 & 0 & 1 \\
c & 0 & 0 & 1 & 1 & 0 \\
d & 0 & 1 & 0 & 0 & 0 \\
e & 1 & 0 & 1 & 0 & 1 \\
\hline
\end{array}$$

図1 もとの行列

$$\begin{array}{c|ccccc|}
 & A & B & C & D & E \\
\hline
c & 0 & 0 & 1 & 1 & 0 \\
a & 1 & 0 & 1 & 1 & 0 \\
e & 1 & 0 & 1 & 0 & 1 \\
b & 0 & 1 & 0 & 0 & 1 \\
d & 0 & 1 & 0 & 0 & 0 \\
\hline
\end{array}$$

図2 入れかえ後

C性という問題が，いつ頃から，またどういうことに関連して出現したのであるか？ そのくわしいことはよく知らない．どうもはじめは遺伝学で取り扱われたようである[3]．ついで，この問題を考察する必要性，またその応用の可能性，有効性が考古学[7]，心理学[15]，最近ではとくに情報検索の分野等で見出されたということである[4]，[5]，[6]．おそらく，今後も多方面に応用されるであろうと思われるのであるが，考古学でも使われるということはまったく想像を絶するといわざるを得ない．

ここでちょっとお断りをしておかなければならない．1972年の暮に，当ＯＲ学会の高橋磐郎教授のお誘いにより，私はある勉強会に参加した．そこでは，主に[4]の紹介の話題であって，私ははじめてC性の問題を知ったのである．もともと私は数学が専攻分野であり，したがってかかる問題の意義，その後の発展の方向，また応用面については残念ながら今日に至るまでほとんど無知に近いのであるから，

$$\begin{array}{|ccc|}
\hline
0 & 1 & 1 \\
1 & 0 & 1 \\
1 & 1 & 0 \\
\hline
\end{array}$$

図3 C性をもたない例

このような解説を書く資格がはじめからないのである．ただ前記の会後まもなく，この問題に対する一応の数学的な解決を得た[13]，[14]ので，そこに至る道筋をお話しすれば，読者各位がこの問題ばかりでなく広く他の問題を考えられるときにあるいはお役に立つかも知れないと思って書くことをお引き受けした次第である．この問題を精力的に追究している Dr. Lipski（ポーランド科学アカデミー）が昨秋来日し，私が勤めている立教大学において，OR学会との共催で講演された[12]．そのときの草稿を読んでいただくほうがよほど良いように思う．ことにその付録文献表は最近までの研究成果をほとんど集めているから誠に貴重なものであろう．これは公刊されていないのでご希望の方にはコピーを差し上げる用意がある．

2. 問題の転換，区間

さて前の例にもどり，Sを行をあらわす記号a, b, c, d, eから成る集合とする．そうすれば列Aはa, eから成るSの部分集合とみなすことができる．

すなわち，$A=\{a, e\}$である．同様に，

$$B=\{b, d\}, \qquad C=\{a, c, e\},$$
$$D=\{a, c\}, \qquad E=\{b, e\}$$

となる．逆にある集合Sと，その部分集合いくつかの集りまたは族(family) $\Phi=\{A, B, \cdots\}$が与えられているならば，それに対応して図1のような01行列を作ることができる．すなわち，Sを行方向に，またΦを列方向にとり，Sの元aがΦに属するAに含まれるならばa行A列の行列要素を1とし，そうでなければ0とすればよいのである．つまり01行列の問題は直ちに集合と部分集合の族の問題にいいかえられるわけである．

そこでC性を後者，すなわち集合と部分集合の族の場合に移すとすれば，図1の例ではつぎのようになる．集合Sに，たとえば，

$$c>a>e>b>d$$

のように順序（全順序，または線形順序）を導入すると，族$\Phi=\{A, B, C, D, E\}$の各部分集合が，すべてSの区間(interval)になるということに他ならない．たとえばCは$\{x|c\geqq x\geqq e\}$とあらわすことができるし，また他も同様である．一般的にいうならば，C性とは，ある集合Sとその部分集合の族Φが与えられたときに，Sに適当な順序を定めて，Φのすべての元，実はSの部分集合がその順序に関して区間になるかどうかの問題になる．

区間ということを厳密に定義しておこう．Iをある順序集合の部分集合とする．いま$x, y\in I$で，$x\geqq y$とすれば，$x\geqq z\geqq y$となるzが必ずまたIに含まれるときIを区間とよぶ．この定義は集合が無限である場合も考えた一般的なものであるが，いまは有限集合のみを対象としているので，区間には常に最大元または上端，および最小元または下端が存在する．以下では上端がa，下端がbである区間を(a, b)で記すことにする．すなわち，$(a, b)=\{x|a\geqq x\geqq b\}$である．

3. 閉　　包

古風な記法であるが，以下においては，Sの部分集合A, Bの和集合または合併集合を$A+B$，またそれらの共通部分をAB，さらにAのSにおける補集合をA'であらわす．周知のように，これらの演算は交換則，結合則，分配則：$A(B+C)=AB+AC$, $A+BC=(A+B)(A+C)$を満たし，また$A\supseteqq B$ならば$A+C\supseteqq B+C$，$AC\supseteqq BC$である．さらに補集合については，$A\supseteqq B$ならば$A'\subseteqq B'$，$(A+B)'=A'B'$，$(AB)'=A'+B'$となる．特殊な記号として，もし$A\supseteqq B$か$A\subseteqq B$ならば，$A\sim B$と記す．すぐわかるように，$A\sim B$ならば，$A+C\sim B+C$，$AC\sim BC$であり，$A'\sim B'$となる．この程度の代数的演算規則でC性の問題は解けるのである．

さて，この種の問題においては，つぎのように考えることが場合によっては有効である．いまΦがC性をもつならば，ΦにさらにSの部分集合を

いくつか追加して，より大きな族を構成したとき，それがまた当然に C 性をもつのではないであろうか？いま $I=(a,b)$，$J=(c,d)$ が順序集合 S の区間であって，$IJ \neq \phi$（空集合）とすれば，あきらかに $I+J$ もまた区間となり，上端は $\max(a,c)$，また下端は $\min(b,d)$ となることがわかる．たとえば図2において，$B+E$ がそうなっている．そこで，Φ が C 性をもつか否かは別として，Φ にさらに S の部分集合を追加する操作 w をつぎのように定める：

$$w\Phi=\{X+Y|X,\ Y\in\Phi,\ XY\neq\phi\}\cup\Phi.$$

そうすれば，Φ が C 性をもつならば，新しいより大きな族 $w\Phi$ もまた C 性をもち，あきらかにその逆も成り立つ．図2の例で，$w\Phi$ がどうなるか試みてみられたい．たとえば $A+D$，$A+E$，$B+E$ 等々が追加されることがわかるであろう．$w\Phi$ にさらに操作 w をほどこしたものを $w^2\Phi$，一般に n 回繰返しほどこした結果得られた族を $w^n\Phi$ と記すならば，

$$\Phi\subseteqq w\Phi\subseteqq w^2\Phi\subseteqq\cdots\subseteqq w^n\Phi\subseteqq\cdots$$

なる族の無限列ができて，各 $w^n\Phi$ は結局 Φ が C 性をもつとき，またそのときに限って C 性をもつことになる．よって，いま，

$$\Gamma=\cup^{\infty}{}_{n=0}\, w\Phi,$$

すなわち上の無限列のすべての合併集合をあらわすものとすれば，この Γ に操作 w をほどこしても，もう Γ はこれ以上大きくならない．いいかえれば，$\Gamma\ni X,Y$ で $XY\neq\phi$ ならば，常に $X+Y\in\Gamma$ となるのである．このような Γ を $c(\Phi)$ と記し，Φ の閉包(closure)とよぶことにしよう．現在 S は有限集合であるから，あきらかにその部分集合全体の族といえどもやはり有限であって，したがって上のような列が無限に続くわけがなく，ある所から先は一定になってしまう．すなわち，

$$\Gamma=w^n\Phi=w^{n+1}\Phi=\cdots$$

となるであろう．再び図2で試みてごらんになるとよい．上述をまとめればつぎのようにいえる．

S の部分集合の族 Φ が C 性をもつのは，その閉包 $c(\Phi)$ が C 性をもつときに限る．

ところで，Φ がもうそれ以上大きくできない，すなわち $\Phi=c(\Phi)$ または $\Phi=w\Phi$ であるという場合がある．このとき Φ は閉じている(closed)とよぶ．たとえば Φ でその任意の二つの部分集合 A,B が常に $A\sim B$ となる場合である．このとき Φ のすべての元は，包含関係 $\supseteqq$ について線形順序，簡単にいえば一番大きい集合から順々に前のものに含まれるという状態になっている：$A\supseteqq B\supseteqq\cdots\supseteqq C$．このような Φ は入れ子族(nested)とよばれる[4]．あきらかに任意の Φ に対する $c(\Phi)$ は閉じている．

なぜ $c(\Phi)$ を問題にするのであるか？ それは Φ より $c(\Phi)$ のほうが一般には大きいので，C 性をもつための条件は，$c(\Phi)$ に課する条件のほうが Φ に対するそれよりも，より単純になるであろうと推察されるからである．これはごく自然な考え方であろう．

4. I 性，区間の特徴

$A_1=(a_1,b_1)$，$A_2=(a_2,b_2)$ が順序集合の二つの区間であれば，A_1A_2 が空集合でなければ，その上端は $\min(a_1,a_2)$，また下端が $\max(b_1,b_2)$ である区間となる．これは C 性の問題を考えるとき，別に際立った性質ではない．

さらに三個の区間 $A_1=(a_1,b_1)$，$A_2=(a_2,b_2)$，$A_3=(a_3,b_3)$ について，それらの共通部分を考えてみると，$A_1A_2A_3$ はやはり空集合でなければ，上端が $\min(a_1,a_2,a_3)$，下端が $\max(b_1,b_2,b_3)$ となり，とにかく $A_1A_2A_3=(a_i,b_j)=A_iA_j$ の形になる．ここで i,j は 1,2,3 のいずれかであって，いま k を i とも j とも異なるものとすれば，たとえ $A_1A_2A_3=\phi$ であっても，$A_iA_j\subseteqq A_k$ となる．記号をかえていいかえれば，いま A,B,C を順序集合 S の三つの区間とすると，

$$AB\subseteqq C\ \text{か}\ BC\subseteqq A\ \text{か}\ CA\subseteqq B$$

のいずれか一つが成り立たねばならないことがわかる（空集合は任意の集合に含まれることに注

意)．これは図4を参照するならば，区間を特徴づける相当に有力な性質であると考えてよいであろう．いいかえれば，普通の部分集合を三個勝手にもってきても上記の条件は必ずしも成り立たないのである．そこで，族Φの任意の三個の部分集合A, B, Cについて，このような条件が満たされるとき，ΦはI性(intersection property) をもつとよぶことにする．したがって，ある族ΦがC性をもつためには，Φはもちろん，その閉包$c(\Phi)$がI性をもつことが必要である．実は後述するようにこれがまた十分条件になるのである．つまり，集合Sの部分集合の族ΦがC性をもつということは，$c(\Phi)$がI性をもつということと同じである．もちろん，四個あるいはそれ以上の個数の区間が満たすべき条件が考えられるけれども，いまの場合は三個の部分集合の関係であるI性が基本的本質的であることがわかるのである，組合せ論的問題では，これに類した場合が多いように思われる．

図3の例は一目でわかるようにΦそれ自身がすでにI性をもたない．ところが図5ではΦはI性をもつけれども，$c(\Phi)$についてはだめである．したがって二つともC性をもたない．図5の場合，$AD=BC=\phi$であるから，Φからどの三つをとってもその中にはAとDか，またはBとCが含まれるのでI性が成り立つ．しかし，いま$E=A+B$, $F=A+C$, $G=B+D$とおけば，目で追っていただくとすぐわかるように，E, F, Gの間ではI性は満たされていない．

5. I性のいいかえ

上述のようにC性の問題と閉包のI性とが深いかかわりがあることがわかった．数学の問題としてはこれで終わりとしてよいのであろうが，応用上はまだまだ追究しなければならない点がたくさんある．その一つは，C性があるのか否かを効率のよい計算手順またはいわゆるアルゴリズムで確かめる問題である[1], [3]．そのために，以下に二種のI性と同値な性質を述べることにする．これらがすぐに役立つかどうかは速断できないのであるが，いずれにせよ，このような考察を抜きにしていきなりアルゴリズムの工夫に走ることは賛成し難い．いかなるアルゴリズムでも，その背後には確かな理論的裏づけが必要であろう．

その一つは，いま問題としているΦのかわりにΦに属する各部分集合の補集合のほうからI性を論ずるやり方である．これは01行列で当面の1のかわりに0の配置について考えることになるから，ちょっと奇妙な感じがしないでもないが，第3節で演算規則について述べたように二つの部分集合の和集合をつくることと，共通部分をとることは，それぞれの補集合を考えると入れかわるので，数学的には別に不自然ではないのである．すなわち，三つの部分集合A, B, Cの間で，もし$AB\subseteq C$ならば，それは$A'+B'\supseteq C'$と同じことである．よってI性はまた，$A'+B'\supseteq C'$か，$B'+C'\supseteq A'$か，$C'+A'\supseteq B'$の少なくとも一つが成り立つことといってよい．それでΦのかわりに$\Phi'=\{X'|X\in\Phi\}$を考えるならば，$c(\Phi)'$はつぎのような操作を繰り返して得られる．すなわち，$X'+Y'\neq S$ならば$X'Y'$をつけ加える．なぜなら，$\phi'=S$また$S'=\phi$であるからである．このようにして$c(\Phi)'=\{Z|Z'\in c(\Phi)\}$に到達するのであ

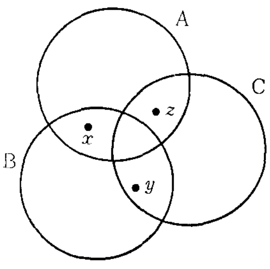

図 4　3個の部分集合の一般的関係

	A	B	C	D
a	1	0	1	0
b	0	1	0	1
c	1	1	0	0
d	0	0	1	1

E	F	G
1	1	0
1	0	1
1	1	1
0	1	1

図 5　ΦがI性をもち，$c(\Phi)$がもたない例．

るが，ちょっと見たところC性とかけ離れたような印象をもたれるであろう．

つぎに考えられるのは，01行列の行と列を転換した方法である．第2節で述べたのと逆に，小文字であらわした各行は，また大文字であらわしたΦの元の集合と考えることができる．たとえば，図1において，行aはA, C, Dの集合とみなすのである．同様に$b=\{B, E\}$，$c=\{C, D\}$等となる．もし族ΦでI性が成り立たないものとすれば，Sの三つの元x, y, zが存在して，$x\in AB$であるが$x\notin C$，また$y\in BC$で$y\notin A$，$z\in CA$で$z\notin B$となるに違いない．これを，x, y, zがΦの部分集合とみるならば，$xy\ni B$で$z\not\ni B$，$yz\ni C$で$x\not\ni C$，さらに$zx\ni A$で$y\not\ni A$といいかえることができる．すなわち，三つのΦの部分集合x, y, zでは，$xy\not\subset z$，$yz\not\subset x$，$zx\not\subset y$となって，行の側からみたI性が成り立たないことになる．01行列では，いままで単にI性と称したのは列のI性とよぶべきものであった．ここに述べたことは，行のI性と，列のI性はまったく対称的に同じであることを示している．したがって．族ΦがC性をもつためには，$c(\Phi)$が行についてI性を満たすことが必要かつ十分であることになる．この結果もまたC性についての素朴な観察を裏切るように思われる．

6. 十分条件

上述のように，C性が成り立つためには，閉包がI性をもつことが必要である．逆にこれが十分であることの証明は比較的に厄介である．詳細は原文献[13]を参照していただくこととして，以下ではその大筋を説明する．

周知のように，実数を定義するには有理数のDedekind 切断(X, X')というやり方がある．ここでXはある実数より小さい有理数の集合であり，またX'はその実数より大きい有理数の集りとなる．そうして二つの実数をあらわす切断(X, X')と(Y, Y')の間の順序は，$X\supseteq Y$したがって$X'\subseteq Y'$である時に$(X, X')\geqq(Y, Y')$とするのである．類似の方法でSに順序を導入することが可能である．

いま$c(\Phi)$がI性を有し，Sと異なる四個の部分集合A, B, C, Dについて，もし$A+B=C+D=S$ならば，$A\sim C$かつ$B\sim D$であるか，または$A\sim D$かつ$B\sim C$となることが証明されるのである．さらに，もし$A, B\in c(\Phi)$が$A+B=S$となるならば，ΦにA'を追加した族の閉包$c(\Phi\cup\{A'\})$がまたI性をもつのである．したがって，$S=A+B$となるようなAやBに対して，ΦにつぎつぎA'およびB'をつけ加えるということを繰返すならば，ついにはさらに大きな族Ψに到達して，そこではもし$A, B\in\Psi$が$A+B=S$となるならば常に$A', B'\in\Psi$となり，かつΨはI性をもつことになるのである．もちろん，Ψは閉じている．さて，いま1組の$P, P'\in\Psi$をとり出して，上述の事実にもとづき，$P\sim X$かつ$P'\sim X'$となる$X\in\Psi$から成る組(X, X')のすべてを考える．そうすると，これら二個の組(A, A')，(B, B')の間では常に$A\sim B$かつ$A'\sim B'$となる．よってもし$A\supseteq B$ならば$A'\subseteq B'$であり，こういう場合は，二つの組または切断に対して$(A, A')\geqq(B, B')$と定義することにする．このことによって，はじめに述べたように，Sに順序を導入することができて，Φに含まれる各部分集合が，この順序に関して区間となるようにすることが可能になるのである．原証明を読んでいただければ，このような順序の入れ方がすべて尽されていることがよくおわかりになると思う．この結果は，Sが無限集合の場合でも正しいことが証明されている[11]．すなわち，区間の集りの特徴はI性にあることが完全に示されたのである．

この簡単な応用の一例として，ΦがC性をもつならば$Q\subset S$がつぎの条件を満たすとき，QをΦに追加した族がまたC性をもつ：$A, B\in c(\Phi)$でQがこの二つおよびA', B'と共通部分をもつならば，常に$A+Q\sim B+Q$，また$A'+Q\sim B'+Q$

となる．とくに Φ が3節で述べた入れ子族であるときは任意の部分集合 Q をつけ加えてもまた C 性をもつことがわかる[4]．この例に限らず，C 性の問題はすべて上述の結果を基礎にしないで論ずることはできないように思う．

7. グラフ理論との関係，その他

部分集合の族 Φ の各元を頂点 (vertex) とし，もし $A, B \in \Phi$ が $AB \neq \phi$ ならば A と B を辺(edge)で結んで得られるグラフを共通部分グラフ(intersection graph)という．とくに Φ が順序集合の区間から成るとき，そのようなグラフを区間グラフ(interval graph)とよび，その特徴づけが以前から考察されたのである[3], [8]．私はグラフ理論をあまりよく知らないし興味をもたないのであるが，前記 Lipski のいうところによれば，これらの結果はいずれもわれわれの得た結果と同値，すなわち一方から他方が導かれる由である．しかし，そのための証明はわれわれのものがより単純であるということである．同人もまた[9]において一個の定理を提示している．Tucker[16]は[8]の結果から01行列が C 性をもつための必要十分条件を出している．それによると，ある01行列が列について C 性を満たすためには5種の禁止部分行列(forbidden submatrix)をふくまないということが要件になっている．ここで部分行列というのは，与えられた01行列の行および列をそれぞれにいくつか指定して，その交差する要素だけをとり出してできる行列のことである．たとえば図3のようなものである．ただしその中3種は形態は類似でも大きさはさまざまである．この結果にもとづいて，C 性判定のよいアルゴリズムが得られるか否かは，にわかに判断できない．

情報検索に応用することを目的として，C 性の一種の拡張が提起されている[10]．すなわち，S から S 自身の中への写像 φ で，Φ に属するすべての部分集合 A が，

$$A=\{a, \varphi(a), \varphi^2(a), \cdots, \varphi^n(a)\}$$

の形であらわすことができるかという問題である．ここで $\varphi^2(a)=\varphi(\varphi(a))$，一般に $\varphi^m(a)=\varphi(\varphi^{m-1}(a))$ を意味する．確かにこのような目的に対しては有用であろうが，いまのところまだ簡単な特徴づけは得られていない．

また，S のいくつかの元を重複して使うことにより Φ に C 性をもたらすことが研究されている[2]．これも実用上は有効であろうと考えられるけれども，やはり現在までの結果はまだ満足すべき段階に達していないようである．

かえりみれば，C 性の問題それ自体は，数学の問題としては実に簡明であるけれども，これを実際に応用するためには，今後さらに追究されなければならない種々の問題をかかえている．

参考文献

[1] Booth, K. S. and Lueker, G. S.: Testing for the consecutive ones property, interval graphs, and graph planarity using PQ-tree algorithms. *J. Comp. System Sci.*, **13** (1976), 335-379.

[2] Ehrich, H. D. and Lipski, W.: On the storage space requirement of consecutive retrieval with redundancy. *Inform. Processing Letters*, **4** (1976), 101-104.

[3] Fulkerson, D. R. and Gross, O. A.: Incidence matrices and interval graphs. *Pacific J. Math.*, 15 (1965), 835-855.

[4] Ghosh, S. P.: File organization; the consecutive retrieval property. *Comm. ACM*, **9** (1972), 802-808.

[5] ———: On the theory of consecutive storage of relevant records. *J. of Information Sciences*, **16** (1973), 1-9.

[6] ———: Consecutive storage of relevant records with redundancy. *Comm. ACM*, **18** (1975), 464-471.

[7] Kendall, D. G.: Incidence matrices, interval graphs and seriation in archeology. *Pacific J. Math.*, **28** (1969), 595-570.

[8] Lekkerkerker, C. G. and Boland, J. C.:

Representation of a finite graph by a set of intervals on the real line. *Fundamenta Math.*, **51** (1962), 45-64.

[9] Lipski, W.: Information storage and retrieval-mathematical foundations II (Combinatorial problems). *Theoretical Computer Science*, **3** (1976), 183-211.

[10] Lipski, W. and Marek, W.: An application of graph theory to information retrieval. *Bull. Acad. Polon. Sci., Sér. Sci. Math. Astronom. Phys.*, **22** (1974), 691-695.

[11] Lipski, W. and Nakano, T.: A note on the consecutive 1's property (infinite case). *Comm. Math. Univ. St. Pauli*, **25** (1977), 149-152.

[12] Lipski, W.: Consecutive 1's property and related topics; recent results. *Lecture at St. Paul's University*, 1977, unpublished.

[13] Nakano, T: A characterization of intervals; the consecutive (one's or retrieval) property. *Comm. Math. Univ. St. Pauli*, **22** (1973), 49-59.

[14] ———: A remark on the consecutivity of incidence matrices. *Comm. Math. Univ. St. Pauli*, **22** (1973), 61-62.

[15] Roberts, F.S.: *Discrete Mathematical Models, with Applications to Social, Biological and Environmental Problems*. Prentice-Hall, Englewood Cliffs, N. J., 1976.

[16] Tucker, A. C.: A structure theorem for the consecutive 1's property. *J. Comb. Theory*, **12**B (1972), 153-162.

なかの・たけお　1926年生
1948年　東京大学理学部数学科卒
現在　立教大学理学部教授

