化学反応の単純化

1　前書き

高校化学などで紹介される化学反応のとらえ方として、代表的な分子の結合状態をいくつか挙げ、その状態の間の遷移確率なるものを与える（ことと本質的に同値な）モデルが紹介される。

$$2\,H_2O \quad \leftrightarrow \quad H_3O^+ + OH^-$$

のような化学反応である。

だが実際の現象はここまで単純ではなく、もっと多く（無限の？）状態の間を行ったり来たりしていると考えた方がしっくり来る。

オキソニウムイオン H_3O^+ と水酸化物イオン OH^- から 2 つの水分子 H_2O に変化する途中で、H_2O と OH^- の中間の位置に H^+ が浮いてるような状態の瞬間を経由することもあれば、H_3O^+ と OH^- がくっついた状態になってから 2 つの H_2O に分離する場合もあるような気がする [1]。水分子 H_2O 内の酸素イオン O^{2-} と陽子 H^+ との距離だって（古典的描像で見ても）一定ではなく、瞬間瞬間で 0.1Å 前後の値をフラフラしているであろう。

だがその一方で、代表的な（この例で言えば H_2O, H_3O^+, OH^- という 3 種類の分子・イオンとして存在する）状態のみを考えても実際の現象をうまく近似しているというのも妙に納得できることであるし、その意味でこの化学反応のモデルには意味があると多くの人は考えるであろう。

では現実（？）には無数に存在する状態の中から“代表的な状態”を抽出するだけでなぜ現象がうまく説明できるのであろうか。また“代表的な状態”なるものはどのようにして選ばれるのであろうか。

前者についてはこの記事の全体を通じて考えていくとして、ここでは後者について大ざっぱに考えてみよう。

代表的な状態、それは大きく分けて次の 2 つの考え方により選ばれていると私は思う。

1 つ目は“準安定”とでも言うべき状態のみを考え、その他の状態は存在時間が短いという理由で無視するという考え方である。上の例で言えば H^+ が他の分子と離れて孤立している状態や H_4O_2 のように共有結合を結べないくらいに原子が寄り集まっている状態がそれにあたる。

もう一つは“似た状態”を 1 つの同じ状態であるとみなしてしまうという方法である。上の例で言えば O^{2-} と H^+ の間の距離が 0.05Å だろうと 0.2Å だろうと同じ水分子 H_2O の状態として同一視してしまっているわけである。

数学的に言えば、より現実に近いモデルとして多くの状態の集合を初めに考え、そのモデルの単純化として、前者は状態集合としてもとのモデルの部分集合を、後者は商集合を考えるわけである。

先ほどの分子間の距離のように、連続的に変化しうるパラメータの値ごとに異なる状態を考えるのが本来なら自然なのかもしれない。だが今回は有限個の状態のみからなる系と、その状態空間の上のマルコフ連鎖を考察する。

このような設定をする意図は 2 つある。

一つは単純に有限個の状態しかないモデルの方が数学的に平易 [2] だろうということである。

もう一つは、状態・時間共に離散的なモデルを解析すれば、多少の障害はあろうが既存の一般論によって時間・空間共に連続な確率モデルの結果を導けると思われ、その意味で一般性を失っていないと思うからで

[1] ただの妄想である。実際にどんな現象が起こっているのか筆者は知らない
[2] 理解に必要な予備知識が非常に少なくて済む、というような意図である。

ある。(きちんと考えたわけではないので見逃しも多々あろうが、一般化の障害となりそうなことは気づいた範囲でコメントしておこうと思う。)

またそもそもの今回の主題である、少ない状態のみを想定するモデルと複雑だが現実をより正確に反映したモデルとの関連は、(時間は連続の方が良いかもしれないが)“有限種類の状態のみを想定するモデル同士”で比較する方がかえって見やすい[3]のではないかという思いもあるが、これはあまり適切な考えではないかもしれない。

今回のテーマは数学ではない。だが書いてる人間が人間だけに数学的視点寄りの考察であり(と言っても数学としてはそれはそれで不完全である)、科学的にも哲学的にもいろいろとおかしな点があると思う。寛大な心で読んでいただくとともに適切なご指摘をいただければ幸いである。

2　予備知識　マルコフ連鎖に関する基本的な定義と性質

本稿では以下の表記や用語を用いる。そこまで特殊な表記ではないと思うが、一般に普及している表記法がどのようなものかをきちんと調べていないため、他の文献とは使い方が微妙に異なって読みにくいかもしれない。

この節の命題群の証明については確率論の教科書、例えば [1] の §7.4 を参照されたい。

記号 2.1. 有限集合 S に対し、$\mathrm{M}(S,\mathbb{R})$ で S^2 から $\mathbb{R}$ への写像の全体を表す。また $S=\{1,2,\ldots,n\}$ のときは $\mathrm{M}(S,\mathbb{R})$ を $\mathrm{M}(n,\mathbb{R})$ とも表す。$\mathrm{M}(S,\mathbb{C})$ などの定義も同様とする。

また $P=(p_{i,j})_{(i,j)\in S^2}\in\mathrm{M}(S,\mathbb{R})$ に対し P^n の $(i.j)$ 成分を $p_{i,j}^{(n)}$ と表す。

$S\times S$ を添え字として実数を並べたものであるから、要するに普通の正方行列の全体である。

記号 2.2. $a=(a_i)_{i\in S}\in\mathbb{C}^S$ に対し、$\sum_{i\in S}a_i$ を $s(a)$ と表す。

定義 2.3. 有限集合 S 上の状態分布とは以下をみたす $a=(a_i)_{i\in S}\in\mathbb{R}^S$ のことをいう。

- $s(a)=1.$
- 任意の $i\in S$ に対して $a_i\geq 0.$

状態分布は確率ベクトルということもある。

定義 2.4. $P=(p_{i,j})_{(i,j)\in S^2}\in\mathrm{M}(S,\mathbb{R})$ は、全ての $j\in S$ に対して $(p_{i,j})_{i\in S}$ が S 上の状態分布であるとき、確率行列という。

S で添え字づけられた確率行列の全体を $\mathcal{P}(S)$ と書くことにする。$S=\{1,\ldots,n\}$ のときの $\mathcal{P}(S)$ は $\mathcal{P}(n)$ とも書く。

マルコフ連鎖とは……と、集合論的な記法できちんと書いても良いのだが、あまり意味はないと思うのでやめておく。

マルコフ連鎖の大まかなイメージとしては S が“あり得る状態の全体”であり、離散的な時刻を考えて状態 $j\in S$ にあるものは次の時刻には確率 $p_{i,j}$ の確率で状態 i に変化する、という変化が次々と起こっていくというものである。

S 上の状態分布 $(a_i)_{i\in S}$ というのは状態 i である確率が a_i であるというようなイメージであり、$a=(a_i)_{i\in S}$ という状態分布は次の時刻には $Pa=(\sum_{k\in S}p_{i,k}a_k)$ に変化するのである。

[3]筆者のように確率論に明るくない場合の話かもしれないが。

単位時間あたりの変化を表すのが P という確率行列であり、P にしたがうマルコフ過程を考える際には P は遷移（確率）行列ということもある。

さて、確率行列についての基本的な性質をお復習いしておこう。

命題 2.5. P を確率行列とする。

(1) P は 1 を固有値に持つ。

(2) P の固有値はどれも絶対値が 1 以下である。

定義 2.6. $P = (p_{i,j}) \in \mathcal{P}(S)$ とする。

(1) P は条件「ある正の整数 N が存在し P^N の全ての成分は正」をみたすときエルゴード的であるという。

(2) P は条件「各 $i, j \in S$ 毎に、正の整数 N で $p_{i,j}^{(N)} > 0$ となるものが存在する」をみたすとき、既約であるという。

(3) $i \in S$ が P（にしたがう確率過程）において非周期的であるとは「『$m > N$ ならば $p_{i,i}^{(m)} > 0$』をみたす正の整数 N が存在する」ことをいう。

命題 2.7. 確率行列に対して以下の 3 条件は互いに同値である。

(1) エルゴード的である。

(2) 既約であり、かつ非周期的な状態を持つ。

(3) 既約であり、かつ任意の状態は非周期的である。

命題 2.8. P を既約な確率行列とする。

(1) P の固有値 1 の重複度は 1 である。

(2) 固有値 1 の固有ベクトルは、正の成分と負の成分が同時には存在しない。

(3) 1 以外の固有値に対する固有ベクトル v は $s(v) = 0$ をみたす。

(4) P がエルゴード的なら 1 以外の P の固有値の絶対値は 1 より真に小さい。

定義 2.9. 確率行列 P の固有値 1 に属する固有ベクトル v で $s(v) = 1$ となるものを P（にしたがう確率過程）の定常状態という。

3 同一視による単純化

3.1 問題の数学的な表現

本節では似ている状態を同一視することによるモデルの単純化を考える。

まず大ざっぱなとらえ方として、次のような状況を想定するのが自然ではないかと思う。

- S：有限集合、$P = (p_{i,j})_{(i,j)\in S^2}$：$S$ 上の確率遷移行列
- $\sim$：S 上の同値関係
- $i \not\sim j$ であれば $p_{i,j} \fallingdotseq 0$.

S はあり得る状態を全て集めた集合を表し、$i \sim j$ であることは i と j とが“似た状態”であるということにあたる二項関係である。最後の条件は「似た状態の間の行き来は、似ていない状態と比べてはるかに起こりやすい」ということにあたる条件である [4]。

さて、$\fallingdotseq 0$ という表現をどのようにとらえたら良いだろうか。これにはいろんな定式化が考えられるだろうが、本稿では極限でそれらの成分が 0 に収束するような、1 つの変数で媒介変数表示された遷移行列の族を考えるという方法をとることにする。

つまり $\sim$ を S 上の同値関係とし、$P(0) = (p_{i,j})$ が $i \not\sim j$ なる成分については $p_{i,j} = 0$ となる、連続写像 $P : [0, \varepsilon) \to \mathcal{P}(S)$ を考えるわけである。

というわけで、本稿の全体を通じて行列の 1 パラメータ族を取り扱う。有限集合上のマルコフ連鎖の解析においては、確率行列の固有空間への分解が基本的な役割を果たす。なので本題に入る前に、行列値連続関数に対してその固有値と固有ベクトルが連続的に変化することを確認しておこう。その 1 つのバージョンとして次のようなことが成り立つ。

補題 3.1. n を正の整数、$P : [0,1] \to \mathrm{M}(n, \mathbb{R})$ を連続写像とし、$\alpha_1, \dots, \alpha_n$ を $P(0)$ の固有値の全体 [5] とする。

このとき次のような連続関数 $\alpha_1(t), \dots, \alpha_n(t), v_1(t), \dots, v_n(t)$（$\alpha_i : [0,1] \to \mathbb{C}$, $v_i : [0,1] \to \mathbb{C}^n$）が存在する：

(i) $\alpha_i(0) = \alpha_i$.

(ii) 任意の $t \in [0,1]$ に対し、$P(t)$ の特性多項式は $(x - \alpha_1(t)) \cdots (x - \alpha_n(t))$ と因数分解される。

(iii) 任意の $t \in [0,1]$ に対して $v_i(t)$ は $P(t)$ の固有値 $\alpha_i(t)$ に対する広義固有空間に属する 0 でないベクトルである。

上の条件に加え、さらに

(iv) $\alpha_i \neq \alpha_j \;\Rightarrow\; (\forall t \in (0,1]\,;\ \alpha_i(t) \neq \alpha_j(t))$

をみたすならば「各 $t \in (0,1]$ に対し、$\{v_1(t), \dots, v_n(t)\}$ は $\mathbb{C}^n$ の基底」となるようにとることができる。

補題 3.2. n を正の整数、$Q \in \mathrm{M}(n, \mathbb{R})$ とするとき、次の条件をみたす Q の近傍 $V \subset \mathrm{M}(n, \mathbb{R})$ が存在する：任意の $P \in V$ に対し、$P(1) = P$, $P(0) = Q$ および前補題の (i), (ii), (iii), (iv) をみたす写像 $P : [0,1] \to V$ が存在する。

要するに $P = P(1)$ と $Q = P(0)$ とが係数が近い行列なら、P の広義固有空間は Q の広義固有空間の近くにあり、遷移行列 P にしたがうマルコフ連鎖の漸近的な振る舞いが Q にしたがうマルコフ連鎖で近似できることが言いたいのである。

それだけのためならこんなに条件が入り組んだ主張ではなく、もっと簡単なものがある気もするのだが、思いつかなかったのでこういう補題を取ることにさせてもらった。何かもっと適当な主張を思いついたらこの部分の文章は更新したいと思う。

[4] もちろん“似ている”というのは普通は推移律が成り立たない関係であるから、突き詰めると同値関係であるという仮定には不自然に感じられるかもしれない。だが複雑な系を少数の状態のみで近似したいという目的からすると「同値関係でまとめる」という発想は自然だろう。他に適当な単語を思いつかなかったので「似ている」という言葉を用いたが、不適切さはこの単語選びにあると言うのが正しい思う。

[5] 重複固有値がある場合、重複度と同じ回数その固有値が現れるとする。以下でも固有値の全体と言えば同じ意味とする。

3.2　数学を使わない直観的な予想

移りやすい状態が短い時間のうちに互いに同一視できるようになるならば、同一視した後の状態空間の上での確率遷移行列はどのようになるであろうか。まずはできるだけ単純な例でその考え方を見てみよう。

なお、この小節で述べることは化学の業界でも前から予想されていたようである [6]。

例 3.3. 確率行列

$$\begin{pmatrix} \frac{1}{3}-t & \frac{1}{2}-t & t & 0 \\ \frac{2}{3} & \frac{1}{2} & 2t & 0 \\ t & 0 & \frac{2}{5}-3t & \frac{1}{2} \\ 0 & t & \frac{3}{5} & \frac{1}{2} \end{pmatrix}$$

にしたがうマルコフ過程を考える。状態の名前は行や列の番号をそのまま使うことにする。t が 0 であれば、これは状態 $1,2$（以下第 1 ブロックと呼ぶ）と $3,4$（以下第 2 ブロック）は完全に断絶された系となるから、t が小さい正の数のときには、第 1 ブロックと第 2 ブロックの中では（$1/t$ の大きさと比べて）一瞬で定常状態に達するように思われる。

t が完全に 0 のときは、1 と 2 の間の確率遷移行列

$$\begin{pmatrix} \frac{1}{3} & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{2} \end{pmatrix}$$

は固有値が 1 と $-\frac{1}{6}$ で、定常状態が ${}^t\left(\frac{3}{7}, \frac{4}{7}\right)$ であり、第 2 ブロックの確率行列

$$\begin{pmatrix} \frac{3}{5} & \frac{1}{2} \\ \frac{2}{5} & \frac{1}{2} \end{pmatrix}$$

は固有値が 1 と $-\frac{1}{10}$ で、${}^t\left(\frac{5}{9}, \frac{4}{9}\right)$ が定常状態である。

t が 0 でなはいが小さい正数である場合には、第 1 ブロックにあるものはほとんど状態 3 や 4 には変化せず、ごく短時間経過した後には“第 1 ブロックにおける定常状態”${}^t\left(\frac{3}{7}, \frac{4}{7}, 0, 0\right)$ に近い分布になるだろう。同様に第 2 ブロックにあれば短時間のうちに ${}^t\left(0, 0, \frac{5}{9}, \frac{4}{9}\right)$ に近い分布になると思われる。

t に対して十分小さな時間しか経過しないうちに各ブロック内では“定常状態”になり、しかもブロック外への流出やブロック外からの流入が多少あっても、ブロック内の移動は速いのですぐに定常状態に戻る。故に開始からほんの少しの時間の経過以後はずっと、各ブロックを“一つの塊”とみなし、それ以降の状態変化は各ブロック間の流入出のみを観察することで記述できるのではないかと考えるのが自然であろう。

t の小ささに対してそれなりに時間が経った後の状態変化を考えてみよう。そのときの確率分布 ${}^t(a_1, a_2, a_3, a_4)$ において各ブロックの中では定常状態で近似できる比率で存在するという仮説により、近似的に $a_1 : a_2 = 3 : 4$, $a_3 : a_4 = 5 : 4$ となっている。第 1 ブロックに確率 $a_1' = a_1 + a_2$, 第 2 ブロックに確率 $a_2' = a_3 + a_4$ で存在するとき、その 1 ステップ後に第 1 ブロックに存在する確率 a_1'' は

$$\begin{aligned} a_1'' =& \left\{ \left(\frac{1}{3} - t\right) a_1 + \left(\frac{1}{2} - t\right) a_2 + ta_3 \right\} + \left\{ \frac{2}{3}a_1 + \frac{1}{2}a_2 + 2ta_3 \right\} \\ =&(a_1 + a_2) - ta_1 - ta_2 + 3ta_3 \\ \fallingdotseq& a_1' - t\frac{3}{7}a_1' - t\frac{4}{7}a_1' + 3t\frac{5}{9}a_2' \\ =&(1-t)a_1' + t\frac{5}{3}a_2' \end{aligned}$$

[6] 歴史については全くと言って良いほど知らないが、本質的に同じことが少なくとも（数学的に正確な定式化や証明を抜きにした予想として）1969 年時点での文献 [?] に見られる。

で近似できるであろう。同様に第 2 ブロックの割合 a_2'' も求めると、結局ブロック間の移動は

$$\begin{pmatrix} a_1'' \\ a_2'' \end{pmatrix} \fallingdotseq \begin{pmatrix} 1-t & \frac{5}{3}t \\ t & 1-\frac{5}{3}t \end{pmatrix} \begin{pmatrix} a_1' \\ a_2' \end{pmatrix}$$

となり、確率行列

$$\tilde{P} = \begin{pmatrix} 1-t & \frac{5}{3}t \\ t & 1-\frac{5}{3}t \end{pmatrix} \tag{1}$$

に基づく変化をしていると考えられるわけである。$\tilde{P}$ の固有値が 1 と $1-\frac{8}{3}t$ であり、定常状態は $\left(\frac{5}{8}, \frac{3}{8}\right)$ である。

また $\left(1-\frac{8}{3}t\right)^n$ が 1 よりははっきりと小さく、0 よりははっきりと大きいような n についての条件は、例えば

$$\left(1-\frac{8}{3}t\right)^n = \frac{1}{2} \iff n = \frac{\log\left(\frac{1}{2}\right)}{\log(1-\frac{8}{3}t)} \fallingdotseq -\log\left(\frac{1}{2}\right)\frac{3}{8}\frac{1}{t}$$

であるから、t^{-1} と同程度の時間スケールでブロック間の反応は進行していくことが予想される。

各ブロック内の状態に存在する確率の比は前述の通り $a_1 : a_2 = 3 : 4,\ a_3 : a_4 = 5 : 4$ であろうから、系全体としての定常状態は近似的に

$$\begin{aligned} &\left(\frac{a_1''}{a_1''+a_2''}\frac{a_1}{a_1+a_2}, \frac{a_1''}{a_1''+a_2''}\frac{a_2}{a_1+a_2}, \frac{a_2''}{a_1''+a_2''}\frac{a_3}{a_3+a_4}, \frac{a_2''}{a_1''+a_2''}\frac{a_4}{a_3+a_4}\right) \\ =&\left(\frac{5}{8}\frac{3}{7}, \frac{5}{8}\frac{4}{7}, \frac{3}{8}\frac{5}{9}, \frac{3}{8}\frac{4}{9}\right) = \left(\frac{15}{56}, \frac{5}{14}, \frac{5}{24}, \frac{1}{6}\right) \end{aligned}$$

となると思われる。

3.3 線形代数的な考察

前小節の例により本節で証明したい主張の大まかなイメージが理解いただけたと思う。

証明は次の小節で述べようと思うが、厳密に証明しようとすると極限に関する多少めんどくさい議論も必要になってくる。

厳密性を求める読者も求めない読者もどちらもいるだろうことを考え、証明の一部にもなってかつ、極限操作については直観的な考察でできる部分を独立させてこの小節で説明しておきたいと思う。

例 3.4. 引き続き、実質的に前小節の例を考える。

$$P(t) = \begin{pmatrix} \frac{1}{3} & \frac{1}{2} & 0 & 0 \\ \frac{2}{3} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{2}{5} & \frac{1}{2} \\ 0 & 0 & \frac{3}{5} & \frac{1}{2} \end{pmatrix} + t \begin{pmatrix} -1 & -1 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 1 & 0 & -3 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} + o(t)$$

の形の行列を考えよう [7]。

[7] $o(t)$ はいわゆるランダウの記号である。基本的には $g(t) = h(t) + o(f(t))$ の形で

$$\lim_{t\to 0}\frac{g(t)-h(t)}{f(t)} = 0$$

のような意図を表す。同様に $g(t) = h(t) + O(f(t))$ は $\limsup_{t\to 0}\frac{|g(t)-h(t)|}{f(t)} < \infty$ である。ここで $f(t)$ の値は非負実数だが、$g(t)$ や $h(t)$ の値域は文脈によって変わるし、関数の定義域も適当に“ $t=0$ の近く ”程度で一定ではないため正確な定義を書くとややこしくなりそうである。だが文脈から容易に判断できると思うので省略する。

また慣習的な表記をきちんと調べていないがこのノートでは、$g(t) = h(t) + o(t)$ を $g(t) \in h(t) + o(f(t))$ と書いたり、$g(t) \equiv h(t) (\mathrm{mod}\ o(f(t)))$ とか $g(t) - h(t) \in o(f(t))$ とか $g(t) + o(f(t)) = h(t) + o(f(t))$ とかの書き方のうちから、その文脈で見やすいと思ったものを使わせていただく。

いくつかの行列に以下のように名前を付けておく。$t=0$ のときの各ブロック内の遷移行列を

$$P_1 = \begin{pmatrix} \frac{1}{3} & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{2} \end{pmatrix}, P_2 = \begin{pmatrix} \frac{2}{5} & \frac{1}{2} \\ \frac{3}{5} & \frac{1}{2} \end{pmatrix}$$

とし、それらを対角化するときの基底の取り替え行列として

$$Q_1 = \begin{pmatrix} \frac{3}{7} & -1 \\ \frac{4}{7} & 1 \end{pmatrix}, Q_2 = \begin{pmatrix} \frac{5}{9} & -1 \\ \frac{4}{9} & 1 \end{pmatrix}$$

をとり、

$$Q = \begin{pmatrix} Q_1 & 0 \\ 0 & Q_2 \end{pmatrix}$$

とする。（ここの 0 は 2×2 のサイズの零行列である。）

また $P(t)$ の t の“ 係数 ”を A とおく。すなわち

$$A = \begin{pmatrix} -1 & -1 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 1 & 0 & -3 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

とする。

このとき

$$Q_1^{-1} = \begin{pmatrix} 1 & 1 \\ -\frac{4}{7} & \frac{3}{7} \end{pmatrix}, Q_2^{-1} = \begin{pmatrix} 1 & 1 \\ -\frac{4}{9} & \frac{5}{9} \end{pmatrix}$$

であり、

$$Q^{-1}P(t)Q = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -\frac{1}{6} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -\frac{1}{10} \end{pmatrix} + t \begin{pmatrix} -1 & 0 & \frac{5}{3} & -3 \\ \frac{4}{7} & 0 & \frac{10}{63} & -\frac{2}{7} \\ 1 & 0 & -\frac{5}{3} & 3 \\ \frac{8}{63} & 1 & \frac{12}{27} & -\frac{4}{3} \end{pmatrix} + o(t)$$

となる。

ここで $t=0$ における固有値 1 の固有ベクトルに対する成分である、第 1 成分と第 3 成分のみを取り出した行列

$$\tilde{P}(t) = E + t \begin{pmatrix} -1 & \frac{5}{3} \\ 1 & -\frac{5}{3} \end{pmatrix} + o(t)$$

が前小節で導出した (1) と同じものであることが（基底の取り替えの意味を考えれば）わかるだろう。

記号が多くなって恐縮だが、後の都合上 $\tilde{P}(t)$ の t の係数を

$$\tilde{A} = \begin{pmatrix} -1 & \frac{5}{3} \\ 1 & -\frac{5}{3} \end{pmatrix}$$

とおいておこう。

$P(t)$ の固有値を考えよう。固有値の連続性により、t について連続な関数 $1, \alpha(t), \beta_1(t), \beta_2(t)$ で、その値が $P(t)$ の固有値になるものが存在する。ここで、$\alpha(0)=1, \beta_1(0) = -\frac{1}{6}, \beta_2(0) = -\frac{1}{10}$ である。

$\alpha(t)$ の値について考える。固有多項式 $\mathrm{Det}(Q^{-1}P(t)Q - E\lambda)$ の行列式の成分のうち、$t=0, \lambda=1$ を代入して 0 にならない成分は 1 以外の固有値が並んだ対角成分のみである。よって

$$\mathrm{Det}(Q^{-1}P(t)Q - E\lambda) = \left(-\frac{1}{6}\right)\left(-\frac{1}{10}\right) \begin{vmatrix} 1-\lambda-t & \frac{5}{3}t \\ t & 1-\lambda-\frac{5}{3}t \end{vmatrix} + o((t, 1-\lambda)^2)$$

と表せる。

ここで $o((t,1-\lambda)^2)$ は

$$\frac{g_1(t,\lambda)}{t^2}, \frac{g_2(t,\lambda)}{t(1-\lambda)}, \frac{g_3(t,\lambda)}{(1-\lambda)^2} \to 0 \quad ((t,\lambda)\to(0,1)\text{ のとき})$$

となる関数 g_1, g_2, g_3 により $g_1(t,\lambda)+g_2(t,\lambda)+g_3(t,\lambda)$ の形で表せる関数を表す [8]。

$o((t,1-\lambda)^2)$ という“高次の項”を切り落とす [9] ことで、$P(t)$ の固有値のうち 1 に近いものである 1 と $\alpha(t)$ は

$$\begin{aligned}(1-\lambda)(\alpha(t)-\lambda) =&\mathrm{Det}(P(t)-\lambda E)\\ \equiv&\mathrm{Det}(E+t\tilde{A}-\lambda E)\ (\mathrm{mod}\ o((t,1-\lambda)^2))\\ =&\mathrm{Det}\begin{pmatrix}1-\lambda-t & \frac{5}{3}t\\ t & 1-\lambda-\frac{5}{3}t\end{pmatrix}\end{aligned}$$

という式をみたすと予想される。1 は $P(t)$ が確率行列であるという仮定から誤差なしで 1 であり、$\alpha(t)$ は $1+ct+o(t)$ と表してみて上の式に代入すると

$$\begin{aligned}(1-\lambda)(1+ct+o(t)-\lambda) &= (1-\lambda)^2+ct(1-\lambda)+o((t,1-\lambda)^2)\\ \mathrm{Det}\begin{pmatrix}1-\lambda-t & \frac{5}{3}t\\ t & 1-\lambda-\frac{5}{3}t\end{pmatrix} &= (1-\lambda)^2-\frac{8}{3}t(1-\lambda)\end{aligned}$$

となることから $t(1-\lambda)$ の係数を比較することで $c=-\frac{8}{3}$ と求められる。

ちなみに $o((-\frac{1}{6}-\lambda,t)^3)$ の項を無視すれば $\beta_1(t)=-\frac{1}{6}+o(t)$ となり、$o((-\frac{1}{10}-\lambda,t)^3)$ を切り捨てることで $\beta_2(t)=-\frac{1}{10}-\frac{4}{3}t+o(t)$ となる。

次に固有ベクトルについて考えよう。$v_1(t), v_\alpha(t), v_{\beta_1}(t), v_{\beta_2}(t)$ をそれぞれ $P(t)$ の固有値 $1, \alpha(t), \beta_1(t), \beta_2(t)$ に対する固有ベクトルとしよう。ただし $s(v_1(t))=1$ としておく。

$P(t)$ に従うマルコフ連鎖を考えると、十分短い時間で固有値 $\beta_i(t)$ の成分は 0 に近い値なってしまうので考えないことにして [10]、固有値 1 と固有値 $\alpha(t)$ に対する固有ベクトルを求めてみよう。

$v_1(t)=a+tb+o(t), a={}^t(a_1,a_2,a_3,a_4), b={}^t(b_1,b_2,b_3,b_4)$ と表せるとする [11]。$P(t)v_1(t)=v_1(t)$ は

$$(P(0)+tA+o(t))(a+tb+o(t)) = P(0)a+t(P(0)b+Aa)+o(t) = a+tb+o(t)$$

となる。まず定数項を比較することで $P(0)a=a$, すなわち a が $P(0)$ の固有値 1 の固有ベクトル（か、零ベクトル）であることがわかるが、それは第 1 ブロック ${}^t(a_1,a_2)$ が P_1 に関する定常状態（の定数倍）で、第 2 ブロック ${}^t(a_3,a_4)$ が P_2 に関する定常状態（の定数倍）であることに他ならない。

故に $a_1:a_2=3:4$, $a_3:a_4=5:4$ である。前節では $a'_1=a_1+a_2$, $a'_2=a_3+a_4$ とおいて比の値 $a'_1:a'_2$ を考えた。これに相当する操作を行列を用いて表現すると次のようになる。行列

$$R=\begin{pmatrix}1&1&0&0\\0&0&1&1\end{pmatrix}$$

（これは Q_1^{-1} の第一行と Q_2^{-1} の第一行を取り出し、ずらして並べたものである）を考え、これを t の係数の等式 $P(0)b+Aa=b$ の両辺に左からかけるのである。すると

$$\begin{pmatrix}-a'_1+\frac{5}{3}a'_2+b_1+b_2\\ a'_1-\frac{5}{3}a'_2+b_3+b_4\end{pmatrix}=\begin{pmatrix}b_1+b_2\\ b_3+b_4\end{pmatrix}$$

[8] 一般的な記号なのか裏をとれていないが、$O(f_1,\ldots,f_n)$ は関数の層の特定の点における茎の $f_1,\ldots,f_n$ が生成するイデアル、$o(f_1,\ldots,f_n)$ はそれに含まれるイデアルの中で極大なものの共通部分、というような気持ちでこのような表記をした。

[9] この考え方の根拠となるのは補題 B.1 である。

[10] と言っても他の固有値と同様の考え方で近似値を求めることは可能であるが。

[11] 実際には $v_1(t)=a+o(1)$ までしか言えない。だが t の係数を比較することは可能であり、議論の筋は崩壊しない（次小節参照）。

という式が得られる。これは

$$\tilde{A}\begin{pmatrix} a_1' \\ a_2' \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

という式に他ならない。この式はさらに

$$\tilde{P}(t)\left(\begin{pmatrix} a_1' \\ a_2' \end{pmatrix} + o(t)\right) = (\tilde{P}(0) + t\tilde{A} + o(t))\left(\begin{pmatrix} a_1' \\ a_2' \end{pmatrix} + o(t)\right) = \begin{pmatrix} a_1' \\ a_2' \end{pmatrix} + o(t)$$

とも書き直せることを指摘しておく。

これより $a_1' : a_2' = 5 : 3$ となることがわかり、前小節と同じ結論

$$v_1(t) = a + o(1) = \left(\frac{15}{56}, \frac{5}{14}, \frac{5}{24}, \frac{1}{6}\right) + o(1)$$

が得られるのである。

$P(t)$ の固有値 $\alpha(t) = 1 - \frac{8}{3}t + o(t)$ に対する固有ベクトルについても全く同様の計算が可能である。$v_\alpha(t) = c + O(t)$, $c = {}^t(c_1, c_2, c_3, c_4)$ と書けるとすると、v_1 のときと同様に $P(0)c = c$ となり、$c_1' = c_1 + c_2$, $c_2' = c_3 + c_4$ とおけば

$$\tilde{A}\begin{pmatrix} c_1' \\ c_2' \end{pmatrix} = -\frac{8}{3}\begin{pmatrix} c_1' \\ c_2' \end{pmatrix}$$

となる。そしてこの式は

$$\tilde{P}(t)\left(\begin{pmatrix} c_1' \\ c_2' \end{pmatrix} + o(t)\right) = \left(1 - \frac{8}{3}t + o(t)\right)\left(\begin{pmatrix} c_1' \\ c_2' \end{pmatrix} + o(t)\right)$$

と書き直せるのである。

以上の考察で、ブロック行列に近い確率行列 $P(t)$ にしたがうマルコフ連鎖では、ごく短時間でブロック内では平衡状態に達し、その後の挙動は $\tilde{P}(t)$ という"各ブロックを一つの状態ととらえ直したマルコフ連鎖"に近い振る舞いをすることが数学的にも証明できそうだと納得していただけたのではないかと思う。

3.4 主結果

さて、本節の主定理を述べようと思うが、基本的には前小節での考察を正確に記号を用意して一般的に述べるだけである。ただ、現状私ができた証明では"陰関数定理"を用いるので、高次の項の滑らかさを仮定せねばならない。

定理 3.5. ε を正の実数、$f(t)$ を $t \in [0, \varepsilon)$ で定義された $f(0) = 0$ なる単調増加連続関数とし、次の条件をみたす確率行列の族 $P(t) \in \mathcal{P}(m)$ を考える：まず $P(t)$ は

$$P(t) = \begin{pmatrix} P_1 & & \\ & \ddots & \\ & & P_n \end{pmatrix} + f(t)A + o(f(t))$$

と表せ、これに関して次が成り立つとする。

- ブロックのサイズは $A_{i,j}$ が $m_i \times m_j$ 行列で、P_i は m_i 次正方行列である。
- 各 P_i はエルゴード的な確率行列である。

- $s = f(t)$ とおくと s をパラメータとする確率行列の族 $P(f^{-1}(s))$ は C^n 級である。

Q_i を

$$Q_i^{-1}P_iQ_i = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \beta_{i,1} & & * \\ \vdots & & \ddots & \\ 0 & O & & \beta_{i,m_i-1} \end{pmatrix} \quad \text{(1 行目が (1,1) 成分以外 0 の上三角行列)}$$

となる正則行列し、

$$Q = \begin{pmatrix} Q_1 & & \\ & \ddots & \\ & & Q_n \end{pmatrix}$$

とする。ただし Q_i の第 1 列 u_i は $s(u_i) = 1$ となるようにとる。

$e_i = {}^t(1, 0 \ldots, 0) \in \mathbb{R}^{m_i}$（第 1 成分のみが 1 でその他の成分が 0 の m_i 次元列ベクトル）とし、$m \times n$ 行列 U を

$$U = \begin{pmatrix} e_1 & & \\ & \ddots & \\ & & e_n \end{pmatrix}$$

で定める。そして

$$\tilde{P}(t) := {}^tUQ^{-1}P(t)QU, \qquad \tilde{A} := {}^tUQ^{-1}AQU$$

とおく。

さらに $\tilde{A}$ の固有値 $a_1, \ldots, a_n$ には重複がないと仮定する。

以上の仮定と記号の下で以下が成り立つ。

(1) $a_i = 0$ となる $i \in \{1, \ldots, n\}$ が存在する。以下ではそれは $i = 1$ であるとする。

(2) $\tilde{P}(t)$ の固有値は $\alpha_1(t) = 1,\ \tilde{\alpha}_i(t) = 1 + a_i f(t) + o(f(t))\ (i \in \{2, \ldots, n\})$ と表せる。

(3) $P(t)$ の固有値は $\alpha_1(t) = 1,\ \alpha_i(t) = 1 + a_i f(t) + o(f(t))\ (i \in \{2, \ldots, n\}),\ \beta_{i,i'}(t) = \beta_{i,i'} + o(1)$ $(i \in \{1, \ldots, n\}, i' \in \{1, \ldots, m_i - 1\})$ と表せる。

(4) 固有値 $\tilde{\alpha}_i(t)$ に属する $\tilde{P}(t)$ の固有ベクトル $\tilde{v}_i$ は $\tilde{A}$ の固有値 a_i の固有ベクトル w_i を用いて $\tilde{v}_i = w_i + o(1)$ と表せる。

(5) 固有値 $\alpha_i(t)$ に属する $P(t)$ の固有ベクトル $v_{\alpha_i}(t)$ はいずれもブロックごとに定常分布であり、その比率は前項の固有ベクトルの係数で表される。正確に書けば、P_i の定常状態 u_i および前項の w_i を $u_i = {}^t(d_{i,k})_{k=1}^{m_i}$, $w_i = {}^t(c_{i,j})_{j=1}^{n} \in \mathbb{C}^n$ と表すと、

$$v_{\alpha_i}(t) = {}^t(c_{i,1}d_{1,1}, \ldots, c_{i,1}d_{1,m_1}, \ldots, c_{i,n}d_{n,1}, \ldots, c_{i,n}d_{n,m_n}) + o(1) \in \mathbb{C}^{m_1} \times \cdots \times \mathbb{C}^{m_n} = \mathbb{C}^m$$

が成り立つ。

記号がやたらごちゃごちゃしているが、これらが前小節で述べたことを一般的に述べただけであることが理解していただけると思う。言うまでのないと思うが、${}^tUQ^{-1}$ と QU で挟む操作は各ブロックの定常状態の成分のみを取り出す操作であることを補足しておく。

証明：

(1) 全ての t に対して $P(t)$ が確率行列であるという仮定から A は m 本の行ベクトルの和が 0 ベクトルになる。エルゴード的な確率行列 P_i の 1 以外の固有値に対する固有ベクトルは全て成分の和が 0 であるから、$\tilde{A}$ も n 本の行ベクトルの和が 0 という性質を持つ。ゆえに $\tilde{A}$ が正則でないので 0 が固有値になる。

(2) S を

$$S^{-1}\tilde{A}S = \mathrm{diag}(0, a_2, \ldots, a_n)$$

となる n 次正方行列とする。ただし $\mathrm{diag}(x_1, \ldots, x_n)$ は $x_1, \ldots, x_n$ が順に対角成分に並ぶ対角行列を表す。このとき

$$S^{-1}(1 - \lambda E + f(t)\tilde{A})S = \mathrm{diag}(1-\lambda, 1-\lambda+a_2 f(t), \ldots, 1-\lambda+a_n f(t))$$

であるから $\tilde{P}(t)$ の特性多項式は

$$\begin{aligned}\mathrm{Det}(\tilde{P}(t) - \lambda E) \equiv& \mathrm{Det}((1-\lambda)E + f(t)\tilde{A}) \pmod{o((f(t), 1-\lambda)^n)} \\ =& \prod_{i=1}^{n}(1 - a_i f(t) - \lambda)\end{aligned}$$

となることより、後述の補題 B.1（高次特異点に対する陰関数定理）により $\tilde{\alpha}_i(t) = 1 + a_i f(t) + o(f(t))$ の形をしていることがわかる。その際 $\tilde{P}(t)$ が確率行列であることから t 毎に $\tilde{\alpha}_i(t) = 1$ となる i がなければならないが、$a_i \neq 0$ $(i \geq 2)$ なので十分小さな $t \neq 0$ に対しては、$\tilde{\alpha}_i(t) = 1$ となるのは $i = 1$ 以外はあり得ない。

(3) $\beta_{i,i'}$ についての主張は固有値の連続性からただちにしたがう。

Q の列ベクトルのうち、固有値 1 の固有ベクトルが先に来るように並び替える基底の取り替え行列を K とすると

$$K^{-1}Q^{-1}P(t)QK = \begin{pmatrix} \tilde{P}(t) & O(f(t)) \\ O(f(t)) & B(t) \end{pmatrix}$$

の形になる。$\tilde{P}(t) - E \in O(f(t))$ であることに注意すると

$$\mathrm{Det}(P(t) - \lambda E) \equiv \mathrm{Det}((1-\lambda)E + f(t)\tilde{A})\mathrm{Det}(B(t) - \lambda E) \pmod{O((f(t), 1-\lambda)^{n+1}}$$

となる。$B(t) - B(0) \in O(f(t))$ および

$$\mathrm{Det}(B(0) - E) = \prod_{(i,i')}(1 - \beta_{i,i'}) \neq 0$$

であることを踏まえると

$$\mathrm{Det}(P(t) - \lambda E) \equiv \mathrm{Det}((1-\lambda)E + f(t)\tilde{A}) \prod_{(i,i')}(1 - \beta_{i,i'}) \pmod{O((f(t), 1-\lambda)^{n+1}}$$

となる。つまり $s = f(t)$ と変数変換すると $P(f^{-1}(s))$ の特性多項式と $\tilde{P}(f^{-1}(s))$ の特性多項式の s による n 階微分とはちょうど $\mathrm{Det}(B(0) - E)$ 倍の関係になっているので、再び補題 B.1 により $\alpha_i(t) = 1 + a_i f(t) + o(f(t))$ の形をしていることがわかる。

(4) 固有ベクトルを並べて作った行列を $S(t) = (\tilde{v}_1(t), \ldots, \tilde{v}_n(t))$ とおく。

$$\tilde{P}(t)S(t) = S(t)\mathrm{diag}(1, \tilde{\alpha}_2(t), \ldots, \tilde{\alpha}_n(t))$$

に $\tilde{P}(t) = E + f(t)\tilde{A} + o(f(t))$, $S(t) = S(0) + (S(t) - S(0))$, $\tilde{\alpha}_i = 1 + a_i f(t) + o(f(t))$ を代入して整理すると

$$
\begin{aligned}
&S(0) + (S(t) - S(0)) + f(t)\tilde{A}S(0) + o(f(t)) \\
=&S(0) + (S(t) - S(0)) + f(t)S(0)\mathrm{diag}(0, a_2, \ldots, a_n) + o(f(t))
\end{aligned}
$$

となる。

この両辺に共通の項を消去し、$f(t)$ で割って $t \to 0$ の極限をとれば $\tilde{A}S(0) = S(0)\mathrm{diag}(0, a_2, \ldots, a_n)$ が得られる。これは $S(0)$ の各列が $\tilde{A}$ の固有ベクトルであることを意味する。

(5) $P(t)v_{\alpha_i}(t) = \alpha_i v_{\alpha_i}(t)$ に $P(t) = P(0) + f(t)A + o(f(t))$, $v_{\alpha_i}(t) = v_{\alpha_i}(0) + (v_{\alpha_i}(t) - v_{\alpha_i}(0))$ および $\alpha_i = 1 + a_i f(t) + o(f(t))$ を代入すると

$$
\begin{array}{rlll}
 & P(0)v_{\alpha_i}(0) & +P(0)(v_{\alpha_i}(t) - v_{\alpha_i}(0)) + f(t)Av_{\alpha_i}(0) & +o(f(t)) \\
= & v_{\alpha_i}(0) & +(v_{\alpha_i}(t) - v_{\alpha_i}(0)) + f(t)a_i v_{\alpha_i}(0) & +o(f(t))
\end{array}
\tag{2}
$$

となる。

まずこの式に $t = 0$ を代入すると

$$P(0)v_{\alpha_i}(0) = v_{\alpha_i}(0)$$

となる。これはブロックごとに $P(0)$ の定常分布（の定数倍）であることを意味する。言い換えると $w_i' = (c_{i,j}')_j \in \mathbb{C}^n$ を用いれば求める形に表せる。

この w_i' が $\tilde{A}$ の固有値 a_i の固有ベクトルであることを示そう。

${}^tUQ^{-1}$ を左からかけることは“定常状態の成分”を取り出すことに他ならないので、${}^tUQ^{-1}P(0) = {}^tUQ^{-1}$ が成り立つ。

このことに注意すると、(2) の両辺に左から ${}^tUQ^{-1}$ をかけることで

$$f(t){}^tUQ^{-1}Av_{\alpha_i}(0) + o(f(t)) = f(t)a_i v_{\alpha_i}(0) + o(f(t))$$

を得る。これより ${}^tUQ^{-1}Av_{\alpha_i}(0) = a_i v_{\alpha_i}(0)$ となる。

定義より $w_i' = {}^tUQ^{-1}v_{\alpha_i}(0)$ であり、また $v_{\alpha_i}(0)$ が定常状態の成分しか持たないことから $QUw_i' = v_{\alpha_i}(0)$ となる。これらのことに注意すると

$$
\begin{aligned}
\tilde{A}w_i' =&({}^tUQ^{-1}AQU)w_i' \\
=&{}^tUQ^{-1}(Av_{\alpha_i}(0)) \\
=&{}^tUQ^{-1}(a_i v_{\alpha_i}(0)) \\
=&a_i w_i'
\end{aligned}
$$

となる。

3.5 多階層構造ついて

例えば $f(t) = o(1)$, $g(t) = o(f(t))$ という収束の速さの異なる関数によって

$$P(t) = \begin{pmatrix} P_1 & 0 & 0 \\ 0 & P_2 & 0 \\ 0 & 0 & P_3 \end{pmatrix} + f(t)\begin{pmatrix} A_{11} & A_{12} & 0 \\ A_{21} & A_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix} + g(t)B + o(g(t))$$

というように表される場合を考えてみよう。

この場合まず、開始後のごく短時間のうちに第 1, 第 2, 第 3 ブロックのそれぞれで定常状態に到達し、その後 $1/f(t)$ 程度の時間スケールで第 1 ブロックと第 2 ブロックの間で $E + f(t)\tilde{A}$ の形の行列で表されるマルコフ連鎖が進行するであろう。そして少しすれば第 1 ブロックと第 2 ブロックとの間でも定常状態に到達した大ブロックを形成する。そして $1/g(t)$ のスケールの時間が経つと今度はその大ブロックと第 3 ブロックとの間でマルコフ連鎖が進行するであろう。

やや不正確だが次の表のような感じである。

時間スケール	現象
$\Theta(1)$	各状態間で進行
$\Omega(1)\sim o(1/f(t))$	小ブロック内で定常状態
$\Theta(1/f(t))$	小ブロック間で進行
$\Omega(1/f(t))\sim o(1/g(t))$	大ブロックで定常状態
$\Theta(1/g(t))$	大ブロック間で進行
$\Omega(1/g(t))$	全体として定常状態

前小節までに加えて数学としてやるべきことは全くないし、かっちりした主張を書いても煩雑なだけだと思うのでやめておく。

ただ有限種類の状態のみ存在し、反応速度のスケールも有限種類しかない場合はそれで良いのだが、状態を無限に用意した系では反応速度のスケールも連続ないし稠密な分布をする系が現れるのが（数学的には）自然だと思う。その場合には別の考察が必要であろう。

4　不安定状態の解消

4.1　モデルの設定

本節の目的は確率行列 P で定まるマルコフ過程が「不安定な状態」を持つ場合、「不安定な状態を除外した状態集合上でのマルコフ過程によってそれを近似することができる」ということを定式化、証明することである。

状態 $s \in S$ が不安定であるというのは「状態 $s \in S$ に存在する確率は極めて小さい」というようなイメージであるが、確率ベクトルとしては s の成分のみに局在したものが定義としては含まれてしまうわけで、確率ベクトルの非存在によって不安定状態の集合 S' を定式化するわけにはいかない。しかし、どんな初期値から出発しても P にしたがって時間発展すれば 0 に向かうような成分は存在する場合がある。そのような状態を不安定状態と呼ぶのは自然であろう。

前節と同様に、大まかな問題設定を書くなら次のようになるだろうか。

- S は有限集合、$P = (p_{i,j})_{(i,j)\in S}$：エルゴード的な S 上の確率行列である。
- $S' \subset S$.
- 十分大きな n および全ての $j \in S$ に対し、$\sum_{i\in S'} p_{i,j}^{(n)} \fallingdotseq 0$

ここでもやはり $P(0)$ では“S' の状態は全て不安定”という条件をみたす確率行列の族 $P(t)$ を考えて、t が小さい値のときに $P(t)$ にしたがうマルコフ過程がどのような挙動をするかを考察してみたいと思う。

定義 4.1.　(1) 確率行列 $P = (p_{i,j}) \in \mathcal{P}(S)$ に対し、次の条件をみたす最大の集合 $S' \subset S$ を P に関する不安定状態の集合という。

- 任意の $i \in S', j \in S \setminus S'$ に対し $P_{i,j} = 0$.
- 各 $j \in S$ に対し、$p_{i,j}^{(n)} > 0$ をみたす $i \in S \setminus S'$ と正の整数 n が存在する。

(2) S 上の確率行列の族 $P(t)$（正確に書けば連続写像 $P : [0, \varepsilon) \to \mathcal{P}(S)$）の不安定状態の集合とは $P(0)$ の不安定状態の集合のことをいう。

(3) 不安定状態の集合の元を不安定状態という。

(2) と (1) の定義は論理的にはとても紛らわしいが、実際には確率行列の族を一つ固定した上での議論しかしないので、混同の恐れはほとんどないと思う。

注意 4.2. 上の定義において $j \in S'$ の状態にあるものは n ステップ後に $p_{i,j}^{(n)}$ 以上の確率で S' から脱出し、その後 S' に戻ってくることは一切ない。全ての $j \in S'$ に対して一定時間（$S' \subset S$ が有限なので、各 j に対する n の上限）で正の確率（$S' \subset S$ は有限集合なのでその下限も正）で S' を脱出するから、時間が十分に経過すれば S' に存在する確率はほぼ 0 になる。

この雰囲気では S' からなくなるまでにはずいぶんと長い時間がかかるように感じられるかもしれないが、その時間は P のみに依存して定まる有限の時間なので、$P(0) = P$ となる確率行列の族 $P(t)$ を固定した上で t を十分に小さくとれば、$P(t)$ にしたがうマルコフ過程ではやはり $P(0)$ における不安定状態 S' は短い時間（t に依存しない有限時間の意味）のうちにすっからかんになってしまうのである。

定義から直ちにわかることとして、定常状態では不安定状態での存在確率は（ほぼ）0 となる。もう少し強く、次のことが成り立つ。

補題 4.3. $P = (p_{i,j}) \in \mathcal{P}(S)$, α を P の $|\alpha| \neq 1$ なる固有値、$v = (v_i)_{i \in S} \in \mathbb{C}^S$ を $\alpha v = v$ なるベクトルとする。$i \in S$ が P に関する不安定状態であるならば $v_i = 0$ である。

証明： S' を P の不安定状態の集合とし、$a = \sum_{i \in S'} |v_i|$ を考える。$j_0 \in S'$, $v_{j_0} \neq 0$ とする。$p_{i_0,j_0}^{(n)} > 0$ となる $n > 0, i_0 \in S \setminus S'$ をとると

$$
\begin{aligned}
a &= \sum_{i \in S'} |v_i| \\
&= \sum_{i \in S'} \left| \sum_{j \in S} p_{i,j}^{(n)} v_j \right| \\
&\leq \sum_{i \in S'} \sum_{j \in S} p_{i,j}^{(n)} |v_j| \\
&= \sum_{j \in S'} \sum_{i \in S'} p_{i,j}^{(n)} |v_j| \qquad (i \in S', j \in S \setminus S' \text{ に対して } p_{i,j}^{(n)} = 0) \\
&\leq \sum_{j \in S' \setminus \{j_0\}} |v_j| + (1 - p_{i_0,j_0}^{(n)}) |v_{j_0}| \\
&= a - p_{i_0,j_0}^{(n)} |v_{j_0}| < a
\end{aligned}
$$

となり矛盾が生じる。

4.2　例

例 4.4. 確率行列

$$P(t) = \begin{pmatrix} \frac{1}{3} & \frac{1}{2}-3t & t & 2t & 0 \\ \frac{2}{3}-t & \frac{1}{2} & 0 & \frac{1}{3} & 0 \\ 0 & t & 1-2t & 0 & \frac{1}{3} \\ t & 2t & 0 & \frac{1}{2}t & \frac{1}{3} \\ 0 & 0 & t & \frac{1}{6}-3t & \frac{1}{3} \end{pmatrix}$$

にしたがうマルコフ連鎖を考えよう。$P(0)$ にしたがうマルコフ連鎖では第 1 状態と第 2 状態にあるものはその中から出られないので、これを一まとめにしたものを第 1 ブロックと呼び、同様の理由で第 3 状態のみからなる集合を第 2 ブロックと呼ぶことにする。第 4 状態と第 5 状態は、十分に時間が経てばいずれも第 1 ブロックか第 2 ブロックに落ちる。どちらにどの比率で落ちるのかをまず考えよう。

計算は省略するが、$P(0)$ の固有値は $1, -\frac{1}{6}, 1, \frac{2}{3}, \frac{1}{6}$ であり、それらの固有ベクトルとして

$$v_1 = \begin{pmatrix} \frac{2}{7} \\ \frac{5}{7} \\ 0 \\ 0 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, v_4 = \begin{pmatrix} -6 \\ -4 \\ -5 \\ 10 \\ 5 \end{pmatrix}, v_5 = \begin{pmatrix} 3 \\ -1 \\ -2 \\ -5 \\ 5 \end{pmatrix}$$

がとれる。v_1 と $_3$ はそれぞれ第 1 ブロック、第 2 ブロックの定常状態である。固有ベクトルからなる基底を用いて、不安定状態に局在する状態ベクトルを表してみると

$$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \frac{4}{5}v_1 + \frac{13}{35}v_2 + \frac{1}{5}v_3 + \frac{1}{15}v_4 - \frac{1}{15}v_5, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \frac{2}{5}v_1 - \frac{4}{35}v_2 + \frac{3}{5}v_3 + \frac{1}{15}v_4 + \frac{2}{15}v_5$$

となる。十分な時間の後には $P(0)$ の 1 でない固有値に対する固有ベクトルの係数は消えて安定状態に落ち着くわけだがその比率は v_1 と v_3 の係数の比になりそうである。つまり第 4 状態からスタートすれば第 1 ブロック : 第 2 ブロック $= 4:1$, 第 5 状態からスタートすれば $2:3$ である。

なお後の話との関連で、係数を並べた列ベクトル ${}^t\left(\frac{4}{5}, \frac{13}{35}, \frac{1}{5}, \frac{1}{15}, -\frac{1}{15}\right)$, ${}^t\left(\frac{2}{5}, -\frac{4}{35}, \frac{3}{5}, \frac{1}{15}, \frac{2}{15}\right)$ は固有ベクトルを並べた行列 $(v_1, v_2, v_3, v_4, v_5)$ の逆行列の第 4 列と第 5 列であることを注意しておく。

$P(t)$ にしたがうマルコフ連鎖では、安定状態にあるものもごく稀に不安定状態になるわけだが、あっという間（$\frac{1}{t}$ に比べて短い時間のうちに、という意味である）に安定状態に戻るだろう。そしてその比率は第 4 状態になったものなら（t は微少なので第 4 列の t の項は無視できて）$4:1$ で、第 5 状態になった場合なら $2:3$ の比率で各ブロックに分配されると思われる。

さて、$P(t)$ にしたがって $\frac{1}{t}$ よりは短いが十分に時間が経過したとき、各ブロックでは定常状態に達し、不安定状態にはほぼ存在していないと予想される。つまり第 1 ブロックに存在する割合 a'_1 と第 2 ブロックに存在する割合 a'_2 を用いて状態ベクトルは近似的に

$$v = {}^t\left(\frac{2}{7}a'_1, \frac{5}{7}a'_1, a'_2, 0, 0\right)$$

と表されるであろう。この状態の次の瞬間の分布は

$$P(t)v = \begin{pmatrix} \frac{1}{3}\frac{2}{7}a_1' + (\frac{1}{2} - 3t)\frac{5}{7}a_1' + ta_2' \\ (\frac{2}{3} - t)\frac{2}{7}a_1' + \frac{1}{2}\frac{5}{7}a_1' \\ t\frac{5}{7}a_1' + (1-2t)a_2' \\ t\frac{2}{7}a_1' + 2t\frac{5}{7}a_1' \\ ta_2' \end{pmatrix} = v + t\begin{pmatrix} -\frac{15}{7}a_1' + a_2' \\ -\frac{2}{7}a_1' \\ \frac{5}{7}a_1' - 2a_2' \\ \frac{12}{7}a_1' \\ a_2' \end{pmatrix}$$

となるだろう。安定状態のブロック内では前節で考察したとおりにブロック内でごく短い時間に定常状態に向かい、不安定状態に登ったものについては先述の割合で安定状態に分配されるという予想の下では、ブロック間の時間発展を記述する行列 $\tilde{P}(t)$ は

$$\tilde{P}(t)\begin{pmatrix} a_1' \\ a_2' \end{pmatrix} = \begin{pmatrix} a_1' \\ a_2' \end{pmatrix} + t\begin{pmatrix} (-\frac{15}{7}a_1' + a_2') + (-\frac{2}{7}a_1') + \frac{4}{5}\frac{12}{7}a_1' + \frac{2}{5}a_2' \\ \frac{5}{7}a_1' - 2a_2' + \frac{1}{5}\frac{12}{7}a_1' + \frac{3}{5}a_2' \end{pmatrix} = \begin{pmatrix} 1 - \frac{37}{35}t & \frac{7}{5}t \\ \frac{37}{35}t & 1 - \frac{7}{5}t \end{pmatrix}\begin{pmatrix} a_1' \\ a_2' \end{pmatrix}$$

となるわけである。

後で引用する都合上、第 j ブロックから不安定状態を経由して第 i ブロックへ移動する割合を表した行列

$$t\begin{pmatrix} \frac{4}{5}\frac{12}{7} & \frac{2}{5} \\ \frac{1}{5}\frac{12}{7} & \frac{3}{5} \end{pmatrix}$$

を間接移動項とでも名付けておこう。

補足 4.5. ごく短い時間であってもブロック内で定常状態とは異なる分布を経由するものが“次の瞬間に”このように変化する、という言い回しは奇妙な印象を受けるかもしれない。上記のことをもう少し正確に言えば次のようになる。変化が 1 回起こるのに要する（平均）時間は不安定状態にある時間を考慮しなければ $\frac{1}{t}$ のスケールであるのに対し、不安定状態にある時間 ε（これは $1/t$ よりははるかに短い時間である）も含めれば $\frac{1}{t} + \varepsilon = \frac{1+t\varepsilon}{t} \fallingdotseq \frac{1}{t(1-t\varepsilon)}$ 程度になる。$P(t)$ が $o(t)$ 程度ずれることによる効果は無視するという発想を受け入れるとすれば、これらの差異も同様に無視できるわけである。

4.3 主定理

定理 4.6. ε を正の実数、$f(t)$ を $t \in [0, \varepsilon)$ で定義された $f(0) = 0$ なる単調増加連続関数とし、次の条件をみたす確率行列の族 $P(t) \in \mathcal{P}(m)$ を考える：$P(t)$ は

$$P(t) = \begin{pmatrix} P_1 & & & P_{1,n+1} \\ & \ddots & & \vdots \\ & & P_n & P_{n,n+1} \\ & & & P_{n+1,n+1} \end{pmatrix} + f(t)\begin{pmatrix} A_{1,1} & \cdots & A_{1,n} & A_{1,n+1} \\ \vdots & & \vdots & \vdots \\ A_{n,1} & \cdots & A_{n,n} & A_{n,n+1} \\ A_{n+1,1} & \cdots & A_{n+1,n} & A_{n+1,n+1} \end{pmatrix} + o(f(t))$$

と表せ、これに関して次が成り立つとする。

- 各 $P_i \quad (i \in \{1, \ldots, n\})$ はエルゴード的な確率行列である。
- 第 $(n+1)$ ブロックに対応する状態は全て不安定である。
- $s = f(t)$ とおくと s をパラメータとする確率行列の族 $P(f^{-1}(s))$ は C^n 級である。

以下正方行列 $P_i, P_{n+1,n+1}, P$ のサイズをそれぞれ m_i, m_{n+1}, m と表す。Q_i を

$$Q_i^{-1}P_iQ_i = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \beta_{i,1} & & * \\ \vdots & & \ddots & \\ 0 & O & & \beta_{i,m_i-1} \end{pmatrix}$$

となる正則行列とし、Q は

$$Q = \begin{pmatrix} Q_1 & & & R_1 \\ & \ddots & & \vdots \\ & & Q_n & R_n \\ & & & R_{n+1} \end{pmatrix}$$

の形の正則行列で、$Q^{-1}P(0)Q$ は上三角行列になるとする [12]。また Q_i $(i \in \{1,\dots,n\})$ の第 1 列 v_i は $s(v_i) = 1$ をみたすとする。

$e_i = {}^t(1, 0\dots, 0) \in \mathbb{R}^{m_i}$（第 1 成分のみが 1 でその他の成分が 0 の m_i 次元列ベクトル）とし、$m \times n$ 行列 U を

$$U = \begin{pmatrix} e_1 & & \\ & \ddots & \\ & & e_n \\ 0 & \cdots & 0 \end{pmatrix}$$

で定める（最下段の 0 は m_{n+1} 次の列零ベクトルを表す）。そして

$$\tilde{P}(t) := {}^tUQ^{-1}P(t)QU, \qquad \tilde{A} := {}^tUQ^{-1}AQU$$

とおく。

さらに $\tilde{A}$ の固有値 $0, a_2, \dots, a_n$ には重複がないと仮定する。

以上の仮定と記号の下で以下が成り立つ。

(1) $P(0)$ の固有値 1 の重複度は n である。

(2) $P(0)$ の 1 以外の固有値の絶対値は 1 より真に小さい。

(3) $P(t)$ の固有値をとる連続関数で、$t = 0$ で 1 となるものは n 個あり、1（定数関数）、および $\alpha_i(t) = 1 + a_i f(t) + o(f(t))$ $(i \in \{2,\dots,n\})$ の形の関数、である。

(4) 固有値 1 および $\alpha_i(t)$ に属する $P(t)$ の固有ベクトル $v_i(t)$ $(i \in \{1, 2, \dots, n)$ は

$$v_{\alpha_i}(t) = {}^t(c_{i,1}v_1, \dots, c_{i,n}v_n, 0) + o(1) \in \mathbb{C}^{m_1} \times \cdots \times \mathbb{C}^{m_n} \times \mathbb{C}^{m_{n+1}} = \mathbb{C}^m$$

と表すことができる。

(5) 前項の $(c_{i,j})_{j=1}^n$ は $\tilde{A}$ の固有値 a_i の固有ベクトルである。

---

[12] 実際にはもう少し弱い条件

$$\lim_{n\to\infty} P(0)^n \begin{pmatrix} R_1 \\ \vdots \\ R_{n+1} \end{pmatrix} = 0$$

で十分である。

証明： 基本的には前節の主定理と同様であるので異なる部分や前節にはなかった主張についてのみ軽くコメントするにとどめる。

(1) および (2) は補題 4.3 からしたがう事実である。

(3) について：

前節と同様に安定ブロックの中の固有値 1 の固有空間を先に来るように基底を取り替えると

$$K^{-1}Q^{-1}P(t)QK = \begin{pmatrix} \tilde{P}(t) & O(f(t)) & * \\ O(f(t)) & B(t) & * \\ O(f(t)) & O(f(t)) & C(t) \end{pmatrix}$$

の形になることが Q の形からわかる。（第 1 ブロックが固有値 1 の固有空間、第 2 ブロックが安定状態のその他の固有空間、第 3 ブロックが不安定状態である。）

前節同様 $\mathrm{Det}(B(0)-E) \neq 0$ である。また補題 4.3 より $\mathrm{Det}(C(0)-E) \neq 0$ もしたがう。

その他は前節の議論を繰り返せばよい。

証明の変更点は以上だが、それよりもこの主張が前小節で考えたような計算法の正当化になっていることをきちんと説明しておかねばならないだろう。

まず Q^{-1} を

$$Q^{-1} = \begin{pmatrix} Q_1^{-1} & & & S_1 \\ & \ddots & & \vdots \\ & & Q_n^{-1} & S_n \\ & & & S_{n+1} \end{pmatrix}$$

と表しておく。具体的には $S_i = -Q_i^{-1}R_iR_{n+1}^{-1}$ であるが、それはどうでもよく、不安定状態を Q の縦ベクトルの一次結合で表したときの係数であることが重要である。$\tilde{P}$ は Q の列ベクトルを基底に取り直して、第 1 ブロックから第 n ブロックまでの $(1,1)$ 成分のみを取り出したものであった。

$${}^tUQ^{-1}P(0)QU = E$$

であるから次の次数の項として $f(t)Q^{-1}AQ$ がどのようなものかを見てみよう。

$$\begin{aligned}
&Q^{-1}AQ \\
&= \begin{pmatrix} Q_1^{-1} & & & S_1 \\ & \ddots & & \vdots \\ & & Q_n^{-1} & S_n \\ & & & S_{n+1} \end{pmatrix} \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} & A_{1,n+1} \\ \vdots & & \vdots & \vdots \\ A_{n,1} & \cdots & A_{n,n} & A_{n,n+1} \\ A_{n+1,1} & \cdots & A_{n+1,n} & A_{n+1,n+1} \end{pmatrix} \begin{pmatrix} Q_1 & & & R_1 \\ & \ddots & & \vdots \\ & & Q_n & R_n \\ & & & R_{n+1} \end{pmatrix} \\
&= \begin{pmatrix} Q_1^{-1} & & & S_1 \\ & \ddots & & \vdots \\ & & Q_n^{-1} & S_n \\ & & & S_{n+1} \end{pmatrix} \begin{pmatrix} A_{1,1}Q_1 & \cdots & A_{1,n}Q_n & \sum_{k=1}^{n+1} A_{1,k}R_k \\ \vdots & & \vdots & \vdots \\ A_{n,1}Q_1 & \cdots & A_{n,n}Q_n & \sum_{k=1}^{n+1} A_{n,k}R_k \\ A_{n+1,1}Q_1 & \cdots & A_{n+1,n}Q_n & \sum_{k=1}^{n+1} A_{n+1,k}R_k \end{pmatrix} \\
&= \begin{pmatrix} Q_1^{-1}A_{1,1}Q_1 + S_1A_{n+1,1}Q_1 & \cdots & Q_1^{-1}A_{1,n}Q_n + S_1A_{n+1,n}Q_n & Q_1^{-1}\sum_{k=1}^{n+1} A_{1,k}R_k + S_1\sum_{k=1}^{n+1} A_{n+1,k}R_k \\ \vdots & & \vdots & \vdots \\ Q_n^{-1}A_{n,1}Q_1 + S_nA_{n+1,1}Q_1 & \cdots & Q_n^{-1}A_{n,n}Q_n + S_nA_{n+1,n}Q_n & Q_n^{-1}\sum_{k=1}^{n+1} A_{n,k}R_k + S_n\sum_{k=1}^{n+1} A_{n+1,k}R_k \\ S_{n+1}A_{n+1,1}Q_1 & \cdots & S_{n+1}A_{n+1,n}Q_n & S_{n+1}\sum_{k=1}^{n+1} A_{n+1,k}R_k \end{pmatrix}
\end{aligned}$$

となる。この (i,j) ブロック（$i,j \in \{1,\ldots,n\}$）は

$$Q_i^{-1}A_{i,j}Q_j + S_iA_{n+1,j}Q_j$$

である。この第 1 項の $(1,1)$ 成分は前節でも考えたもので、第 j ブロックの“定常状態”から第 i ブロックに移動する割合を表す。第 2 項はどのようなものであるかというと

- $f(t)A_{n+1,j}Q_j$ の第 1 列は第 j ブロックから各不安定状態に移動する確率（の $f(t)$ のスケールの項）を並べたものであり、
- S_i の第 1 行は各不安定状態から第 i ブロックへ移動する確率を並べたもの（前節参照）

に他ならない。つまり $f(t)S_iA_{n+1,j}Q_j$ の $(1,1)$ 成分はまさに前小節で考えた間接移動項（の $f(t)$ のスケールの成分）に他ならないのである。

補足 4.7. 状態が有限種類でない系を考える場合、“個々の状態は全て不安定”とでも呼ぶべき事態が数学的には普通にあり得ると思う。ここの議論を一般化する場合には何らかの“安定な状態が存在する”という意味合いの仮定をおく必要があるだろうが、何が最適な仮定なのかは今の私にはよくわからない。

A　補足というか言い訳的な何か

マルコフ連鎖というのは、言ってみれば「（今の）自分自身の状態のみで次の瞬間の自分自身の状態が（確率的に）わかる」というものである。

$$2\,\mathrm{H_2O} \quad \leftrightarrow \quad \mathrm{H_3O^+} + \mathrm{OH^-}$$

という化学反応式は 4 つの水素原子と 2 つの酸素原子のみからなる孤立した系において、その系がどのように時間変化するかを述べたものではなく、アボガドロ数に近い数の原子が入り交じった系において起こるであろう反応の局所的なところを切り取ったものであるはずである。

一つの $\mathrm{H_3O^+}$ に注目したとき、それが $\mathrm{H_2O}$ に変化する確率は、高校化学的な単純な描像ですら $\mathrm{OH^-}$ の濃度という環境に依存した値となる。

そのため今回考えたマルコフ連鎖の系は、このような単純な化学反応式ですら、状況が直接は対応していないのである。

今回のモデルを“実際の”現象に適用できるとすれば、次のような状況・方法であろうか。

(1) アボガドロ数程度の原子からなる閉じた系を考えること。上の例で言えば、不安定状態として消去されるのは陽子が孤立しているなどの状態であり、同一視されるのは結合距離や分子間の距離が違う状態である。その結果として 3 種類の分子のとり得る個数ごと（つまりアボガドロ数程度の膨大な）に現れる“準安定”な状態“のみ”を考えれば良い、という近似である。

(2) 分子内での結合の変化のように、少数の原子だけで完結した記述が可能な化学反応については、現実的な個数の安定状態の想定でモデルが構築できるかもしれない。

(3) 水溶液中の $\mathrm{H_2O}$ のような、多少の反応によって濃度がほとんど変化しない物質のみが消費・生産される反応においてはそこそこ良い近似ができるかもしれない。

個々の現象の理解に応用するには今回のモデルは決して有力ではないと思うが、それでも今回のモデル自体が無意味であるとは現時点では私は思っていない。

始めに述べたことであるが改めて明言しておきたい。今回このようなモデルを考えた（私自身の）動機は

- 短時間でも存在する不安定な状態を無視できるのはなぜか、
- 分子内の原子間の距離が多少変わっても「結合している」状態であると一まとめに考えても支障がないのはなぜか

ということに対する理解を得たかったというものである。

なぜ単純化できるのか、という哲学的な問いに対しては今回のモデルが一つの解答を与えているとは思う。

とは言えある程度一般的な化学反応式、せめて上で挙げた程度の単純な化学反応式くらいは取り込めて、かつアボガドロ数的な複雑なものには思いを馳せなくてもよいモデルは哲学的にも、科学的にも求められているであろう。残念ながらそこまでは至れなかったことは私としても残念であるし、何らかの解答があればご教授いただければ幸いである。

B　微積分からの補題

補題 B.1 (特異点における陰関数定理)**.** n を非負整数とする。$(x,s)=(0,0)$ の近傍 $U\subset\mathbb{C}\times\mathbb{R}$ で定義された複素数値 C^n 級関数 $f(x,s)$ が条件

$$\frac{\partial^{k+l}f}{\partial x^k\partial s^l}(0,0)=0 \quad (\forall(k,l)\text{ such as }k+l<n)$$
$$\frac{\partial^n f}{\partial x^n}(0,0)\neq 0$$

をみたすとし[13]、x の n 次多項式

$$\sum_{k=0}^{n}\frac{1}{k!(n-k)!}\frac{\partial^n f}{\partial x^k\partial s^{n-k}}(0,0)x^k$$

は重根を持たないとする。

このとき $s=0$ の近傍で定義された C^1 級関数 $\alpha_1(s),\dots,\alpha_n(s)$ で

$$f(\alpha_i(s),s)=0$$
$$\sum_{k=0}^{n}\frac{1}{k!(n-k)!}\frac{\partial^n f}{\partial x^k\partial s^{n-k}}(0,0)x^k s^{n-k}=\frac{\partial^n f}{\partial x^n}(0,0)\prod_{i=1}^{n}(x-\alpha_i'(0)s)$$

をみたすものが存在する。

証明：

$$h(x,s)=\sum_{k=0}^{n}\frac{1}{k!(n-k)!}\frac{\partial^n f}{\partial x^k\partial s^{n-k}}(0,0)x^k s^{n-k}$$

とおき、a を x の多項式 $h(x,1)$ の根の一つ、$a_2,\dots,a_n$ をその他の根とする。

$s=0$ の近傍において、“直線”$x=at$ の近くに $f(x,s)=0$ となる点が一意に存在することをまず示そう。

存在性は正確には次の主張である。

$$\forall\varepsilon>0\exists\delta>0\forall s\in B_\delta\exists x\in\mathbb{C}(|x-as|<|s|\varepsilon\text{ かつ }f(x,s)=0)$$

ただし $B_\delta=\{s\in\mathbb{R}\mid |s|<\delta\}$ である。

ε が小さいほどこの条件はきつくなるので、必要なら後で小さい値に取り直しても良い。

s が十分小さければ曲線 $C=\{x\in\mathbb{C}\mid |x-as|=|s|\varepsilon\}$ に対して

$$\int_C\frac{1}{f(x,s)}dx\neq 0$$

となることを示そう。すると $f(x,s)$ の零点が $|x-as|<|s|\varepsilon$ の範囲に存在することになる。

[13] 複素変数と実変数が混ざった関数に対して C^n 級とか $\frac{\partial}{\partial x},\frac{\partial}{\partial s}$ のような表記をするのは見慣れないがないが、意図は伝わると思うのであえて説明しない。

$g(x,s) = f(x,s) - h(x,s)$ とおくと

$$\begin{aligned}\frac{1}{f(x,s)} =& \frac{1}{(x-as)\prod_{i=2}^n (x-a_i s) + g(x,s)} \\ =& \frac{1}{x-as}\frac{1}{\prod_{i=2}^n (x-a_i s) + \frac{g(x,s)}{x-as}}\end{aligned}$$

となる。

C 上では $x = \left(a + \varepsilon e^{i\theta}\right) s$ $(\theta \in \mathbb{R})$ と書けるので

$$\prod_{i=2}^n (x - a_i s) = s^{n-1}\prod_{i=2}^n (a - a_i + \varepsilon e^{i\theta})$$

となる。これは ε を小さく取り直せば 0 にならず、かつ偏角もほぼ一定となる。加えて s を 0 に近づけたときの 0 への収束の速さは s^{n-1} 程度である。

$g(x,s)$ は n 階までの導関数が全て $(x,s) = (0,0)$ で 0 となることから

$$\frac{g(x,s)}{(|x|+|s|)^n} \to 0 \quad (x, s \to 0)$$

となる。このことに注意すると、C 上で

$$\begin{aligned}\left|\frac{1}{s^{n-1}}\frac{g(x,s)}{x-as}\right| =& (|a+\varepsilon e^{i\theta}| + 1)^n \frac{|g(x,s)|}{(|x|+|s|)^n}\frac{|s|}{|x-as|} \\ \leq& \frac{|g(x,s)|}{(|x|+|s|)^n}\frac{(|a|+2)^n}{\varepsilon} \to 0 \quad (s \to 0)\end{aligned}$$

となるから、ε としてある値より小さい任意の正数を固定したとき、適当に δ を定めれば任意の $s \in B_\delta$ に対して

$$\frac{1}{\prod_{i=2}^n (x-a_i s) + \frac{g(x,s)}{x-as}} = \frac{1}{s^{n-1}}\frac{1}{\prod_{i=2}^n (a - a_i + \varepsilon e^{i\theta}) + o(1)}$$

は C 上 0 にならず、かつその偏角もほぼ一定となる。故に

$$\int_C \frac{1}{f(x,s)} dx \neq 0$$

となる。

次に x の一意性を示そう。正確に書けば

$$\exists\varepsilon > 0 \exists\delta > 0 \forall s \in B_\delta \forall x, y \in \mathbb{C}\left[(|x-as|, |y-as| < |s|\varepsilon \text{ かつ } f(x,s) = f(y,s) = 0) \Rightarrow x = y\right]$$

である。

$$\frac{\partial f}{\partial x}(x,s) = \prod_{i=2}^n (x - a_i s) + (x - as)\sum_{i=2}^n \prod_{j\neq i}(x - a_j s) + \frac{\partial g}{\partial x}(x,s)$$

であるが $|x-as| \leq |s|\varepsilon$ の範囲において、この第 1 項は $s^{n-1}\prod_{i=2}^n (a-a_i)$ に近い値をとり、第 2 項は ε を小さくとればいくらでも 0 に近い値しかとらなくなり、第 3 項は (x,s) が $(0,0)$ に近づけば 0 に近づく。故に ε, δ を適当に小さく取り直すと $|x-as| \leq |s|\varepsilon, |s| < \delta$ において $\frac{\partial f}{\partial x}$ の偏角はいくらでも狭い範囲に収めることができる。

このことから $f(x,s) = 0$ となる x は s 毎に 2 つは存在しないことがしたがう。

以下 s にこの x を対応させる写像を $a(s)$ とおく。

$a(s)$ の微分可能性および微分係数を求めよう。まずどんな $\varepsilon > 0$ に対しても十分 0 に近い s に対しては

$$\left|\frac{a(s)}{s} - a\right| = \frac{|a(s) - as|}{|s|} < \varepsilon$$

が成り立つことから $a'(0) = a$ であることがわかる。

次に $s \neq 0$ における微分について考える。まず

$$\begin{aligned}\frac{\partial f}{\partial x}(a(s), s) =& \prod_{i=2}^{n}(a(s) - a_i s) + (a(s) - as)\sum_{i=2}^{n}\prod_{j\neq i}(a(s) - a_j s) + \frac{\partial g}{\partial x}(a(s), s)\\ =& s^{n-1}\prod_{i=2}^{n}(a - a_i) + o(s)\\ \neq& 0\end{aligned}$$

であるから、$(a(s), s)$ を原点にずらせて $n = 1$ の場合に上記の結果を適用すると $a(s)$ は微分可能、特に連続であることがわかる。

$$\begin{aligned}0 =& f(a(s+h), s+h) - f(a(s), s)\\ =& \frac{f(a(s+h), s+h) - f(a(s), s+h)}{a(s+h) - a(s)}\frac{a(s+h) - a(s)}{h}h + \frac{f(a(s), s+h) - f(a(s), s)}{h}h\end{aligned}$$

であるから

$$\begin{aligned}a'(s) &= \lim_{h\to 0}\frac{a(s+h) - a(s)}{h}\\ &= \lim_{h\to 0}\frac{\frac{f(a(s),s+h) - f(a(s),s)}{h}}{\frac{f(a(s+h),s+h) - f(a(s),s+h)}{a(s+h) - a(s)}}\\ &= -\frac{\frac{\partial f}{\partial s}(a(s), s)}{\frac{\partial f}{\partial x}(a(s), s)}\end{aligned}$$

が成り立つ [14]。最後の等号は $a(s)$ の連続性および、既に見た性質「δ, ε が十分に小さければ $0 < |s| < \delta$, $|x - as| < |s|\varepsilon$ の範囲において

$$\frac{\partial f}{\partial x} \fallingdotseq \prod_{i=2}^{n}(as - a_i s) \neq 0$$

となること」を用いた。

f は C^1 級であるから $a(s)$ は微分可能で、かつ $a'(s)$ は連続であることがわかった。

最後に $\lim_{s\to 0} a'(s) = a(0)$ を示せば全ての証明が完了する。

既に見たように δ を小さくとれば $s \in B_\delta$ に対して

$$\frac{\partial f}{\partial x} \fallingdotseq \prod_{i=2}^{n}(as - a_i s)$$

と近似できる。同様に

$$\begin{aligned}\left.\frac{\partial f}{\partial s}(x, s)\right|_{(x,s)=(a(s),s)} &= \left[-a\prod_{i=2}^{n}(x - a_i s) - (x - as)\sum_{i=2}^{n}a_i\prod_{j\neq i}(x - a_j s) + \frac{\partial g}{\partial s}\right]_{(x,s)=(a(s),s)}\\ &\fallingdotseq -a\prod_{i=2}^{n}(as - a_i s)\end{aligned}$$

[14] 細かいことだが $a(s+h) = a(s)$ となる h に対しては $f(a(s), s+h) - f(a(s), s) = 0$ であるから適当に場合分けして $\varepsilon - \delta$ 式の定義に照らし合わせると $a'(s) = \frac{\partial f}{\partial s}(a(s), s) = 0$ が確認できる。

と近似できるので、

$$a'(s) = -\frac{\frac{\partial f}{\partial s}}{\frac{\partial f}{\partial x}} \to a \quad (s \to 0 \text{ のとき})$$

参考文献

[1] 舟木直久，確率論，朝倉書店，2004（ISBN978-4-254-11600-7）

[2] James Wei, J.C.W.Kuo, Lumping Analysis in Monomolecular Reaction Systems. Analysis of the Exactly Lumpable System, Ind. Eng. Chem. Fundamen., 1969, 8 (1), pp 114-123 (DOI: 10.1021/i160029a019)

公開	初版	2017 年 5 月 12 日	公開
	第 2 版	2017 年 9 月 24 日	公開
著者	河下希		
公開場所	邪道数学研究所		
	http://cubicsphere.web.fc2.com/index.html		