マルコフ連鎖入門　第3回

永幡幸生
新潟大学
nagahata@eng.niigata-u.ac.jp

2021 第1学期

Markov 連鎖

以下に挙げる Perron-Frobenius の定理は確率行列との相性が良い。

定理 (Perron-Frobenius)

A を非負の正方行列として既約とする。このとき以下が成り立つ。
(1) A はスペクトル半径 (固有値の絶対値の最大値) $\rho(A)$ と等しい固有値を持つ。
(2) $\rho(A)$ に対する正の固有ベクトル u が存在する。
(3) $\rho(A)$ は A の全ての成分に対して単調増大。
(4) $\rho(A)$ は単純特性根。

注意：Perron-Frobenius の定理はいくつかの書き方があるようである。基本的には [V] の教科書に従っている。

証明

定理の証明のために用語、補題等を準備する。

x を非負のベクトルとして

$$r_x := \min\{\frac{(Ax)_i}{x_i}; x_i > 0\}$$

と定義する。さらに

$$r := \sup\{r_x; x \geq 0, x \neq 0\}$$

と定義する。
注意：r の定義で全ての非負の x に関する上限を取ったが実際は $\|x\| = 1$ だけに制限しても構わない。
また

$$Az \geq rz$$

を満たす非負のベクトル z を極ベクトル (extremal vector) と呼ぶ。

補題 (極ベクトル)

A を非負の正方行列として既約とする。このとき以下が成り立つ。
(1) $r > 0$
(2) 極ベクトル z は正で A の r に対する固有ベクトル。すなわち
$Az = rz$

証明
(1) 正のベクトル x に対して Ax も正のベクトルになることから明らかに $r_x > 0$ になっているので定義より $r > 0$
(2) まず既約なので十分大きく n をとれば $(I+A)^n$ は正の正方行列になる。
極ベルトルの定義から $y = Az - rz \geq 0$ になるが $y \neq 0$ とすると $w = (I+A)^n z > 0$ に対して

$$Aw - rw = A(I+A)^n z - r(I+A)^n z = (I+A)^n y > 0$$

このとき

$$r_w = \min_i \{\frac{Aw_i}{w_i}\} = \min_i \{\frac{rw_i + ((I+A)^n y)_i}{w_i}\} > r$$

より r の定義に矛盾する。すなわち $Az = rz$
この下で $w = (I+A)^n z = (1+r)^n z > 0$ なので z は正。 □

補題 (行列の比較)

A を非負の正方行列として既約とする。B を複素正方行列とし $|B| \leq A$ とする。
β を B の固有値とすると

$$|\beta| \leq r$$

さらに等号が成立する必要十分条件は $|B| = A$ で $\beta = re^{i\phi}$ と書けたとき対角行列 D でその対角成分 d_i は $|d_i| = 1$ を満たすものを用いて

$$B = e^{i\phi} DAD^{-1}$$

と書ける。

証明

$y \neq 0$ として $By = \beta y$ とすると各成分を比較することで

$$|\beta||y| \leq |B||y| \leq A|y|$$

すなわち $|\beta| \leq r_{|y|} \leq r$ である。
$|\beta| = r$ を仮定すると $|\beta||y| \leq A|y|$ なので $|y|$ は極ベクトル。先の補題より $|y|$ は A の固有値 r に対応する固有ベクトル。すなわち $r|y| = A|y|$。上の式と比較すると

$$r|y| = |B||y| = A|y|$$

でなければならない。また先の補題より $|y|$ は正のベクトルなので $|B| \leq A$ とこの式から $|B| = A$ でなければいけない。

$|y| > 0$ となるベクトル y に対して

$$D = \begin{pmatrix} y_1/|y_1| & & \\ & \ddots & \\ & & y_n/|y_n| \end{pmatrix}$$

は補題の条件を満たしていて $y = D|y|$ と書ける。
$\beta = re^{i\phi}$ として $By = \beta y$ をこれらを使って書き換えると

$$BD|y| = By = \beta y = e^{i\phi} rD|y|,\ \text{すなわち}\ e^{-i\phi} D^{-1} BD|y| = r|y|$$

と書ける。$|e^{-i\phi}D^{-1}BD| = |B|$ と前ページで証明したことより
$e^{-i\phi}D^{-1}BD = |B| = A$ でなければならない。つまり
$B = e^{i\phi}DAD^{-1}$ である。 $\square$

Markov 連鎖

補題 (行列の比較) で $B = A$ とすることで

系 (行列の比較)

A を非負の正方行列として既約とする。このとき r はスペクトル半径にもなる。

正方行列から j 行と j 列を取り除いてできた行列を B とする。(取り除く行、列は 2 つ以上でもよい。) このような行列を主部分行列 (principal submatrix) と呼ぶ。
行の交換に対応する行列 P を使えば PAP^{-1} で取り除きたい行、列を下 (右) から順に並べ直すことが可能なので B は A のブロック行列の左上の部分と思ってよい。

補題 (スペクトル半径の比較)

A を非負の正方行列として既約とする。このとき A の主部分行列 B に対して $\rho(A) > \rho(B)$

証明
前ページに書いたように行の交換に対応する行列 P を用いて

$$A' = PAP^{-1} = \begin{pmatrix} B & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix}$$

と変換して $\rho(A')$ と $\rho(B)$ を比較をしているものと思い直してよい。

$$C = \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix}$$

と定義すると明らかに $\rho(B) = \rho(C)$ でかつ $0 \leq C \leq A'$, $C \neq A$ が成り立つ。従って補題 (行列の比較) が適用でき
$\rho(B) < \rho(A)$ □

Markov 連鎖

問題

各成分が微分可能な関数で与えられる行列 $A(t)$ を $A(t) = (a_1(t), a_2(t), \ldots, a_n(t))$ の形で書く。このとき

$$\frac{d}{dt} det A(t) = \sum_i \det(a_1(t), \ldots, \frac{d}{dt} a_i(t), \ldots, a_n(t))$$

を示せ、但し行列 (ベクトル) の微分は各成分の微分とする。
特に B を定数行列として $A(t) = tI - B$ を考えることで

$$\frac{d}{dt} \det(tI - B) = \sum_i \det(tI - B_i)$$

を示せ、但し B_i は B の i 行、列を取り除いた主部分行列とする。

定理の証明に戻る。
(1) A はスペクトル半径 (固有値の絶対値の最大値) $\rho(A)$ と等しい固有値を持つ。
(2) $\rho(A)$ に対する正の固有ベクトル u が存在する。
この 2 つは補題 (極ベクトル) と系 (行列の比較) から出る。
(3) $\rho(A)$ は A の全ての成分に対して単調増大。
は補題 (行列の比較) から出る。
(4) $\rho(A)$ は単純特性根。
は補題 (スペクトル半径の比較) より A_i を A の i 行、列を取り除いた主部分行列とするとき $\rho(A_i) < \rho(A)$ であるので
$\det(\rho(A)I - A_i) > 0 \ (\forall i)$ である。問題とあわせることにより

$$\frac{d}{dt}\det(tI - A)|_{t=\rho(A)} = \sum_i \det(tI - A_i)|_{t=\rho(A)} > 0$$

より特性方程式において $\rho(A)$ は単純特性根であることが分かる。 $\square$

Markov 連鎖

命題

A を非負の正方行列として既約として正のベクトル u が A の固有ベクトルであるとする。すなわち $Au = \lambda u$ とすると $\lambda = \rho(A)$ である。

注意：[Sai] の教科書の中ではこの命題は Perron-Frobenius の定理の一部になっている。確率論的にはこの事実から $\rho(P) = 1$ がすぐに分かる。(ブロック行列に分けて考える毎に $\rho(P_i) = 1$ が分かる。)

証明

${}^tA$ に関しても Perron-Frobenius の定理は適用可能なので正のベクトル v があって ${}^tvA = \rho(A){}^tv$ を満たす。

$$\rho(A){}^tvu = ({}^tvA)u = {}^tv(Au) = \lambda{}^tvu$$

を満たしていて $u, v > 0$ であるから $\rho(A) = \lambda$ になる。 □

Markov 連鎖

命題

A を正の正方行列とする。このとき A の固有値 λ のうち $\lambda \neq \rho(A)$ となるものは $|\lambda| < \rho(A)$ である。

注意：[Sai] の教科書の中ではこの命題は Perron-Frobenius の定理の一部になっている。

証明

${}^tA$ に関しても Perron-Frobenius の定理は適用可能なので正のベクトル v があって ${}^tvA = \rho(A){}^tv$ を満たす。一方 A の固有値 $\lambda(\neq \rho(A))$ とその固有ベクトル x に対して

$$\lambda x_i = \sum_i a_{i,j}x_j, \qquad |\lambda||x_i| = |\lambda x_i| = |\sum_i a_{i,j}x_j| \leq \sum_i a_{i,j}|x_j|$$

が成り立つ。

Markov 連鎖

2 つめの不等式の両辺に $v_i > 0$ をかけて i に関して和をとると v が $\rho(A)$ に対応する固有ベクトルだったので

$$|\lambda| \sum_i v_i |x_i| \leq \sum_{i,j} v_i a_{i,j} |x_j| = \rho(A) \sum_j v_j |x_j|$$

もし $|\lambda| = \rho(A)$ が成り立つのであればこの不等式でも、1 つ前の不等式でも等号が成立している。すなわち

$$|\sum_i a_{i,j} x_j| = \sum_i a_{i,j} |x_j|$$

が成り立つ。A の全ての成分が正なのでベクトル x はは定数倍 (複素数) することで正のベクトルになる。
先の命題より $|\lambda| = \rho(A)$ であれば $\lambda = \rho(A)$ になる。 □

系

A を非負の正方行列とし、$\exists k$ s.t. A^k は正の正方行列とする。このとき A の固有値 λ のうち $\lambda \neq \rho(A)$ となるものは $|\lambda| < \rho(A)$ である。

注意：実際周期 (d とする) のある確率行列 P に対しては $\rho(P) = 1$ となるが $\omega = \exp(i2\pi/d)$ として $\omega, \omega^2, \ldots, \omega^{d-1}$ が出てくる。

問題

A の固有値を用いて A^k の固有値を表すことで系を証明せよ。

命題 (確率行列の極限)

P を確率行列とし既約、周期 1 (非周期) とする。このとき $\lim_{n\to\infty} P^n$ が存在し

$$\lim_{n\to\infty} P^n = \begin{pmatrix} \mu \\ \vdots \\ \mu \end{pmatrix}$$

と書け、$\mu > 0$ は唯一の定常分布になる。

証明

確率行列なので $u = {}^t(1, \ldots, 1)$ とすれば

$$Pu = u$$

を満たしている。Perron-Frobenius の定理を適用することで $\rho(P) = 1$ であることが分かる。

Markov 連鎖

一方既約で周期 1 なので十分大きな k をとれば P^k は正の正方行列になる。従って命題より 1 以外の固有値 λ は $|\lambda| < \rho(P) = 1$ を満たしている。
また当然

$$vP = v$$

を満たすベクトル v も存在するがこちらも Perron-Frobenius の定理から $v > 0$ となる。これを規格化して

$$\mu = \frac{1}{\sum_i v_i} v$$

とおくことで唯一の定常分布を得られる。(当然 $\mu > 0$)

Markov 連鎖

$Q^{-1}PQ = \Lambda$ と対角化 (必要ならばジョルダン標準形に) するが $\lambda_1 = 1$ とする。このとき $Q = (q_1, q_2, \ldots, q_n)$, $Q^{-1} = {}^t(q_1', q_2', \ldots, q_n')$ と書いたとき q_1, q_1' は固有値 1 に対する固有ベクトルであり ${}^t q_1' q_1 = 1$ を満たしているので特に

$$q_1 = {}^t(1, 1, \ldots, 1), \quad {}^t q_1' = (\mu_1, \mu_2, \ldots, \mu_n) = \mu$$

と取ってよい。ここで 1 以外の固有値 λ が $|\lambda| < 1$ であることから

$$\begin{aligned} \lim_{n\to\infty} P^n &= \lim_{n\to\infty} Q\Lambda^n Q^{-1} \\ &= Q \begin{pmatrix} 1 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{pmatrix} Q^{-1} = \begin{pmatrix} 1 & 0\cdots0 \\ 1 & 0\cdots0 \\ \vdots & \vdots \\ 1 & 0\cdots0 \end{pmatrix} Q^{-1} = \begin{pmatrix} \mu \\ \vdots \\ \mu \end{pmatrix} \end{aligned}$$

□