Chứng minh sơ cấp cho định lý Perron – Frobenius

Trương Phước Nhân, 25/01/2019

Nội dung của bài viết trình bày một chứng minh sơ cấp cho định lý Perron - Frobenius (xem thêm bài viết “**Định lý Perron – Frobenius**” để nắm cách giải bài toán này dựa trên nguyên lý ánh xạ co của Banach)

Đầu tiên ta cần nhắc lại nội dung cơ bản của định lý Perron – Frobenius:

Bài toán. (Perron – Frobenius)

Giả sử cho trước ma trận A cấp $n\times n$ sao cho tất cả các phần tử đều là các số thực không âm.

Khi đó:

(1) Ma trận A có ít nhất một giá trị riêng thực không âm, ta tạm kí hiệu giá trị riêng này là $\lambda(A)$. Giá trị riêng $\lambda(A)$ này có giá trị lớn hơn hoặc bằng trị tuyệt đối của tất cả các giá trị riêng còn lại của A. Nếu tất cả các phần tử của ma trận A đều là các số thực dương thì đánh giá vừa nêu ở trên là nghiêm ngặt.

(2) Nếu ma trận A có tất cả các phần tử đều là các số thực dương thì số bội của giá trị riêng $\lambda(A)$ bằng 1 và ta luôn có thể chuẩn hóa vector riêng tương ứng sao cho tất cả các thành phần đều là các số thực dương.

(3) Nếu ma trận A có một vector riêng v có tất cả các thành phần đều là các số thực dương thì giá trị riêng tương ứng của vector v chính là $\lambda(A)$.

Đầu tiên ta sẽ kiểm tra lại các khẳng định vừa nêu cho trường hợp đơn giản $n=2$ và cố gắng mở rộng các kết quả và phương pháp phân tích cho trường hợp tổng quát.

I. Trường hợp $n=2$

Xét ma trận $A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, a,b,c,d là các số thực không âm.

Đa thức đặc trưng của ma trận A có dạng $P_A(t)=\det(tI-A)=t^2-(a+d)t+(ad-bc)$.

(1) Do biệt thức tương ứng của $P_A(t)$ bằng $\Delta=(a-d)^2+4bc\ge 0$ nên

Bằng cách giải phương trình $P_A(t)=0$, ta tính ra được các giá trị riêng của ma trận A có dạng

$$\lambda_1=\frac{(a+d)+\sqrt{(a-d)^2+4bc}}{2} \text{ và } \lambda_2=\frac{(a+d)-\sqrt{(a-d)^2+4bc}}{2}.$$

Dễ dàng nhận thấy rằng $\lambda_1\ge 0$ và $\lambda_1\ge|\lambda_2|$, $\frac{a+d}{2}$ là một số thực không âm.

Nếu ma trận A có tất cả các phần tử đều là các số thực dương thì $\frac{a+d}{2}$ cũng sẽ là một số thực dương nên $\lambda_1>|\lambda_2|$.

(2) Nếu ma trận A có tất cả các phần tử đều là các số thực dương thì biệt thức $\Delta>0$, nên đa thức đặc trưng $P_A(t)$ có hai nghiệm thực phân biệt. Do đó, giá trị riêng λ_1 có số bội bằng 1 và $\lambda_1>|\lambda_2|\ge 0$.

Để thuận tiện cho việc trình bày phép chứng minh ta sử dụng kí hiệu λ thay cho kí hiệu λ_1.

Giả sử $x=\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ là vector riêng tương ứng với giá trị riêng λ.

Khi đó,

$$Ax=\lambda x$$
$$\Rightarrow \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} \lambda x_1 \\ \lambda x_2 \end{pmatrix}$$
$$\Rightarrow \begin{cases} ax_1+bx_2=\lambda x_1 \\ cx_1+dx_2=\lambda x_2 \end{cases}$$

Từ cách định nghĩa ta suy ra rằng $x_1\neq 0$ hoặc $x_2\neq 0$, ta tạm giả sử $x_1\neq 0$.

Khi đó,

$$\begin{cases} a+b\dfrac{x_2}{x_1}=\lambda \\ c+d\dfrac{x_2}{x_1}=\lambda\dfrac{x_2}{x_1} \end{cases}$$
$$\Rightarrow \begin{cases} \dfrac{x_2}{x_1}=\dfrac{\lambda-a}{b} \\ (\lambda-d)\dfrac{x_2}{x_1}=c>0 \end{cases}.$$

Nhận thấy rằng để chứng minh yêu cầu của bài toán ta cần chỉ ra rằng $\dfrac{x_2}{x_1}>0$.

Thật vậy, từ các phân tích vừa thu được ta chỉ cần chỉ ra rằng $\lambda>a$ hoặc $\lambda>d$ nhưng điều này thì lại là hiển nhiên do $\lambda>\dfrac{a+d}{2}$.

(3) Giả sử ma trận A có một vector riêng $v=\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ có tất cả các thành phần đều là các số thực dương với λ_v là giá trị riêng tương ứng.

Khi đó,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}=\begin{pmatrix} \lambda_v v_1 \\ \lambda_v v_2 \end{pmatrix}$$

Bằng phương pháp lập luận tương tự, ta thu được

$$\begin{cases} b\dfrac{v_2}{v_1}=\lambda_v-a \\ (\lambda_v-d)\dfrac{v_2}{v_1}=c\geq 0 \end{cases}.$$

Nhận xét rằng do $\dfrac{v_2}{v_1}>0$ nên $\lambda_v\geq a$ và $\lambda_v\geq d$.

Do đó $\lambda_v\geq\dfrac{a+d}{2}$ nên $\lambda_v=\dfrac{(a+d)+\sqrt{(a-d)^2+4bc}}{2}=\lambda_1$.

Điều này kết thúc phép chứng minh của định lý Perron – Frobenius cho trường hợp $n=2$.

II. Trường hợp tổng quát

Nhằm mục đích giảm thiểu sự trùng lặp trong các lập luận và phân tích đã gặp phải khi khảo sát trường hợp cơ bản khi $n=2$ ta tiến hành phép chứng minh như sau:

Chứng minh (3)

Giả sử ma trận A có một vector riêng v có tất cả các thành phần đều là các số thực dương và λ_v là giá trị riêng tương ứng, tức là $Av=\lambda_v v$.

Nhận xét rằng $v=\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}=\begin{pmatrix} v_1 & 0 & \cdots & 0 \\ 0 & v_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & v_n \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}=C\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$, nên

$$AC\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}=\lambda_v C\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

$$\Rightarrow C^{-1}AC\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}=\lambda_v\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \text{ trong đó } C^{-1}=\begin{pmatrix} v_1^{-1} & 0 & \cdots & 0 \\ 0 & v_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & v_n^{-1} \end{pmatrix}.$$

Nhận thấy rằng ma trận C có tất cả các phần tử đều là các số thực không âm nên ma trận $C^{-1}AC$ cũng có tất cả các phần tử đều là các số thực không âm. Do hai ma trận A và $C^{-1}AC$ đồng dạng nên chúng có cùng tập phổ.

Không giảm tính tổng quát của bài toán ta có thể giả sử rằng $v=\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$.

Khi đó, $\lambda_v=\sum_{j=1}^{n} a_{ij}\,, 1\le i\le n$.

Do đó λ_v là số thực không âm và λ_v là một số thực dương ngoại trừ trường hợp $A=0$.

Trước khi tiếp tục chứng minh phần còn lại của (3) ta trang bị cấu trúc chuẩn $\|.\|_\infty$ cho không gian $\mathbb{C}^n$ được xác định bởi hệ thức $\|z\|_\infty := \max_{i=1,\dots,n}|z_i|$ với $z=\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$.

Nhận thấy rằng, với mỗi $z\in\mathbb{C}^n$, thành phần thứ i của vector Az bằng $a_{i1}z_1+a_{i2}z_2+\dots+a_{in}z_n$.

Khi đó,

$$\begin{aligned} |a_{i1}z_1+a_{i2}z_2+\dots+a_{in}z_n| &\le |a_{i1}||z_1|+\dots+|a_{in}||z_n| \\ &\le \left(\sum_{j=1}^{n} b_{ij}\right)\max_{i=1,\dots,n}|z_i| \end{aligned}$$

$$= \lambda_v \|z\|_\infty .$$

Do đó,

$$\|Az\|_\infty \le \lambda_v \|z\|_\infty .$$

Giả sử z' là một giá trị riêng bất kì của ma trận A và λ' là giá trị riêng tương ứng.

Khi đó,

$$\|Az'\|_\infty = |\lambda'|.\|z'\|_\infty \le \lambda_v \|z'\|_\infty .$$

Do đó, $\lambda_v \ge |\lambda'|$.

Nhận xét: Nếu tất cả các phần tử của ma trận A đều là các số thực dương thì $\|Az\|_\infty < \lambda_v \|z\|_\infty$ trừ trường hợp $z_1 = z_2 = ... = z_n$, hay nói cách khác số bội của giá trị riêng λ_v bằng 1.

Chứng minh (1), (2).

Ý tưởng chứng minh nằm ở chỗ ta sẽ "xấp xỉ" ma trận A bởi các ma trận có tất cả các phần tử đều là các số thực dương như sau:

Giả sử A là ma trận có tất cả các phần tử là các số thực không âm.

Đặt A_r là ma trận thu được từ ma trận A bằng cách thay các phần tử 0 bằng phần tử $\frac{1}{r}$, trong đó $r > 0$.

Ta tiến hành chứng minh bài toán theo ba bước sau:

Bước i. Ta chứng minh rằng các giá trị riêng của ma trận A_r tiến dần đến các giá trị riêng của ma trận A khi $r \to \infty$.

Đầu tiên ta cần đến kết quả phụ sau trong quá trình lập luận:

Bổ đề 1.

Cho trước đa thức $f(z) = z^n + a_{n-1}z^{n-1} + ... + a_0$, $a_i \in \mathbb{C}$, $|a_0|,...,|a_{n-1}| < M$.

Khi đó, nếu $z_0 \in \mathbb{C}$ là một nghiệm nào đó của đa thức f thì $|z_0| < 1 + nM$.

Chứng minh bổ đề 1.

Ta sẽ chứng minh khẳng định của bài toán bằng phương pháp phản chứng, giả sử rằng z_0 là một nghiệm nào đó của đa thức f sao cho $|z_0| < 1 + nM$.

Khi đó, $f(z_0) = 0$, nên $z_0^n = -a_{n-1}z_0^{n-1} - a_{n-2}z_0^{n-2} - ... - a_0$.

Bằng cách lấy môđun cả hai vế của biểu thức vừa nêu ở trên ta nhận thấy rằng

$$\begin{aligned} |z_0^n| &= |-a_{n-1}z_0^{n-1} - a_{n-2}z_0^{n-2} - ... - a_0| \\ &= |-a_{n-1}z_0^{n-1} - a_{n-2}z_0^{n-2} + ... + a_0| \end{aligned}$$

Thực hiện phép chia cho x^n, ta thu được

$$\begin{aligned} 1 &= \left|a_{n-1}z_0^{n-1}/z_0^n + a_{n-2}\, z_0^{n-2}/z_0^n + ... + a_0/z_0^n\right| \\ &\le \left|\frac{a_{n-1}}{z_0}\right| + \left|\frac{a_{n-2}}{z_0^2}\right| + ... + \left|\frac{a_0}{z_0^n}\right| \\ &\le \frac{M}{|z_0|} + \frac{M}{|z_0^2|} + ... + \frac{M}{|z_0^n|} \\ &\le \frac{M}{(1+nM)} + \frac{M}{(1+nM)^2} + ... + \frac{M}{(1+nM)^n} \\ &\le n\frac{M}{(1+nM)} \\ &< 1. \end{aligned}$$

Điều mâu thuẫn này kết thúc phép chứng minh bổ đề 1.

Giả sử các trị riêng của toán tử A_r là $\lambda_1^{(r)},...,\lambda_n^{(r)}$.

Từ kết quả nhận được từ bổ đề 1) ta nhận thấy rằng $\{(\lambda_1^{(r)},...,\lambda_n^{(r)})\}_{r=1}^{\infty}$ là một dãy bị chặn.

Bằng cách áp dụng nguyên lý Bolzano – Weierstrass ta suy ra dãy này chứa ít nhất một dãy con hội tụ, tạm giả sử là dãy con $\{(\lambda_1^{(r_j)},...,\lambda_n^{(r_j)})\}_{j=1}^{\infty}$, nên $A_{r_j} \to A$.

Do đó, $P_{A_{r_j}} \to P_A$ khi $j \to \infty$, trong đó P_A là đa thức đặc trưng của ma trận A.

Đặt $\lambda_k := \lim\limits_{j\to\infty} \lambda_k^{(r_j)}, k = 1,...,n$.

Khi đó, $P_A(t) = \prod\limits_{k=1}^{n}(t - \lambda_k)$, do $P_{A_{r_j}}(t) = \prod\limits_{k=1}^{n}\left(t - \lambda_k^{(r_j)}\right)$.

Điều này chỉ ra rằng các giá trị riêng của ma trận A_{r_j} tiến dần đến các giá trị riêng của ma trận A khi $j \to \infty$.

Bước ii. Ta sẽ chứng minh tồn tại một giá trị riêng của ma trận A_r có tất cả các thành phần đều là các số thực dương.

Đầu tiên ta sẽ chứng minh khẳng định sau:

Bổ đề 2.

Cho trước ma trận A có tất cả các phần tử đều là các số thực dương thỏa mãn $Av' = \lambda_{v'} v'$ trong đó các thành phần của vector v' đều là các số thực không âm và $v' \neq 0$.

Khi đó, tất cả các thành phần của vector v' đều là các số thực dương.

Chứng minh bổ đề 2.

Do tất cả các thành phần của vector v' đều là các số thực không âm nên điều tương tự cũng sẽ xảy ra đối với vector Av'.

Hơn nữa do tất cả các phần tử của ma trận A đều là các số thực dương nên tất cả các phần tử của vector Av' cũng đều là các số thực dương.

Thật vậy nếu vector Av' có ít nhất một thành phần bằng 0 thì tất cả các thành phần của vector v' đều phải bằng 0, mâu thuẫn với giả thiết $v' \neq 0$.

Điều này kết thúc phép chứng minh bổ đề 2.

Từ bổ đề 2) ta nhận thấy rằng nếu ta chứng minh nếu một ma trận A có một giá trị riêng v' với tất cả các thành phần đều là các số thực không âm thì điều này cũng có nghĩa là giá trị riêng v' sẽ có tất cả các phần tử đều là các số thực dương, trong đó ta tạm giả sử ma trận A có tất cả các phần tử đều là các số thực dương.

Đặt $D := \left\{ d = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix} \middle| 0 \le d_i \le 1, \forall i = 1,...,n \right\}$.

Với mỗi ma trận $A = (a_{ij})_{n\times n}$, ta xét chuẩn $\|A\| = \max\limits_{i=1,...,n} \sum\limits_{j=1}^{n} |a_{ij}|$.

Để chứng minh sự tồn tại của vector v' ta xét hàm $f(d) = \|dAd^{-1}\|$ với mọi $d \in \operatorname{int} D$.

Nhận xét rằng hàm f liên tục trên $\operatorname{int} D$.

Với mọi $\varepsilon > 0$, ta định nghĩa tập D^{ε} như sau:

" $d \in D^{\varepsilon}$ khi và chỉ khi $d_i \ge \varepsilon$ với mọi $i = 1,...,n$ và $\sum_{i=1}^{n} d_i = 1$".

Bổ đề 3.

Với mọi $\varepsilon > 0$ đủ nhỏ, tồn tại phần tử $d'' \in D^{\varepsilon}$ sao cho $f(d') \ge f(d'')$ với mọi $d' \in \operatorname{int} D$.

Chứng minh bổ đề 3.

Nhận thấy rằng $D^{\varepsilon} \subset \operatorname{int} D$ và D^{ε} là một tập đóng, bị chặn nên nó là một tập compact.

Do hàm f liên tục trên $D^{\varepsilon} \subset \operatorname{int} D$ và D^{ε} là một tập compact nên hàm f đạt giá trị cực tiểu trên tập D^{ε}, ta tạm giả sử $f(d'') = \max_{d' \in D^{\varepsilon}} f(d')$.

Giả sử phần tử d' không nằm trong tập D^{ε}, tức tồn tại một chỉ số i sao cho $d'_i < \varepsilon$ và $d'_1 + ... + d'_{i-1} + d'_{i+1} + ... d'_n > 1 - \varepsilon$.

Do đó tồn tại một chỉ số $j \ne i$ sao cho $d'_j > \dfrac{1-\varepsilon}{n-1}$.

Khi đó,

$$f(d') = \left\| d'Ad'^{-1} \right\| \ge d'_j a_{ji} d_i'^{-1} > \frac{(1-\varepsilon)a_{ji}}{(n-1)\varepsilon}.$$

Bằng cách chuyển qua giới hạn khi cho $\varepsilon \to 0$, tức là nếu có một phần tử nào đó của d tiến gần đến 0, ta nhận thấy rằng $f(d) \to \infty$.

Do đó, hàm f đạt được giá trị cực tiểu tại d''.

Điều này kết thúc phép chứng minh bổ đề 3.

Tiếp theo ta kí hiệu lại ma trận $d'' = \begin{pmatrix} d''_1 & 0 & \cdots & 0 \\ 0 & d''_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d''_n \end{pmatrix}$ dưới dạng vector $d'' = (d''_1,...,d''_n)$.

Bổ đề 4. d'' là một giá trị riêng tương ứng của ma trận A với giá trị riêng tương ứng $\lambda = f(d'')$.

Chứng minh bổ đề 4.

Không giảm tính tổng quát của bài toán, bằng cách thay ma trận A bởi ma trận $d''Ad''^{-1}$, ta có thể giả sử rằng $d'' = (1,...,1)$ (xem lại phương pháp lập luận đã sử dụng để chứng minh (3)).

Khi đó, $\max_{i=1,...,n} \sum_{j=1}^{n} a_{ij} = \lambda$.

Từ kết quả nêu trong bổ đề 3),

$$\max_{i=1,...,n} \sum_{j=1}^{n} d_i a_{ij} d_j^{-1} \ge \lambda \text{ với mọi } d_k > 0, \forall k.$$

Đặt $S = \left\{ i \mid \sum_{j=1}^{n} a_{ij} < \lambda \right\}$.

Để hoàn thành chứng minh khẳng định của bài toán ta chỉ cần chứng minh rằng $S = \varnothing$.

Giả sử rằng $S \ne \varnothing$.

Với mọi $\varepsilon > 0$ đủ nhỏ, bằng cách đặt $d_i = \begin{cases} 1-\varepsilon, i \notin S, \\ 1, i \in S, \end{cases}$ ta nhận thấy rằng $\max_{i=1,...,n} \sum_{j=1}^{n} d_i a_{ij} d_j^{-1} < \lambda$.

Điều mâu thuẫn này chỉ ra rằng $S = \varnothing$, đpcm.

Kết hợp các bổ đề 2), 3) và 4) ta hoàn thành phép chứng minh của bước ii).

ta nhận thấy rằng các ma trận A_r luôn có ít nhất một giá trị riêng có tất cả các thành phần đều là các số thực dương, ta tạm giả sử là $\lambda_1^{(r)}$.

Bước iii. Kiểm tra lại các tính chất nêu ở mục (1) và (2).

Bằng cách sử dụng các phân tích thu được khi chứng minh (3) và các bước i), ii), ta nhận thấy rằng:

Với mọi chỉ số $r > 0$, ma trận A_r luôn có một giá trị riêng dương có giá trị lớn hơn hoặc bằng trị tuyệt đối của tất cả các giá trị riêng còn lại của A_r, ta tạm giả sử là $\lambda_1^{(r)}$.

Bằng cách áp dụng lại phương pháp lập luận đã sử dụng ở bước i) ta sẽ chỉ ra được một dãy con $\{A_{r_j}\}_{j=1}^{\infty}$ với dãy các giá trị riêng tương ứng $\{(\lambda_1^{(r_j)},...,\lambda_n^{(r_j)})\}_{j=1}^{\infty}$ hội tụ đến dãy các giá trị riêng của ma trận A ban đầu.

Đặt $\lambda_k = \lim\limits_{j\to\infty} \lambda_k^{(r_j)}, \forall k = 1,...,n$.

Khi đó, $\left(\lambda_1,...,\lambda_n\right)$ là dãy các giá trị riêng tương ứng của toán tử A.

Nhận xét rằng:

(1) $\lambda_1^{(r_j)} > 0$ với mọi $j \Rightarrow \lambda_1$ là một số thực không âm;

(2) Với mọi $2 \le k \le n$, $\lambda_1^{(r_j)} > \left|\lambda_k^{(r_j)}\right|, \forall j \Rightarrow \lambda_1 \ge |\lambda_k|, \forall 2 \le k \le n$.

Tổng hợp các bước i), ii) và ii) ta hoàn tất phép chứng minh cho định lý Perron – Frobenius.

<u>Tài liệu tham khảo:</u>

[1]. Suyeon Khim, The Frobenius – Perron, ??/??/????.

[2]. Trương Phước Nhân, Định lý Perron – Frobenius, 29/08/2017.