Remarks on two theorems in linear algebra *

Yifan Ren†
Zhiyuan College, Shanghai Jiao Tong University
Tong suo Wu‡
School of Mathematical Sciences, Shanghai Jiao Tong University

Abstract. In this note, we use the concept of a polynomial ring to give a direct proof to Cayley-Hamilton Theorem. We also give an elementary proof to Birkhoff theorem on Bi-stochastic matrices.

Key Words: polynomial ring; Cayley-Hamilton Theorem; Bi-stochastic matrices; Birkhoff theorem.

2010 AMS Classification: Primary: 13P05; 13B25; Secondary: 12D05; 05C50.

Recent rediscovery [1] on eigenvalues and eigenvectors aroused interests of many readers, and leaves the hope that many more new discoveries can be found even in the traditional field of studies. In this short note, we record two things happened in the classroom discussions.

1. A new proof of Cayley-Hamilton Theorem

The main result concerning the eigen polynomial $f_1(x) =: \det(xE - A)$ of a square matrix $A$ is the Cayley-Hamilton theorem:

**Theorem 1.1.** Let $A$ be a square matrix over a commutative ring $S$. Then the eigen polynomial $f_1(x) =: \det(xE - A)$ of $A$ annihilates $A$, i.e., $f_1(A) = 0$.

---

*This research was supported by the NSF of Shanghai (Grant No. 19ZR1424100).
†ivan_ren@sjtu.edu.cn
‡Corresponding author. tswu@sjtu.edu.cn
Recall that a column stochastic matrix is a nonnegative matrix, in which the sum of entries of each column is 1. A matrix is called bi-stochastic, if both $A$ and $A^T$ are column \(\square\)

The following is also an immediate consequence of Lemma 1.2:

**Proposition 1.3.** Let $R$ be any ring, and let $r \in R$. Then for any $f(x) \in R[x]$, there exists a polynomial $q(x) \in R[x]$ such that $f(x) = q(x)(x - r) + f(r)$.

Note that $xr = rx$ holds in $R[x]$, thus if we accept $x^n r$ as the standard form of a term in $R[x]$, then there is another version of Proposition 1.3. Thus the generalized Bézout theorem ([2, Page 81]) also follows from Proposition 1.3 (actually, Lemma 1.2).

**Remark.** We hope that this proof could be read by college students. (Actually, the first author is a first-year undergraduate and, the proof is resulted from class discussions.) For this, $S$ could be regarded as a number field $F$, and note that $R[x]$ is exactly the same with the polynomial ring $F[x]$ on the number field $F$, except that $F$ is replaced by a more general ring $R$, e.g., $R = \mathbb{Z}[x], \mathbb{Z}[(\sqrt{2})], M_n(F)$.

Note that the key is Lemma 1.2, thus one needs not to know the concept of modules, and the equality $(M_n(F))[x] = M_n(F[x])$ is essentially not needed to know. What is needed is the concept of a polynomial ring $R[x]$ over a general ring $R$, in which $R = M_n(F)$ is needed in the proof.

**2. Bi-stochastic matrices and Birkhoff theorem**

Recall that a column stochastic matrix is a nonnegative matrix, in which the sum of entries of each column is 1. A matrix is called bi-stochastic, if both $A$ and $A^T$ are column
stochastic.

Recall the following famous theorem of Birkhoff (see e.g., [4, Theorem 8.7.2] or [5, Theorem 5.5.1]):

**Theorem 2.1.** A nonnegative real matrix $M \in M_n(\mathbb{R})$ is bi-stochastic if and only if it is the mass center of some permutation matrices.

In order to prove the theorem, techniques from convex geometry is often applied, and one can refer to [5, Theorem 5.5.1]) or the first edition of [4]. In the second edition of [4], the authors have a noble try to give a more easy and elementary approach to the theorem. While the key is to prove Lemma 8.7.1, by taking advantage of $\det(xE - A)$. In the proof of the lemma, there is an obstacle that we can not overcome. In the following, we provide an alternative way:

**Lemma 2.2.** ([4, Lemma 8.7.1]) Let $A \in M_n(\mathbb{R})$ be a bi-stochastic matrix, which is not a permutation matrix. Then there is a rearrangement $i_1, \ldots, i_n$ of $1, 2, \ldots, n$ such that $a_{k_i k} \neq 0$ holds for all $k$.

*Proof.* If an entry of $A$ is 1, then the result follows by induction assumption. In the following, assume $a_{ij} \neq 1, \forall i, j$.

Let $a_{i_1 j_1} \neq 0$. Then there exists an integer $j_2 \neq j_1$ such that $a_{i_1 j_2} \neq 0$ holds. Clearly, we also have $i_2 \neq i_1$ such that $a_{i_2 j_2} \neq 0$. Again by assumption, there is $j_3 \neq j_2$ such that $a_{i_2 j_3} \neq 0$ holds. If further $j_3 \neq j_1$ holds true, then we have $i_3 \neq i_2$ such that $a_{i_3 j_3} \neq 0$ holds. If further $i_3 \neq i_1$, then we continue the process. If $a_{i_3 j_1} \neq 0$, then we finish to obtain

$$
\begin{pmatrix}
  a_{i_1 j_1} & a_{i_1 j_2} & a_{i_1 j_3} \\
  a_{i_2 j_2} & a_{i_2 j_3} \\
  a_{i_3 j_1} & a_{i_3 j_2} & a_{i_3 j_3}
\end{pmatrix}.
$$

On the other hand, if $a_{i_3 j_1} = 0$, then we continue the previous discussion. Clearly, row indexes (and column index) will repeat after finite steps. Without loss of generality, we assume row indexes will repeat first, and assume further that the first repeat appears in $j_{r+1} = j_1$. This implies that both set $i_1, \ldots, i_r$ and set $j_1, \ldots, j_r$ has cardinality $r$, and

$$a_{i_t j_t} \neq 0, a_{i_t j_{t+1}} \neq 0, \forall t = 1, \ldots, r.$$

Now we are ready to construct a matrix $B$ in the following:

$$b_{i_t j_t} = 1, b_{i_t j_{t+1}} = -1, \forall t = 1, \ldots, r.$$

Now for $c = \min_{1 \leq t \leq r} \{a_{i_t j_t}, a_{i_t j_{t+1}}\}$, the matrices $A + cB$ and $A - cB$ are both bi-stochastic, and at least one of them has zero entries fewer than $A$. Note also that neither has newly added zero entry. This completes the proof by induction. $\square$
References

[1] P.B. Denton, S.J. Parke, T. Tao and X.N Zhang. Eigenvectors from eigenvalues. arXiv:1908.03795v1 [math.RA] 10 Aug 2019.

[2] F.R. Gantmacher. The Theory of Matrices vol.1. AMS Chelsea Publishing, American Mathematical Society Providence, Rhode Island 2000.

[3] Y.Q. Guo, K.P. Shum and Z.P. Wang. A Course on Advanced Algebra. Science Press, 2014 Beijing. (In Chinese)

[4] R.A. Horn and C.R. Johnson. Matrix Analysis. Cambridge University Press. New York. 2’ed Ed. 2013. (1’st Ed. 1985)

[5] D. Serre. Matrices: Theory and Applications. GTM 216, Springer. New York/Berlin/Heidelberg 2002.

[6] S.M. Shi and E.F. Wang. Advanced Algebra. Higher Education Press, Beijing, Third Ed. 1988. (In Chinese)