A binomial-like matrix equation

Alin Bostan, Thierry Combot

A binomial-like matrix equation

Alin Bostan and Thierry Combot

Abstract

We show that a pair of matrices satisfying a certain algebraic identity, reminiscent of the binomial theorem, must have the same characteristic polynomial. This is a generalization of Problem 4 (11th grade) from the Romanian National Mathematical Olympiad 2011.

1 Introduction.

Let $\mathcal{M}_n(\mathbb{C})$ denote the ring of $n \times n$ complex matrices. The following problem was proposed during the 2011 edition of the Romanian National Mathematical Olympiad [1].

Problem 1. *Suppose that $A, B \in \mathcal{M}_2(\mathbb{C})$ satisfy $A^2 + B^2 = 2AB$.*

(i) *Prove that $AB = BA$.*

(ii) *Prove that $\mathrm{Tr}(A) = \mathrm{Tr}(B)$.*

A natural question is whether the assumptions on the size of the matrices, and on the equality they satisfy, are essential to ensure the conclusions. The aim of this short note is to prove the following generalization of (ii).

Theorem 1. *Let n and k be two positive integers, and let $A, B \in \mathcal{M}_n(\mathbb{C})$ satisfy*

$$\sum_{\ell=0}^{k} (-1)^{\ell} \binom{k}{\ell} A^{k-\ell} B^{\ell} = 0. \tag{1}$$

Then, the characteristic polynomials of A and B coincide.

The extension is twofold: the size of the given matrices is arbitrary, and so is also the exponent k in the *binomial-like equation* (1); on the other hand, the conclusion is much more general.

Concerning part (i) of Problem 1, it is not difficult to construct examples of matrices satisfying (1), and which do not commute. In fact, apart from the trivial cases $n = 1$ and $k = 1$, the case $(n, k) = (2, 2)$ is the only one for which equality (1) also implies that $AB = BA$.

To our knowledge, all the previous solutions to Problem 1 first prove that A and B commute, and then use this information to show that their traces are equal. Therefore, they are not directly generalizable to arbitrary sizes n, nor to arbitrary powers k.

2 Equality of spectra, and two special cases.

For a matrix M, let $\sigma(M)$ denote its spectrum, that is, the set of the eigenvalues of M.

We first show that equation (1) entails the equality $\sigma(A) = \sigma(B)$. Let $\lambda \in \mathbb{C}$ be an eigenvalue of B, and let $v \in \mathbb{C}^n$ be an eigenvector associated to λ. The equality $Bv = \lambda v$ yields $B^\ell v = \lambda^\ell v$ for all $\ell \geq 0$, so that (1) implies

$$0 = \sum_{\ell=0}^{k} (-1)^\ell \binom{k}{\ell} A^{k-\ell} B^\ell v = \sum_{\ell=0}^{k} (-1)^\ell \binom{k}{\ell} A^{k-\ell} \lambda^\ell v = (A - \lambda I)^k v,$$

and thus λ is also an eigenvalue of A. This proves $\sigma(B) \subseteq \sigma(A)$.

The opposite inclusion readily follows by taking the transpose of equation (1), and using the fact that eigenvalues are stable under transposition. Indeed, the previous reasoning shows that $\sigma(A) = \sigma(A^t) \subseteq \sigma(B^t) = \sigma(B)$. Therefore, $\sigma(A) = \sigma(B)$.

For matrices A and B of size 2, the equality $\sigma(A) = \sigma(B)$ already implies the equality of the characteristic polynomials $\chi_A(x)$ and $\chi_B(x)$, since either both A and B have the same eigenvalue with multiplicity 2, or they admit the same two distinct eigenvalues. This proves Theorem 1 in the special case $n = 2$.

In size n, a similar argument applies to the *generic* case where one of the characteristic polynomials $\chi_A(x)$ or $\chi_B(x)$ has only simple roots.

3 The general case.

To treat the case when the matrix size n is arbitrary, we will use the notion of *generalized eigenvector*. This will enable us to prove that not only the spectra of A and B are the same as sets (as showed above), but also the algebraic multiplicities of the eigenvalues of A and B are equal.

Let λ be an eigenvalue of the matrix $A \in \mathcal{M}_n(\mathbb{C})$. Recall that the *algebraic multiplicity* of λ is the multiplicity of λ as a root of the characteristic polynomial $\chi_A(x)$. The *generalized eigenspace* of A associated to λ is

$$E_\lambda^\infty(A) = \{v \in \mathbb{C}^n \mid (A - \lambda I)^p v = 0 \text{ for some } p \geq 1\}.$$

The non-zero elements of $E_\lambda^\infty(A)$ are called *generalized eigenvectors* of A associated to the eigenvalue λ.

The following lemma expresses the algebraic multiplicity of any eigenvalue in terms of the dimension of the corresponding generalized eigenspace. This is a classical result, see e.g. [2, Lemma 4.2.4] for an elementary proof. It is also a direct consequence of the Jordan canonical form, but this more advanced notion is not necessary here.

Lemma 1. *Let A be a $n \times n$ matrix with complex coefficients, and let $\lambda \in \mathbb{C}$ be an eigenvalue of A. Then, the algebraic multiplicity of λ is equal to the dimension of the $\mathbb{C}$-vector space $E_\lambda^\infty(A)$.*

Since $\sigma(A) = \sigma(B)$, the characteristic polynomials χ_A and χ_B admit the following factorizations in $\mathbb{C}[x]$, for some positive integers e_i and f_j:

$$\chi_A(x) = (x-\lambda_1)^{e_1}\cdots(x-\lambda_m)^{e_m}, \quad \chi_B(x) = (x-\lambda_1)^{f_1}\cdots(x-\lambda_m)^{f_m}.$$

To conclude the proof of Theorem 1, it is sufficient by Lemma 1 to show that $E_\lambda^\infty(B) \subseteq E_\lambda^\infty(A)$ for any eigenvalue λ. Indeed, this implies that $f_i \le e_i$ for all i, and the equality $n = e_1 + \cdots + e_m = f_1 + \cdots + f_m$ then yields the equality $e_i = f_i$ for all i, and thus $\chi_A(x) = \chi_B(x)$.

Before proceeding to the proof of the needed inclusion $E_\lambda^\infty(B) \subseteq E_\lambda^\infty(A)$ (Lemma 3 below), we first show that equality (1) still holds if A and B are replaced by $A-\lambda I$, and by $B-\lambda I$, respectively. This is done in the following Lemma.

Lemma 2. *Under the assumptions of Theorem 1, the following equality holds for any $\lambda \in \mathbb{C}$,*

$$\sum_{\ell=0}^{k} (-1)^\ell \binom{k}{\ell} (A-\lambda I)^{k-\ell}(B-\lambda I)^\ell = 0. \tag{2}$$

Proof. Let $u, v \in \mathbb{N}$ be such that $u+v \le k$. The coefficient of A^uB^v in the left-hand side of equation (2) is equal to

$$c_{u,v} = (-\lambda)^{k-u-v} \sum_{\ell=v}^{k-u} (-1)^\ell \binom{k}{\ell}\binom{k-\ell}{u}\binom{\ell}{v}.$$

The identity

$$\binom{k}{\ell}\binom{k-\ell}{u}\binom{\ell}{v} = \binom{k}{k-u}\binom{k-u}{v}\binom{k-u-v}{\ell-v}$$

implies that

$$c_{u,v} = (-\lambda)^{k-u-v}\binom{k}{k-u}\binom{k-u}{v}\sum_{\ell=v}^{k-u}(-1)^\ell\binom{k-u-v}{\ell-v},$$

and thus $c_{u,v}$ is zero whenever $u+v<k$, and

$$c_{u,v} = (-1)^v\binom{k}{v} \quad \text{if} \quad u+v=k.$$

To summarize, the left-hand side of equation (2) is equal to

$$\sum_{u+v=k} c_{u,v}A^uB^v = \sum_{v=0}^{k}(-1)^v\binom{k}{v}A^{k-v}B^v,$$

which is zero, by assumption. $\square$

The final Lemma proves the inclusion $E_\lambda^\infty(B) \subseteq E_\lambda^\infty(A)$. As indicated above, this is sufficient to finish the proof of Theorem 1.

Lemma 3. *Under the assumptions of Theorem 1, if $(B-\lambda I)^p v = 0$ for some $p \geq 1$, then $(A-\lambda I)^{p+k-1} v = 0$. Therefore $E_\lambda^\infty(B) \subseteq E_\lambda^\infty(A)$.*

Proof. Using Lemma 2, we multiply equation (2) by $(A-\lambda I)^s$ at the left and by $(B-\lambda I)^t$ at the right, which gives for all $s, t \geq 0$,

$$(A-\lambda I)^{s+k}(B-\lambda I)^t v = \\ \sum_{\ell=1}^{k} (-1)^{\ell+1} \binom{k}{\ell} (A-\lambda I)^{s+k-\ell}(B-\lambda I)^{t+\ell} v. \tag{$\mathsf{E}(s,t)$}$$

Using $(B-\lambda I)^{p+\ell} v = 0$, for all $\ell \geq 0$, equality $\mathsf{E}(0, p-1)$ implies

$$(A-\lambda I)^k (B-\lambda I)^{p-1} v = 0.$$

Let us prove by induction that

$$(A-\lambda I)^{k+j}(B-\lambda I)^{p-j-1} v = 0, \quad \text{for all} \quad 0 \leq j \leq p-1. \tag{3}$$

The case $j=0$ has just been treated. Suppose now that (3) is known for $j = 0, \ldots, m-1$, and let us prove it for $j = m$. Equality $\mathsf{E}(m, p-m-1)$ shows that the left-hand side of (3) for $j = m$ is a linear combination of the vectors

$$(A-\lambda I)^{r+k-1}(B-\lambda I)^{p-r} v, \quad \text{for} \quad m-k+1 \leq r \leq m.$$

Now, these vectors are all zero, since

- either $r \leq 0$, in which case $(B-\lambda I)^{p-r} v = 0$, by assumption,
- or $0 < r \leq m$, in which case $(A-\lambda I)^{r+k-1}(B-\lambda I)^{p-r} v = 0$, by the induction hypothesis.

Therefore, equality (3) is fully proved. For $j = p-1$, it yields the conclusion of Lemma 3. □

4 Final comments.

Theorem 1 is also valid when replacing the field of complex numbers $\mathbb{C}$ by any field $\mathbb{K}$, of arbitrary characteristic. Our proof is straightforwardly adapted to this more general setting.

Part (i) of Problem 1 can be proved in the following way: by (ii), the 2×2 matrix $A-B$ has zero trace. Its square $(A-B)^2$ being by assumption equal to $AB-BA$, it also has zero trace. It follows that $A-B$ is nilpotent, and thus $AB-BA = (A-B)^2 = 0$. This proof works, mutatis mutandis, if the base field $\mathbb{C}$ is replaced by a field $\mathbb{K}$ of characteristic different from 2.

However, there exist 2×2 matrices over $\mathbb{F}_2$ such that $A^2 + B^2 = 2AB$ and $AB \neq BA$. Here is an example:

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \quad AB = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, BA = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, A^2 + B^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Using the fact that the generalized eigenspaces of A and B are equal under (1), we can always restrict A and B to one generalized eigenspace $E_\lambda^\infty(A) = E_\lambda^\infty(B)$. With Lemma 2, (A, B) can replaced by $(A - \lambda I, B - \lambda I)$. So, to find all matrices satisfying the relation (1), one just needs to study the case where A and B are both nilpotent. Still, finding explicitly all nilpotent matrices A, B satisfying (1) might not be an easy task, except in the case $k \geq 2n - 1$ where all nilpotent matrices satisfy the equation. In particular, Theorem 1 gives a necessary condition, but not a sufficient one, given for example that in the case $(k, n) = (2, 2)$, an additional condition is that the matrices commute.

As mentioned in the Introduction, in size higher than 2, the commutation of A and B is not a consequence of equation (1). For any $n, k \geq 2$, $(n, k) \neq (2, 2)$ there exist matrices $A, B \in \mathcal{M}_n(\mathbb{C})$ satisfying equation (1) such that $AB \neq BA$. For $n \geq 3$ and $k \geq 2$, the following matrices do the job

$$A = \begin{bmatrix} 0 & 0 & \cdots & & 0 \\ & & \cdots & & \\ 0 & \cdots & 0 & 1 & 0 \\ 0 & \cdots & 0 & 0 & 1 \\ 0 & \cdots & 0 & 0 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & & \cdots & & 0 \\ & & \cdots & & \\ 0 & \cdots & 0 & 0 & 0 \\ 0 & \cdots & 0 & 0 & 1/2 \\ 0 & \cdots & 0 & 0 & 0 \end{bmatrix}.$$

For $n = 2, k \geq 3$, all pairs of nilpotent matrices satisfy (1), but not all nilpotent matrices of size 2 commute, e.g.

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}.$$

Acknowledgments. We thank the referees for their useful remarks. This work has been supported in part by the Microsoft Research – INRIA Joint Centre and by Université Paris Diderot – Paris 7.

References

[1] Balaj, M., Romanian national mathematical olympiad, April 2011, available at `http://www.onm2011.isjbihor.ro/subiecte/finala11_sol.pdf` (in Romanian).

[2] Weintraub, S. H., *A guide to advanced linear algebra*, The Dolciani Mathematical Expositions, vol. 44, Mathematical Association of America, Washington, DC, 2011, MAA Guides, 6.

Algorithms Project, Inria Paris-Rocquencourt, 78153 Le Chesnay, France
Alin.Bostan@inria.fr

Observatoire de Paris, 77 Avenue Denfert-Rochereau, 75014 Paris, France
Combot@imcce.fr