THE BYRNES–GAUGER THEOREM

In 1978, C. Brynes and M. Gauger published a simple criteria for two matrices to be similar: "Characteristic free, improved decidability criteria for the similarity problem", Linear and Multilinear Algebra 5 (1977), 153–158. The criteria required equality of the characteristic polynomials. In 1979, while studying their proof, I realized to my surprise that one could dispense with this assumption. I was anticipated by the paper of J. Dixon, "An isomorphism criterion for modules over a principal ideal domain", same journal 8 (1979) 69–72.

In this note I present my version of the improved result. The proof requires familiarity with the invariant factors of a matrix and the tensor product of two matrices.

THEOREM 0.1 [Byrnes-Gauger] *Let* $\nu_{A,B} = \nu(A \otimes I_n - I_m \otimes B^t)$, *where* $A \in M_{m\times m}(F)$ *and* $B \in M_{n\times n}(F)$. *Then*

$$\nu_{A,A} + \nu_{B,B} \geq 2\nu_{A,B}, \tag{1}$$

with equality if and only if $m = n$ *and* A *and* B *are similar.*

REMARK. Dixon's result, in matrix terms, states that

$$\nu_{A,A}\nu_{B,B} \geq \nu_{A,B}^2, \tag{2}$$

with equality if and only if A and B are similar to direct sums of a third matrix C.

From the inequality $(\nu_{A,A} + \nu_{B,B})/2 \geq \sqrt{\nu_{A,A}\nu_{B,B}}$, we see that (2) implies (1). Also if $m = n$ and equality holds in (1), then Dixon's result implies A and B are similar.

We need some preliminary results.

THEOREM 0.2 [Cecioni 1908, Frobenius 1910] *Let* $L : U \to U$ *and* $M : V \to V$ *be linear transformations over* F. *Then the vector space* $Z_{L,M}$ *of all linear transformations* $N : U \to V$ *satisfying* $MN = NL$ *has dimension*

$$\sum_{k=1}^{s}\sum_{l=1}^{t} \deg\gcd(d_k, D_l),$$

where $d_1, \ldots, d_s$ *and* $D_1, \ldots, D_t$ *are the invariant factors of* L *and* M, *respectively.*

PROOF. See C.C. MacDuffee, "Theory of matrices", Chelsea 1946, 90–92 or N. Jacobson, "Basic Algebra I", W.H. Freeman and Company, 1974, 197–200.

THEOREM 0.3 *Let* $A \in M_{m\times m}(F), B \in M_{n\times n}(F)$. *Let* $T : M_{m\times n}(F) \to M_{m\times n}(F)$ *be defined by*

$$T(X) = AX - XB,\ X \in M_{m\times n}(F).$$

Then $[T]_\beta^\beta = A \otimes I_n - I_m \otimes B^t$, *where* β *is the standard basis for* $M_{m\times n}(F)$.

COROLLARY 0.1

$$\nu_{A,B} = \sum_{k=1}^{s}\sum_{l=1}^{t} \deg\gcd(d_k, D_l),$$

where

$$d_1|d_2|\cdots|d_s \quad \text{and } D_1|D_2|\cdots|D_t$$

are the invariant factors of A *and* B, *respectively.*

LEMMA 0.1 [Byrnes-Gauger] *Suppose*
$m_1 \le m_2 \le \cdots \le m_s$ *and* $n_1 \le n_2 \le \cdots \le n_s$ *are integer sequences.*
Then

$$\sum_{k=1}^{s}\sum_{l=1}^{s} \{\min(m_k, m_l) + \min(n_k, n_l) - 2\min(m_k, n_l)\} \ge 0.$$

Further, equality occurs if and only if the sequences are identical.

PROOF. **Case 1**: $k = l$.

The terms to consider here are of the form

$$m_k + n_k - 2\min(m_k, n_k)$$

which is obviously ≥ 0. Also, the term is equal to zero iff $m_k = n_k$.
Case 2: $k \ne l$; without loss of generality take $k < l$.

Here we pair the off-diagonal terms (k, l) and l, k.

$$\begin{aligned} & \{\min(m_k, m_l) + \min(n_k, n_l) - 2\min(m_k, n_l)\} \\ & \quad +\{\min(m_l, m_k) + \min(n_l, n_k) - 2\min(m_l, n_k)\} \\ = & \{m_k + n_l - 2\min(m_k, n_l)\} + \{m_l + n_k - 2\min(m_l, n_k)\} \\ \ge & 0. \end{aligned}$$

Since the sum of the diagonal terms and the sum of the pairs of sums of off–diagonal terms are non-negative, the sum is non–negative. Also, if the sum is zero, so must be the sum along the diagonal terms and $m_k = n_k$ for all k.

PROOF OF THE MAIN THEOREM.

$$\begin{aligned}
& \nu_{A,A} + \nu_{B,B} - 2\nu_{A,B} \\
= & \sum_{k_1=1}^{s}\sum_{k_2=1}^{s} \deg\gcd(d_{k_1}, d_{k_2}) + \sum_{l_1=1}^{t}\sum_{l_2=1}^{t} \deg\gcd(D_{l_1}, D_{l_2}) \\
& \quad -2\sum_{k=1}^{s}\sum_{l=1}^{t} \deg\gcd(d_k, D_l).
\end{aligned}$$

We now extend the definitions of $d_1, \ldots, d_s$ and $D_1, \ldots, D_t$ by renaming them as follows, with $N = \max(s, t)$:

$$\begin{aligned}
& \underbrace{1, \ldots, 1,}_{N-s} d_1, \ldots, d_s & \rightarrow & \quad f_1, \ldots, f_N \\
\text{and} \quad & \underbrace{1, \ldots, 1,}_{N-t} D_1, \ldots, D_t & \rightarrow & \quad F_1, \ldots, F_N.
\end{aligned}$$

Then

$$\begin{aligned}
\nu_{A,A} + \nu_{B,B} - 2\nu_{A,B} \quad = \quad & \sum_{k=1}^{N}\sum_{l=1}^{N} \{\deg\gcd(f_k, f_l) + \deg\gcd(F_k, F_l) \\
& \quad -2\deg\gcd(f_k, F_l)\}. \qquad (3)
\end{aligned}$$

We now let $p_1, \ldots, p_r$ be the distinct monic irreducibles in $m_A m_B$ and write

$$\left.\begin{array}{rl}
f_k = & p_1^{a_{k1}} \quad p_2^{a_{k2}} \quad \ldots \quad p_r^{a_{kr}} \\
F_k = & p_1^{b_{k1}} \quad p_2^{b_{k2}} \quad \ldots \quad p_r^{b_{kr}}
\end{array}\right\} \qquad 1 \le k \le N$$

where $\{a_{ki}\}_{i=1}^{r}$, $\{b_{ki}\}_{i=1}^{r}$ are monotonic increasing sequences of non-negative integers. Then

$$\begin{aligned}
& \gcd(f_k, F_l) = \prod_{i=1}^{r} p_i^{\min(a_{ki}, b_{li})} \\
\Rightarrow \quad & \deg\gcd(f_k, F_l) = \sum_{i=1}^{r} \deg p_i \, \min(a_{ki}, b_{li}),
\end{aligned}$$

$$\deg\gcd(f_k, f_l) = \sum_{i=1}^{r} \deg p_i \, \min(a_{ki}, a_{li}),$$
$$\deg\gcd(F_k, F_l) = \sum_{i=1}^{r} \deg p_i \, \min(b_{ki}, b_{li}).$$

Then equation (3) may be rewritten as

$$\begin{aligned}
&\nu_{A,A} + \nu_{B,B} - 2\nu_{A,B} \\
&= \sum_{k=1}^{N}\sum_{l=1}^{N}\sum_{i=1}^{r} \deg p_i \{\min(a_{ki}, a_{li}) + \min(b_{ki}, b_{li}) \\
&\qquad -2\min(a_{ki}, b_{li})\} \\
&= \sum_{i=1}^{r} \deg p_i \sum_{k=1}^{N}\sum_{l=1}^{N} \{\min(a_{ki}, a_{li}) + \min(b_{ki}, b_{li}) \\
&\qquad -2\min(a_{ki}, b_{li})\}.
\end{aligned}$$

We apply the Byrnes–Gauger lemma 0.1 to the latter double sum and since $\deg p_i > 0$, we have

$$\nu_{A,A} + \nu_{B,B} - 2\nu_{A,B} \geq 0,$$

proving the first part of the theorem.

Next we show that equality in the above is equivalent to similarity of A and B:

$$\begin{aligned}
& \nu_{A,A} + \nu_{B,B} - 2\nu_{A,B} = 0 \\
\Leftrightarrow \; & \sum_{i=1}^{r} \deg p_i \sum_{k=1}^{N}\sum_{l=1}^{N} \{\min(a_{ki}, a_{li}) + \min(b_{ki}, b_{li}) \\
& \qquad -2\min(a_{ki}, b_{li})\} = 0 \\
\Leftrightarrow \; & \text{sequences } \{a_{ki}\}, \{b_{ki}\} \text{ identical (by lemma 0.1)} \\
\Leftrightarrow \; & A \text{ and } B \text{ have same invariant factors} \\
\Leftrightarrow \; & A \text{ and } B \text{ are similar } (\Rightarrow m = n).
\end{aligned}$$