ANOTHER ADDENDUM TO KRONECKER'S THEORY OF PENCILS

M. D. Atkinson

SCS-TR-29

August 1983

School of Computer Science
Carleton University
Ottawa, Ontario
K1S 5B6
Canada

This research was supported by the Natural Sciences and Engineering Research Council of Canada.

Abstract: A classical theory due to Kronecker and Weierstrass shows how to decide when two pairs of matrices are equivalent under multiplication by non-singular matrices. Recent work in algebraic computational complexity has demanded a broader definition of equivalence. The classical theory is here extended to the new definition and illustrated by a result on pairs of 3 x 3 matrices: - with the new definition there are only a finite number of inequivalent pairs of 3 x 3 matrices.

ANOTHER ADDENDUM TO KRONECKER'S THEORY OF PENCILS

1. Introduction. There is an old theory dating back to Kronecker and Weierstrass which shows that every pair of $m \times n$ matrices is equivalent, under pre- and post-multiplication by non-singular matrices, to a canonical pair whose entries are largely zeros and ones. For some recent work in algebraic computational complexity [1,4] a weaker, though similar, definition has been appropriate. In [4] Ja'Ja' considered this new equivalence relation and attempted to find computable invariants to distinguish equivalence classes and canonical representatives for each class. Our intention here is to give a substitute for the arguments in section 3 of that paper which unfortunately contain an error that invalidates the main result. As an example of the corrected theory we prove a result about 3×3 matrices.

Let (G_1,G_2) and (H_1,H_2) be two pairs of $m \times n$ matrices over some field F. If there exist non-singular $m \times m$ and $n \times n$ matrices P and Q such that

$$PG_1Q = H_1 \quad \text{and} \quad PG_2Q = H_2$$

we say that the pair (G_1,G_2) is Kronecker equivalent to the pair (H_1,H_2). The classical theory of Kronecker canonical forms [3] shows how to determine a canonical representative from each equivalence class.

The new equivalence relation (tensor equivalence) is that there should exist non-singular matrices P and Q, and a non-singular 2×2 matrix $\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}$ such that

$$PG_1Q = x_{11}H_1 + x_{12}H_2 \quad \text{and} \quad PG_2Q = x_{21}H_1 + x_{22}H_2 .$$

Notice that, if (G_1,G_2) and (H_1,H_2) are Kronecker equivalent, then they are certainly tensor equivalent. The main problems lie in deciding which Kronecker equivalence classes are in the same tensor equivalence class.

Another way of expressing the definition is in terms of the linear spaces

$\mathcal{X} = \langle G_1, G_2 \rangle$ and $\mathcal{Y} = \langle H_1, H_2 \rangle$; (G_1, G_2) is tensor equivalent to (H_1, H_2) if and only if

$$\mathcal{Y} = \{PXQ : X \in \mathcal{X}\}.$$

One other application, further to those mentioned above, is noteworthy. If U, V and W are vector spaces of dimensions m, n and 2 respectively then every bilinear mapping $\phi : U \times V \to W$ corresponds to one and only one tensor equivalence class. This follows from a routine algebraic argument which we omit. The consequence is that the possible bilinear mappings can be classified in terms of the tensor equivalence classes.

In the next section we develop the general theory of tensor equivalence while in the final section we show how a distinguished representative may be selected from each equivalence class.

2. Tensor equivalence. Two $m \times n$ matrices G_1 and G_2 define a homogeneous pencil $\mu G_1 + \lambda G_2$ (μ and λ are indeterminates). The Kronecker theory shows how this pencil is made up of a regular part and a singular part. Our concern here is with regular pencils ($m = n$ and $\det(\mu G_1 + \lambda G_2) \neq 0$) and when two such are tensor equivalent. Then, in conjunction with the results of [4, section 2] (which we recommend be read with this paper), we will have a satisfactory method for deciding when two arbitrary pairs of matrices are tensor equivalent.

Much of the discussion concerns homogeneous polynomials in μ, λ and we preface it with a remark about when such polynomials are "the same". We regard two polynomials $f(\mu, \lambda)$, $g(\mu, \lambda)$ as essentially the same polynomial if they are projectively equivalent, i.e. $f(\mu, \lambda) = kg(\mu, \lambda)$ for some non-zero constant k. In particular each linear polynomial $a\mu + b\lambda$ determines and is determined by a unique ratio a/b (the ratio ∞ corresponds to $b = 0$).

As in [4] we let $\mathcal{D}_k(\mu, \lambda)$ be the greatest common divisor of all minors of order k in the $n \times n$ regular pencil $\mu G_1 + \lambda G_2$. Then the classical homogeneous invariant polynomials are defined by

$$i_k(\mu,\lambda) = \mathcal{D}_{n-k+1}(\mu,\lambda) / \mathcal{D}_{n-k}(\mu,\lambda), \qquad 1 \leq k \leq n.$$

(It can be shown that these quotients are indeed polynomials and, moreover, that $i_k(\mu,\lambda)$ divides $i_{k-1}(\mu,\lambda)$). The invariant polynomials factor into powers of, say r, projectively distinct irreducible polynomials

$$i_1(\mu,\lambda) = [\phi_1(\mu,\lambda)]^{\tau_{11}} [\phi_2(\mu,\lambda)]^{\tau_{12}} \ldots [\phi_r(\mu,\lambda)]^{\tau_{1r}},$$

$$i_2(\mu,\lambda) = [\phi_1(\mu,\lambda)]^{\tau_{21}} [\phi_2(\mu,\lambda)]^{\tau_{22}} \ldots [\phi_r(\mu,\lambda)]^{\tau_{2r}},$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

$$i_n(\mu,\lambda) = [\phi_1(\mu,\lambda)]^{\tau_{n1}} [\phi_2(\mu,\lambda)]^{\tau_{n2}} \ldots [\phi_r(\mu,\lambda)]^{\tau_{nr}}.$$

Of course $\tau_{st} \leq \tau_{s-1,t}$ for each relevant s,t and, to avoid trivialities, we take $\tau_{11}, \tau_{12}, \ldots, \tau_{1r} \neq 0$, although neither of these facts plays a significant part in what follows.

The invariant polynomials (or their irreducible factorizations) characterize the Kronecker equivalence class of the pencil $\mu G_1 + \lambda G_2$. We must investigate these factorizations under general tensor equivalence. To this end let (H_1,H_2) be any pair of $n \times n$ matrices equivalent to (G_1,G_2). Then there exist non-singular matrices P, Q and a non-singular matrix $\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}$ such that

$$PH_1Q = x_{11}G_1 + x_{12}G_2 \quad \text{and} \quad PH_2Q = x_{21}G_1 + x_{22}G_2.$$

The pencil $\mu PH_1Q + \lambda PH_2Q$ has the same invariant polynomials as $\mu H_1 + \lambda H_2$. On the other hand

$$\mu PH_1Q + \lambda PH_2Q = (\mu x_{11} + \lambda x_{21})G_1 + (\mu x_{12} + \lambda x_{22})G_2$$

and therefore the invariant polynomials for $\mu H_1 + \lambda H_2$ are obtained from those of $\mu G_1 + \lambda G_2$ by replacing μ by $\mu x_{11} + \lambda x_{21}$ and λ by $\mu x_{12} + \lambda x_{22}$; and their factorizations are

$$[\phi_1(\mu x_{11} + \lambda x_{21}, \mu x_{12} + \lambda x_{22})]^{\tau_{11}} \quad \ldots\ldots \quad [\phi_r(\mu x_{11} + \lambda x_{21}, \mu x_{12} + \lambda x_{22})]^{\tau_{1r}},$$

$$\cdots\cdots$$

$$[\phi_1(\mu x_{11} + \lambda x_{21}, \mu x_{12} + \lambda x_{22})]^{\tau_{n1}} \quad \ldots\ldots \quad [\phi_r(\mu x_{11} + \lambda x_{21}, \mu x_{12} + \lambda x_{22})]^{\tau_{nr}}.$$

As a consequence we see that the family of column vectors

$$c_1 = \begin{bmatrix} \tau_{11} \\ \vdots \\ \tau_{n1} \end{bmatrix}, \quad c_2 = \begin{bmatrix} \tau_{12} \\ \vdots \\ \tau_{n2} \end{bmatrix}, \ldots\ldots, \quad c_r = \begin{bmatrix} \tau_{1r} \\ \vdots \\ \tau_{nr} \end{bmatrix}$$

remains invariant under tensor equivalence.

From now on we shall take the field F to be algebraically closed so that the irreducible polynomials $\phi_i(\mu,\lambda)$ are linear, say

$$\phi_i(\mu,\lambda) = \alpha_i\mu + \beta_i\lambda \quad \text{for some ratio} \quad \rho_i = \alpha_i/\beta_i .$$

The corresponding polynomials for the equivalent pencil $\mu H_1 + \lambda H_2$ are

$$\phi_i(\mu x_{11} + \lambda x_{21}, \mu x_{12} + \lambda x_{22}) = (\alpha_i x_{11} + \beta_i x_{12})\mu + (\alpha_i x_{21} + \beta_i x_{22})\lambda$$

and are determined by the ratios

$$\frac{\alpha_i x_{11} + \beta_i x_{12}}{\alpha_i x_{21} + \beta_i x_{22}} = \frac{x_{11}\rho_i + x_{12}}{x_{21}\rho_i + x_{22}} .$$

Summing up what we have obtained so far, we know that each regular pencil $\mu G_1 + \lambda G_2$ determines a family of columns $c_1, c_2, \ldots, c_r$ and corresponding distinct ratios $\rho_1, \rho_2, \ldots, \rho_r$ Moreover, if $\mu H_1 + \lambda H_2$ is an equivalent pencil, it determines the same family of columns and the associated ratios have the form $(x_{11}\rho + x_{12})/(x_{21}\rho + x_{22})$ for some non-singular matrix $\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}$.

At this point it is convenient to adopt a notation which recognizes that the r invariant columns may not be distinct. We describe the r invariant columns of some pencil by the column signature

$$c_1^{m_1} c_2^{m_2} \ldots c_t^{m_t}$$

which indicates that m_1 columns are equal to c_1, m_2 columns are equal to c_2, and so on ; thus $\sum m_i = r$. Moreover we order the columns lexicographically so that $c_1 > c_2 > \dots > c_t$.

Then, of the r ratios $\rho_1, \dots, \rho_r$, a set R_1 of m_1 of them is associated with column c_1, a set R_2 of m_2 of them is associated with column c_2, etc. The sequence $(R_1, R_2, \dots, R_t)$ will be called the ratio signature of the pencil. With these notations we may state the fundamental result on tensor equivalence of regular pencils.

THEOREM. Let $\mathcal{P}, \mathcal{Q}$ be two regular pencils of $n \times n$ matrices. Then $\mathcal{P}$ and $\mathcal{Q}$ are tensor equivalent if and only if

(a) they have the same column signature, and

(b) for some non-singular matrix $\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}$ the mapping $\rho \to \dfrac{x_{11}\rho + x_{12}}{x_{21}\rho + x_{22}}$

maps the ratio signature of $\mathcal{P}$ onto the ratio signature of $\mathcal{Q}$.

Proof. The discussion above has shown that, if $\mathcal{P}$ and $\mathcal{Q}$ are equivalent then (a) and (b) hold. Conversely suppose that (a) and (b) both hold. Let $\mathcal{P}$ be the pencil $\mu G_1 + \lambda G_2$. It follows, again from the remarks above, that the equivalent pencil $\widetilde{\mathcal{P}} = \mu(x_{11}G_1 + x_{12}G_2) + \lambda(x_{21}G_1 + x_{22}G_2)$ has the same column signature and ratio signature as $\mathcal{Q}$. But these two signatures determine completely the invariant factors of a pencil and so $\widetilde{\mathcal{P}}$ and $\mathcal{Q}$ are Kronecker equivalent. Thus $\mathcal{P}$ is tensor equivalent to $\mathcal{Q}$.

3. Equivalence class representatives. As it stands it is not evident how criterion (b) of the theorem can be verified. In [2] we give an algorithm for checking this condition whose execution time is proportional to $n^2 \log n$. For the present however we shall consider how to pick a uniquely defined representative from each tensor equivalence class.

The crucial fact that we use is a result from group theory. The group PGL(2,F) of all transformations

$$\rho \to (x_{11}\rho + x_{12})/(x_{21}\rho + x_{22}), \quad x_{11}x_{22} \neq x_{12}x_{21}$$

acts as a sharply triply transitive group on the set $F \cup \{\infty\}$ of ratios ; in other words if $(\alpha_1,\alpha_2,\alpha_3)$ and $(\beta_1,\beta_2,\beta_3)$ are triples of distinct ratios there is a unique transformation in PGL(2,F) which carries each α_i onto β_i. In particular, every three ratios α,β,γ can be uniquely transformed into $\infty,0,1$ respectively and we just have to find a reasonable choice of which three ratios to transform.

The most straightforward case would be when R_1, R_2 and R_3 all had size 1 when we could take the representative of the equivalence class so that $R_1 = \{\infty\}$, $R_2 = \{0\}$ and $R_3 = \{1\}$. Since the only element of PGL(2,F) which fixes ∞,0 and 1 is the identity transformation no further conditions can be placed on the remaining ratios. So, in this case, subject to R_1, R_2, R_3 being $\{\infty\}$, $\{0\}$, $\{1\}$, different ratio signatures give inequivalent pencils.

Other cases are rather trickier to handle and we need to define an ordering on sequences of sets. Let the field F be totally ordered in any way ; for example, if F is the field of complex numbers, we can define $w < z$ if $\mathrm{Re}\, w < \mathrm{Re}\, z$, or if $\mathrm{Re}\, w = \mathrm{Re}\, z$ and $\mathrm{Im}\, w < \mathrm{Im}\, z$. If S and T are subsets of F of the same size we define $S < T$ if, when S and T are each arranged in decreasing order, S lexicographically precedes T. Now if

$$\mathcal{S} = (S_1,S_2, \ldots,S_t) \quad \text{and} \quad \mathcal{T} = (T_1,T_2, \ldots,T_t)$$

are sequences of sets with $|S_i| = |T_i|$, $i=1,2,\ldots,t$ we can define $\mathcal{S} < \mathcal{T}$ if $\mathcal{S}$ lexicographically precedes $\mathcal{T}$ (with the lexicographic order induced by the ordering on component sets).

Using this notion we can now describe a unique representative for each tensor equivalence class. There are four different cases.

(i) $|R_1| = |R_2| = 1$. We take the representative so that $R_1 = \{\infty\}$, $R_2 = \{0\}$ and $1 \in R_3$. Of the m_3 ways of doing this we choose the one which minimizes the sequence $(R_3 - \{1\},R_4,\cdots,R_t)$.

(ii) $|R_1| = 1$, $|R_2| > 1$. Here we take the representative so that $R_1 = \{\infty\}$ and $0,1 \in R_2$. There are $m_2(m_2 - 1)$ ways of doing this and we take the one which minimizes $(R_2 - \{0,1\}, R_3, \ldots, R_t)$.

(iii) $|R_1| = 2$. In this case we take $R_1 = \{\infty, 0\}$ and $1 \in R_2$. There are $2m_2$ ways in which this can be done and we choose the one which minimizes $(R_2 - \{1\}, R_3, \ldots, R_t)$.

(iv) $|R_1| > 2$. We take $\infty, 0, 1 \in R_1$. Of the $m_1(m_1 - 1)(m_1 - 2)$ ways of doing this we take the one which minimizes $(R_1 - \{\infty, 0, 1\}, R_2, \ldots, R_t)$.

Computationally the last case is the hardest. We have to consider all $m_1(m_1 - 1)(m_1 - 2)$ ordered triples from R_1. For each triple (α, β, γ) we calculate the unique transformation which carries (α, β, γ) to $(\infty, 0, 1)$ and apply it to the remaining ratios. Then we have to compare $(R_1 - \{\infty, 0, 1\}, R_2, \ldots, R_t)$ with the smallest such sequence found so far. The amount of work involved is of the order of $n^4 \log n$ operations.

If $n < 3$ some verbal changes have to be made to the exposition above.

Finally we illustrate our results by taking $n = 3$ and showing that there are precisely 6 tensor equivalence classes of 3×3 regular pencils. Using that $\sum_{i,j} \tau_{ij} = 3$ and $\tau_{st} \leqslant \tau_{s-1,t}$ it is easily seen that there are only the following possible column signatures :

$$\text{(i)} \begin{bmatrix}3\\0\\0\end{bmatrix}, \quad \text{(ii)} \begin{bmatrix}2\\1\\0\end{bmatrix}, \quad \text{(iii)} \begin{bmatrix}2\\0\\0\end{bmatrix}\begin{bmatrix}1\\0\\0\end{bmatrix}, \quad \text{(iv)} \begin{bmatrix}1\\1\\0\end{bmatrix}\begin{bmatrix}1\\0\\0\end{bmatrix},$$

$$\text{(v)} \quad c^3 \text{ where } c = \begin{bmatrix}1\\0\\0\end{bmatrix} \quad \text{and} \quad \text{(vi)} \begin{bmatrix}1\\1\\1\end{bmatrix}$$

Each of these column signatures corresponds to just one equivalence class since the ratio signature contains at most 3 distinct ratios, and these can be taken to be $\infty, 0, 1$. Representatives for the classes can be displayed using the

Kronecker-Weierstrass theory. For example,

$$G_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad G_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

represents case (v).

If the regularity condition is omitted only finitely many other tensor equivalence classes are possible and they can be found using the results of [4]. This type of argument can be used to show that there are only a finite number of inequivalent $2 \times 3 \times n$ tensors.

REFERENCES

[1] M. D. ATKINSON AND S. LLOYD, Bounds on the ranks of some 3-tensors, Linear Algebra Appl. 31 (1980) pp. 19-31.

[2] M. D. ATKINSON, Affine and projective equivalence of sets of complex numbers.

[3] F. R. GANTMACHER, The Theory of Matrices, vol. 1 and 2, Chelsea, New York, 1974.

[4] J. JA'JA', An addendum to Kronecker s theory of pencils, Siam J. Appl. Math. 37 (1979), pp. 700-712.

CARLETON UNIVERSITY
School of Computer Science

Bibliography of SCS Reports

SCS-TR-1 THE DESIGN OF CP-6 PASCAL
Jim des Rivieres and Wilf R. LaLonde, June 1982.

SCS-TR-2 SINGLE PRODUCTION ELIMINATION IN LR(1) PARSERS:A SYNTHESIS
Wilf R. LaLonde, June 1982.

SCS-TR-3 A FLEXIBLE COMPILER STRUCTURE THAT ALLOWS DYNAMIC PHASE ORDERING
Wilf R. LaLonde and Jim des Rivieres, June 1982.

SCS-TR-4 A PRACTICAL LONGEST COMMON SUBSEQUENCE ALGORITHM FOR TEXT COLLATION
Jim des Rivieres, June 1982.

SCS-TR-5 A SCHOOL BUS ROUTING AND SCHEDULING PROBLEM
Wolfgang Lindenberg, Frantisek Fiala, July 1982.

SCS-TR-6 ROUTING WITHOUT ROUTING TABLES
Nicola Santoro, Ramez Khatib, July 1982.

SCS-TR-7 CONCURRENCY CONTROL IN LARGE COMPUTER NETWORKS
Nicola Santoro, Hasan Hural, July 1982.

SCS-TR-8 ORDER STATISTICS ON DISTRIBUTED SETS
Nicola Santoro, Jeffrey B. Sidney, July 1982.

SCS-TR-9 OLIGARCHICAL CONTROL OF DISTRIBUTED PROCESSING SYSTEMS
Moshe Krieger, Nicola Santoro, August 1982.

SCS-TR-10 COMMUNICATION BOUNDS FOR SELECTION IN DISTRIBUTED SETS
Nicola Santoro, Jeffrey B. Sidney, September 1982.

SCS-TR-11 A SIMPLE TECHNIQUE FOR CONVERTING FROM A PASCAL SHOP TO C SHOP
Wilf R. LaLonde, John R. Pugh, November 1982.

SCS-TR-12 EFFICIENT ABSTRACT IMPLEMENTATIONS FOR RELATIONAL DATA STRUCTURES
Nicola Santoro, December 1982.

SCS-TR-13 ON THE MESSAGE COMPLEXITY OF DISTRIBUTED PROBLEMS
Nicola Santoro, December 1982.

SCS-TR-14 A COMMON BASIS FOR SIMILARITY MEASURES INVOLVING TWO STRINGS
R.L. Kashyap and B.J. Oommen, January 1983.

SCS-TR-15 SIMILARITY MEASURES FOR SETS OF STRINGS
R.L. Kashyap and B.J. Oommen, January 1983.

SCS-TR-16 THE NOISY SUBSTRING MATCHING PROBLEM
R.L. Kashyap and B.J. Oommen, January 1983.

SCS-TR-17 DISTRIBUTED ELECTION IN A CIRCLE WITHOUT A GLOBAL SENSE OF ORIENTATION
E. Korach, D. Rotem, N. Santoro, January 1983.

SCS-TR-18 A GEOMETRICAL APPROACH TO POLYGONAL DISSIMILARITY AND THE CLASSIFICATION OF CLOSED BOUNDARIES
R.L. Kashyap and B.J. Oommen, January 1983.

SCS-TR-19 SCALE PRESERVING SMOOTHING OF POLYGONS
R.L. Kashyap and B.J. Oommen, January 1983.

SCS-TR-20 NOT-QUITE-LINEAR RANDOM ACCESS MEMORIES
Jim des Rivieres, Wilf LaLonde and Mike Dixon, August 1982, Revised March 1, 1983.

SCS-TR-21 SHOUT-ECHO SELECTION IN DISTRIBUTED FILES
D. Rotem, N. Santoro, J. B. Sidney, March 1983.

SCS-TR-22 DISTRIBUTED RANKING
E. Korach, D. Rotem, N. Santoro, March 1983

SCS-TR-23 A REDUCTION TECHNIQUE FOR SELECTION IN DISTRIBUTED FILES : I
N. Santoro, J. B. Sidney, April 1983.

SCS-TR-24 LEARNING AUTOMATA POSSESSING ERGODICITY OF THE MEAN : THE TWO ACTION CASE
M.A.L. Thathachar and B. J. Oommen, May 1983

SCS-TR-25 ACTORS - THE STAGE IS SET :
John R. Pugh, June 1983

SCS-TR-26 ON THE ESSENTIAL EQUIVALENCE OF TWO FAMILIES OF LEARNING AUTOMATA
M. A. L. Thathachar and B. J. Oommen, May 1983

SCS-TR-27 GENERALIZED KRYLOV AUTOMATA AND THEIR APPLICABILITY TO LEARNING IN NONSTATIONARY ENVIRONMENTS
B. J. Oommen, June 1983

SCS-TR-28 ACTOR SYSTEMS: SELECTED FEATURES
Wilf R. LaLonde, July 1983.

SCS-TR-29 ANOTHER ADDENDUM TO KRONECKER'S THEORY OF PENCILS
M. D. Atkinson, August 1983

SCS-TR-30 SOME TECHNIQUES FOR GROUP CHARACTER REDUCTION
M. D. Atkinson and R. A. Hassan, August 1983

SCS-TR-31 AN OPTIMAL ALGORITHM FOR GEOMETRICAL CONGRUENCE
M. D. Atkinson, August 1983