Linear algebra in bounded arithmetic

Satoru Kuroda Gunma Prefectural Women's University

1 Introduction

Proving mathematical theorems in theories of bounded arithmetic is a fruitful problem. It sheds a light on the structure of various branches in mathematics as well as giving a new insight into the propositional proof complexity.

Among such problems, theorems in linear algebra gained much attention. Skelley and Cook [7] gave formal systems which treat operations of matrices and studied the proof complexity of theorems in linear algebra, especially those concerning the determinant.

Soon after, Tzameret and Cook [8] showed, based on the earlier result by Hrubes and Tzameret [4], that the multiplicativity of the determinant

$$\det(AB) = \det(A)\det(B)$$

can be proved in \mathbf{VNC}^2 .

In complexity theory, many properties of matrices can be computed in the complexity class DET, logspace reducible to the determinant (cf. von zur Gathen [9]). Since DET is equivalent to the logspace counting class #L, we might expect that a fairly amount of linear algebra can be done in the theory for #L or its mild extensions.

For the class #L, a theory $\mathbf{V}\#\mathbf{L}$ is defined by Cook and Fontes [2] by formalizing the power of matrix. However, there seems to be a large gap between $\mathbf{V}\#\mathbf{L}$ and \mathbf{VNC}^2 .

There are complexity classes in between, $\#SAC^1$ and TC^1 . So we may consider the corresponding theory. The main contribution of this article is the construction of such theories $V\#SAC^1$ and VTC^1 and the inclusions of the complexity classes are preserved:

$$\mathbf{V} \# \mathbf{L} \subseteq \mathbf{V} \# \mathbf{SAC}^1 \subseteq \mathbf{VTC}^1 \subseteq \mathbf{VNC}^2.$$

Turing our attensions to the provability of determinant identities, there are several options for the choice of the definition. Tzamaret and Cook adopted NC^2 algorithm based on Schur complement. On the other hand, Skelley and Cook [7] formlized Berkowitz's #L algorithm in his theory of linear algebra. In the final section, we propose the third approach which uses the algorithm by Mahajan and Vinay [5]. This algorithm is based on the fact that the determinant can be defined by way of clow sequences and the algorithm is purely combinatorial. So we may expect essentially different proofs.

It seems that there are many interesting problems in the relation between bounded arithmetic and linear algebra. For instance, many theorems in combinatorics are proved using linear algebra methods. So proving nontrivial upper bounds on the provability of linear algebra theorems will be a great leap forward in proving theorems in other branches of mathematics in bounded arithmetic.

We omit the detailed proofs for most of the results. Rather we concentrate on presenting overview of the state of the art of the connection between linear algebra and weak systems of arithmetic.

2 The theory VTC^1

The theories constructed in the following sections are based on the theory \mathbf{VTC}^0 . For the detail of the theory, we refer the textbook by Cook and Nguyen [3].

We define a theory whose provably total functions are exactly those computable by polynomial size and logarithmic depth circuits with majority gates. We encode circuits by $d \times w$ two dimensional array so that d and w denote the depth and the width. So gates are of the form $\langle x, y \rangle$ for x < d and y < w where x determines the layer in which the gate is placed.

Without loss of generality, we may assume that TC^1 circuits consists only of majority gates

$$majority_m(x_0, \dots, x_{m-1}) = \begin{cases} 1 & \text{if } \sum x_i > \lfloor m \rfloor/2, \\ 0 & \text{otherwise.} \end{cases}$$

So we encode TC^1 circuits by a string $E \subseteq ([d] \times [w])^2$ for logarithmic d which represents the input-output relation in such a way that

 $\langle x, y \rangle$ receives an input from $\langle x', y' \rangle \Leftrightarrow E(x', y', x, y) \land x' < x$.

We construct the theory \mathbf{VTC}^1 by expanding \mathbf{VTC}^0 with a single axiom which states that any TC^1 circuit can be evaluated. We adopt \mathbf{VTC}^0 as our base theory instead of \mathbf{V}^0 since the evaluation of majority gates requires TC^0 functions.

We occationally denote the cardinality of Σ^B_0 definable sets as

$$|\{x < a : \varphi(x)\}|$$

for $\varphi(x) \in \Sigma_0^B$.

Define the formula $\delta_{TC}(d, w, E, I, Y)$ as

$$\begin{split} \delta_{TC}(d, w, E, I, Y) \Leftrightarrow \\ \forall y < w \ (Y(0, y) \leftrightarrow I(y)) \land 0 < \forall x < d \ \forall y < w \\ Y(x, y) \ \leftrightarrow \ |\{\langle x', y'\rangle \in [d] \times [w] : x' < x \land E(x', y', x, y) \land Y(x', y')\}| \\ &> \lfloor |\{\langle x', y'\rangle \in [d] \times [w] : E(x', y', x, y) \land Y(x', y')\}| \rfloor /2 \end{split}$$

Definition 1 Let

 $TCV \equiv \forall d, w \; \forall C, I \; \exists Y \subseteq [|d|] \times [w] \; \delta_{TC}(|d|, w, E, I, Y).$

We define the theory \mathbf{VTC}^1 to be $\mathbf{VTC}^0 + TCV$.

Theorem 1 A function is computable by a uniform TC^1 circuit family if and only if it is Σ_1^B definable in \mathbf{VTC}^1 .

3 Coding arithmetic circuits

Our first goal is to show that a certain fragment of PI^{-1} -proofs are coded and its consistency is provable in \mathbf{VTC}^1 . To this end, it is required that arithmetic circuits used in such proofs can be evaluated in \mathbf{VTC}^1 .

Definition 2 Arithmetic TC^1 circuis are arithmetic circuits of $O(\log n)$ depth and $n^{O(1)}$ size with unbounded fan-in + gates and fan-in two × and ÷ gates.

We code the computation of arithmetic TC^1 circuits in a similar manner as that for TC^1 circuits. An arithmetic circuit with +, \times and \div gates is a pair $\langle C, E \rangle$ such that $C : [d] \times [w] \to 2$ and $E \subseteq ([d] \times [w])^2$. The intended meaning of C is that for x < d and y < w,

$$C[x, y] = 0 \Rightarrow \langle x, y \rangle \text{ is } +,$$

$$C[x, y] = 1 \Rightarrow \langle x, y \rangle \text{ is } \times and$$

$$C[x, y] = 2 \Rightarrow \langle x, y \rangle \text{ is } \div.$$

For a *times* gate or a \div gate, we assume that it receives two inputs

$$input_0(x, y, E) = \min\{\langle x', y' \rangle \in [d] \times [w] : x' < x \land E(x', y', x, y)\}$$

and

$$input_1(x, y, E) = \min\{\langle x', y' \rangle \in [d] \times [w] : x' < x \land E(x', y', x, y) \land \langle x', y' \rangle > input_0(x, y, E)\}$$

If $\langle x, y \rangle$ is a \times gate then its output is given by the multiplication of outputs of its two inputs. If $\langle x, y \rangle$ is a \div gate then outputs of $input_0(x, y, E)$ and $input_1(x, y, E)$ are treated as the nominator and denominator respectively. If the output of $input_1(x, y, E)$ is 0 then the output if $\langle x, y \rangle$ is undefined and is assigned -1.

Note that unbounded fan-in + gates are computed by vector summation which can be computed by TC^0 circuits.

We encode the value of each gate in an arithmetic TC^1 circuit by an integer, that is a number with its signature. The only case in which the gate has a negative value is when it is undefined as stated above.

Now define the formula $\delta_{arithTC}(d, w, C, E, I, Y)$ as

$$\begin{split} \delta_{arithTC}(d, w, C, E, I,) \Leftrightarrow \\ \forall y < w(Y[0, y] = \langle +, I[y] \rangle) \wedge \\ 0 < \forall x < d \; \forall y < w \begin{pmatrix} (C[x, y] = 0 \land SumComp(x, y, E, Y)) \lor \\ (C[x, y] = 1 \land MultComp(x, y, E, Y)) \lor \\ (C[x, y] = 2 \land DivComp(x, y, E, Y)) \end{pmatrix} \end{split}$$

where

$$SumComp(x, y, E, Y) \equiv \begin{pmatrix} \forall \langle x', y' \rangle \in [d] \times [w] \ (x' < x \land E(x', y', x, y) \to Y[x', y'] \neq -1) \land \\ Y[x, y] = \sum \{ \langle x', y' \rangle : x' < x \land y < w \land E(x', y', x, y) \} \end{pmatrix} \lor \\ \begin{pmatrix} \exists \langle x', y' \rangle \in [d] \times [w] \ (x' < x \land E(x', y', x, y) \land Y[x', y'] = -1) \land \\ Y[x, y] = -1 \end{pmatrix} \end{pmatrix} MultComp(x, y, E, Y) \equiv$$

$$\left(\begin{array}{c} \bigwedge_{k=0,1} Y[input_k(x,y,E,Y)] \neq -1 \land \\ Y[x,y] = Y[input_0(x,y,E,Y)] \cdot Y[input_1(x,y,E,Y)] \end{array} \right) \lor \\ \left(\bigvee_{k=0,1} Y[input_k(x,y,E,Y)] = -1 \land Y[x,y] = -1 \right)$$

 $DivComp(x, y, E, Y) \equiv$

$$\begin{pmatrix} \bigwedge_{k=0,1} Y[input_k(x,y,E,Y)] \neq -1 \land Y[input_1(x,y,E)] \neq 0 \land \\ Y[x,y] = Y[input_0(x,y,E,Y)]/Y[input_1(x,y,E,Y)] \end{pmatrix} \lor$$
$$\begin{pmatrix} \left(\bigvee_{k=0,1} (Y[input_k(x,y,E,Y)] = -1 \land Y[x,y] = -1) \lor Y[input_1(x,y,E)] = 0 \right) \land \\ Y[x,y] = -1 \end{pmatrix} \end{pmatrix}$$

Theorem 2 \mathbf{VTC}^1 proves the following:

 $\forall d, w \; \forall C, E, I \; C : [|d|] \times [w] \to 2 \to \exists Y \leq t(d, w) \; \delta_{arithTC}(|d|, w, C, E, I, Y).$

(Proof). It is easy to see that each gate in arithmetic TC^1 circuits can be computed by some TC^0 circuit. so we ca construct any arithmetic TC^1 circuit by replacing each gate by the corresponding TC^0 circuit.

Now we associate an algebraic proof system to \mathbf{VTC}^1 .

Definition 3 $PI^{-1}(TC)$ -proofs are PI^{-1} -proofs in which all formulas are arithmetic TC^1 circuits.

It is easy to see that there exists a Σ_0^B formula $Prf_{PI^{-1}(TC)}(P, \langle C, E \rangle = \langle C', E' \rangle)$ which states that P is a $PI^{-1}(TC)$ -proof of the equiation $\langle C, E \rangle = \langle C', E' \rangle$ for arithmetic TC^1 circuits $\langle C, E \rangle$ and $\langle C', E' \rangle$). Also there exists a Σ_0^B formula Sat(C, E, C', E', I) which states that the equation $\langle C, E \rangle = \langle C', E' \rangle$ holds for all input I.

According to these formalization, we prove that \mathbf{VTC}^1 proves the soundness of $PI^{-1}(TC)$ -proofs.

Theorem 3 VTC^1 proves the following:

 $\forall \langle C, E \rangle \exists P \ Prf_{PI^{-1}(TC)}(P, \langle C, E \rangle = \langle C', E' \rangle) \rightarrow \forall I \ Sat(C, E, C', E', I).$

(Proof Sketch). By Theorem 2, \mathbf{VTC}^1 can evaluate arithmetic TC^1 circuits. So the claim follows from by induction on the number of inference rules in a given $PI^{-1}(TC)$ -proof. \Box

4 The theory $V # SAC^1$

We define the theory for $\#SAC^1$ by formalizing the following characterization.

Theorem 4 A function is in $\#SAC^1$ if and only if it is computable by a uniform family of arithmetic circuits of polynomial size and degree with + and \times gates.

We code criticity by a pair $\langle C, E \rangle$ with $C \subseteq [d] \times [n]$ and $E \subseteq ([d] \times [n])^2$ where d and n denotes the depth and the width as before. The intended meaning is that C determines the type of the gate $\langle x, y \rangle$ in such a way that

$$C(x, y) \Rightarrow \langle x, y \rangle \text{ is a } \times \text{ gate,} \neg C(x, y) \Rightarrow \langle x, y \rangle \text{ is a } + \text{ gate,}$$

and E gives the input-output relation.

The idea of the formalization is to code the computation of a given circuit on a put of lenth n by two lists D which codes the degree of each gate and V which codes the output of each gate. The degree of the gate has the upperbound n^k for $k \in \omega$ and if the degree of a gate exceeds n^k then the output is set to 0.

Specifically, we formalize this by the formula $\delta^k_{\#SAC^1}(d, n, C, E, I, D, V)$ which is defined by

$$\begin{split} &\delta^k_{\#SAC^1}(d,n,C,E,I,D,V) \leftrightarrow \\ &\forall y < n(D[0,y] = 1 \land V[0,y] = I[y]) \land \\ &0 < \forall x < d \; \forall y < w \left(\begin{array}{c} (C(x,y) \land MultDegComp^k(x,y,n,E,D,V)) \\ \lor (\neg C(x,y) \land SumDegComp^k(x,y,n,E,D,V)) \end{array} \right) \end{split}$$

where

$$\begin{split} &MultDegComp^{k}(x,y,n,E,D,V)\Leftrightarrow\\ &\left(\begin{array}{l} D[input_{0}(x,y,E)]+D[input_{1}(x,y,E)] < n^{k}\wedge\\ D[x,y]=D[input_{0}(x,y,E)]+D[input_{1}(x,y,E)]\wedge\\ V[x,y]=V[input_{0}(x,y,E)]\cdot V[input_{1}(x,y,E)] \end{array}\right)\vee\\ &\left(\begin{array}{l} D[input_{0}(x,y,E)]+D[input_{1}(x,y,E)] \geq n^{k}\wedge\\ D[x,y]=n^{k}\wedge V[x,y]=0 \end{array}\right) \end{split}$$

appendix

$$\begin{aligned} MultDegComp^{k}(x, y, n, E, D, V) \Leftrightarrow \\ & \left(\begin{array}{c} D[input_{0}(x, y, E)] + D[input_{1}(x, y, E)] < n^{k} \wedge \\ D[x, y] = D[input_{0}(x, y, E)] + D[input_{1}(x, y, E)] \wedge \\ V[x, y] = V[input_{0}(x, y, E)] \cdot V[input_{1}(x, y, E)] \end{array}\right) \vee \\ & \left(\begin{array}{c} D[input_{0}(x, y, E)] + D[input_{1}(x, y, E)] \geq n^{k} \wedge \\ D[x, y] = n^{k} \wedge V[x, y] = 0 \end{array}\right) \end{aligned}$$

and

 $SumDegComp^k(x,y,n,E,D,V) \Leftrightarrow$

$$\left(\begin{array}{l} \forall \langle x', y' \rangle in[d] \times [w] \ (x' < x \land E(x', y', x, y) \to D[x', y'] < n^k \land \\ D[x, y] = \max\{D[x', y'] : x' < x \land y < w \land E(x', y', x, y)\} \land \\ V[x, y] = \sum\{V[x', y'] : x' < x \land y < w \land E(x', y', x, y)\} \\ \left(\begin{array}{l} \exists \langle x', y' \rangle in[d] \times [w] \ (x' < x \land E(x', y', x, y) \land D[x', y'] \ge n^k \land \\ D[x, y] = n^k \land V[x, y] = 0 \end{array} \right)$$

Definition 4 For $k \in \omega$. $\#SAC^1V_k$ denotes the formula

$$\forall d, n \; \forall C, E, I \exists D \le dn^{k+1} \; \exists V \le dn^{k+1} \; \delta^k_{\#SAC^1}(d, n, C, E, I, D, V)$$

The theory $\mathbf{V} \# \mathbf{SAC}^1$ is axiomatized by

$$\mathbf{VTC}^0 + \{\#\mathsf{SAC}^1V_k : k \in \omega\}.$$

Theorem 5 A function is computable by $\#SAC^1$ circuits if and only if it is Σ_1^B -definable in $\mathbf{V}\#\mathbf{SAC}^1$.

5 Proving properties of the determinant

In this section we consider the problem of whether properties of the determinant can be proved in \mathbf{VTC}^1 or $\mathbf{V}\#\mathbf{SAC}^1$. It is already proved by Tzameret and Cook [8] that \mathbf{VNC}^2 proves basic properties of the determinant, namely,

Theorem 6 (Tzameret-Cook) VNC^2 proves the followings:

- 1. the multiplicativity of the determinant: det(AB) = det(A)det(B),
- 2. the cofactor expansion
- 3. Cayley-Hamilton Theorem

The proof heavily depends on how we define the determinant in the theory and the above theorem uses the recursive definition via Schur complement.

On the other hand, several other algorithms for the determinant are known. Berkowitz [1] gave a #L algorithm for the characteristic polynomial in the following manner.

Berkowitz's Algorithm:

Let A be a $n \times n$ matrix and define A_0, A_1, \ldots, A_n as

$$A = A_0 = \begin{pmatrix} a_{11} & R_1 \\ S_1 & A_1 \end{pmatrix} \text{ and } A_{k-1} = \begin{pmatrix} a_{kk} & R_k \\ S_k & A_k \end{pmatrix}.$$

Then define $B_1 f, B_2, \ldots, B_n$ by

$$B_{i} = \begin{pmatrix} 1 & 0 & \cdots \\ -a_{ii} & 1 & \cdots \\ -R_{i}S_{i} & -a_{ii} & \cdots \\ -R_{i}A_{i}S_{i} & -R_{i}S_{i} & \cdots \\ \vdots & \vdots & \ddots \\ -R_{i}A_{i}^{i-2}S_{i} & -R_{i}A_{i}^{i-3}S_{i} & \cdots \end{pmatrix}$$

Finally, let

$$P(A) = B_1 B_2 \cdots B_n.$$

It is easily seen that this algorithm can be done in #L. Skelley and Cook [7] formlized this algorithm in their theory of linear algebra and showed that the multiplicativity of the determinant can be proved in the system LAP expanded by two axioms

$$CHT \equiv p_A(A) = 0,$$

$$det(A) = 0 \to \forall B \ AB \neq I.$$

It is also easy to see that these two definitions of the determinant are equivalent provably in \mathbf{VNC}^2 .

Theorem 7 Let $P(A) = (p_n, ..., p_0)$ be the sequence which are coefficients of the characteristic polynomial of A. Then \mathbf{VNC}^2 proves that

$$\det(xI - A) = p_n x^n + \dots + p_0.$$

$$\det(xI - A) = p_n x^n + \dots + p_0.$$

has a Polynomial Identity proof with division provably in \mathbf{VNC}^2 . Specifically, we formalize the following proof in \mathbf{VNC}^2 .

Let A be a $n \times n$ matrix. If n = 1 then the proof is trivial. Otherwise, let

$$A = \begin{pmatrix} a_{11} & R_1 \\ S_1 & A_1 \end{pmatrix}$$

where R_1 , S_1 and A_1 are $1 \times (n-1)$, $(n-1) \times 1$ and $(n-1) \times (n-1)$ matrices respectively. By the definition of det(A), we have

$$\det(xI - A) = (x - a_{11}) \det(xI - A_1) - R_1 \det(xI - A_1)(xI - A_1)^{-1}S_1.$$

Let

$$p_{A_1}(x) = q_{n-1}x^{n-1} + \dots + q_0.$$

and

$$B(x) = \sum_{k=2}^{n} (q_{n-1}A^{k-2} + \dots + q_{n-k+1}I)x^{n-k}$$

By a simple calculation from the definition of $p_A(x)$, we have

$$p_A(x) = (x - a_{11})p_{A_1}(x) - R_1 B(x)S_1$$

and by the inductive hypothesis, we have

$$p_{A_1}(x) = \det(xI - A_1).$$

So it suffice to show that

$$B(x) = \det(xI - A_1)(xI - A_1)^{-1}$$

Let B'(x) be obtained from B(x) by replacing the coefficient $p_k^{A_1}$ by the corresponding coefficient in det $(xI - A_1)$. Then by the inductive hypothesis, it suffices to show that

$$B'(x) = \det(xI - A_1)(xI - A_1)^{-1}.$$

Let adj(A) denote the adjunct of A. Then as in Soltys [6], we can show that

$$adj(xI - A_1)(xI - A_1) = \det(xI - A_1)I$$

which proves the claim.

Now as in [8], we convert the PI^{-1} -proof into **VNC**²-Frege proof.

There is yet another algorithm. Mahajan and Vinay [5] constructed a $\#SAC^1$ algorithm which utilizes clow sequences. Specifically, a clow in [n] is a closed walk $\langle w_1, \ldots, w_l \rangle$ where $w_1, \ldots, w_l < n$ and the first element, called the head of the clow head(C), can only occur once in the walk. A clow sequence is a sequence $\langle C_1, \ldots, C_k \rangle$ of clows such that

- $head(C_1) < \cdots < head(C_k)$ and
- the sum of edges in the sequence is [n].

Then weight $w(C) = w(\langle w_1, \ldots, w_l \rangle)$ of a clow on a $n \times n$ matrix $A = (a_{ij})$ is defined to be

$$w(C) = a_{w_1w_2}\cdots a_{w_{l-1}w_l}a_{w_lw_1}$$

and the weight of a clow sequence $W = \langle C_1, \ldots, C_k \rangle$ is

$$w(W) = \prod_{1 \le i \le k} w(C_i)$$

The sign sgn(W) is defined to be $(-1)^{n+k}$.

The it is known that

Theorem 8 Let A be a $n \times n$ matrix. Then

$$\det(A) = \sum_{W : clow sequence in [n]} sgn(W)w(W).$$

Using this theorem and divide and conquer technique, a $\#SAC^1$ algorithm for the deteminant is given in [5]. Moreover, the algorithm can be formalized in $V\#SAC^1$.

Theorem 9 There is a Σ_0^B definition of the $\#SAC^1$ circuit for the determinant by Mahajan and Vinay provably in $\mathbf{V}\#\mathbf{SAC}^1$.

We finish the article by addressing a problem.

Problem 1 Let $\det_{clow}(A)$ be the function computing the determinant computed by Mahajan and Vinay algorithm. Show that the multiplicativity of the \det_{clow} is provable in $\mathbf{V}\#\mathbf{SAC}^1$.

References

- Berkowitz, On computing the determinant in small parallel time using a small number of processors. Information Processing Letters, 18(3):147–150, (1984).
- [2] Cook, S. and Fontes, L., Formal Theories for Linear Algebra. Logical Methods in Compter Science, Volume 8, Issue 1. (2012)
- [3] Cook, S., and Nguyen, P. Logical Foundations of Proof Complexity. Cambridge: Cambridge University Press. Iddo Tzameret, Stephen A. Cook: Uniform, Integral, and Feasible Proofs for the Determinant Identities. J. ACM 68(2): 12:1-12:80 (2010)
- [4] Hrubes, P., and Tzameret, I., Short Proofs for the Determinant Identities. SIAM J. Comput. 44(2) 340-383 (2015)

- [5] Mahajan, M., Vinay, V., A Combinatorial Algorithm for the Determinant. In: Proceedings of the Eight Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 1997 (1997)
- [6] Soltys, M., Berkowitz's Algorithm and Clow Sequences, Electronic Journal of Linear Algebra, 9:42-54, (2002)
- [7] Soltys, M., and Cook, S., The proof complexity of linear algebra, Annals of Pure and Applied Logic, 130(1–3), 277-323, (2004)
- [8] Tzameret, I., Cook, S. Uniform, Integral and Feasible Proofs for the Determinant Identities Journal of the ACM, Vol. 68, No. 2 (2021)
- [9] von zur Gathen, J., Parallel linear algebra. In John H. Reif, editor, Synthesis of Parallel Algorithms, pages 574–617. Morgan and Kaufman, (1993)

群馬県立女子大学・文学部 黒田 覚