Mika Hirvensalo

A Method for Computing the Characteristic Polynomial and Determining Semidefiniteness

Turku Centre *for* Computer Science

TUCS Technical Report
No 727, December 2005

A Method for Computing the Characteristic Polynomial and Determining Semidefiniteness

Mika Hirvensalo
TUCS–Turku Centre for Computer Science and
Department of Mathematics, University of Turku
FIN-20014 Turku, Finland
mikhirve@cs.utu.fi

The research was done while the author was
visiting the Institute for Quantum Computing,
Waterloo, Canada in 2005.
The author is grateful to IQC for general hospitality.

Supported by the Academy of Finland under grant 208797.

TUCS Technical Report
No 727, December 2005

Abstract

We point out that to determine whether a given self-adjoint matrix A is positive semidefinite is equivalent to determining whether the characteristic polynomial of A is *alternating*, and give a new algorithm for computing the characteristic polynomial. The said algorithm uses $\mathcal{O}(n^3 \log n)$ multiplications and additions, and one division when computing any coefficient of the characteristic polynomial of an $n \times n$ matrix. The algorithm is essentially division-free, and hence it can be also used to compute the determinant for $n \times n$ matrices over a great variety of commutative rings. It takes $\mathcal{O}(n^3 \log n)$ ring multiplications, additions, $\mathbb{Z}$-module multiplications, and one $\mathbb{Z}$-module division to compute the determinant.

Keywords: Determinant, Characteristic Polynomial, Positive Semidefinite, Polynomial Time

TUCS Laboratory
Discrete Mathematics for Information Technology

1 Introduction

In the usual *Hilbert space formalism* of quantum mechanics, the states of quantum systems are identified with self-adjoint mappings that have unit trace and are positive semidefinite. In quantum physics literature, term "positive semidefinite" is sometimes replaced by "positive", and we will also use the latter term in this article.

It follows that when studying quantum mechanics, it is sometimes necessary to discover whether a given matrix is indeed a valid representation of a state of a quantum system or not. That is, one should find out if the matrix is self-adjoint, has unit trace, and is positive semidefinite.

The two first conditions, namely the self-adjointness and the unit trace property are very straightforward to verify, but the third one, the positive semidefiniteness seems to be much more complicated. The following criterion occurs in the literature very frequently: Matrix A is positive semidefinite if and only if the principal minors of A are all nonnegative. The mentioned criterion is theoretically very interesting but appears very tedious to verify, since for an $n \times n$-matrix, there are 2^n principal minors, including the two trivial ones.

In this article, we are interested only in finite-dimensional Hilbert spaces, and hence we can identify the Hilbert space with the Cartesian product $\mathbb{C}^n$. Having this identification, we also assume that a linear mapping $\mathbb{C}^n \to \mathbb{C}^n$ is specified by a given $n \times n$-matrix A over complex numbers.

Here we will not pay any particular attention to the algorithmic aspects of representing complex numbers. Instead, we will present the main result just by counting how many multiplications and additions are needed. In fact, when introducing the algorithm for computing the characteristic polynomial, we only assume that the matrix entries are from a commutative ring, and we will also count the number of $\mathbb{Z}$-module multiplications (defined later).

We present our criterion for positive semidefiniteness as a theorem below. The notations and terminology involved in the theorem are explained in the next section.

Theorem 1. *Define an infinite sequence of expressions as follows:*

The kth expression is defined as

$$\sum_{r=1}^{k}(-1)^{k-r} \sum_{\substack{l_1+\ldots+l_r=k \\ 1\le l_1\le\ldots\le l_r\le k}} N_k(l_1,\ldots,l_r)\,\mathrm{Tr}(A^{l_1})\cdot\ldots\cdot\mathrm{Tr}(A^{l_r}), \tag{1}$$

where $N_k(l_1,\ldots,l_r)$ is the number of permutations in S_k having cycle structure $(l_1,\ldots,l_r)$. Hence a few first expressions are as follows:

$$\begin{cases} \mathrm{Tr}(A) \\ \mathrm{Tr}(A)^2 - \mathrm{Tr}(A^2) \\ \mathrm{Tr}(A)^3 - 3\,\mathrm{Tr}(A)\,\mathrm{Tr}(A^2) + 2\,\mathrm{Tr}(A^3) \\ \mathrm{Tr}(A)^4 - 6\,\mathrm{Tr}(A)^2\,\mathrm{Tr}(A^2) + 3\,\mathrm{Tr}(A^2)^2 + 8\,\mathrm{Tr}(A)\,\mathrm{Tr}(A^3) - 6\,\mathrm{Tr}(A^4) \\ \ldots \end{cases} \tag{2}$$

Then, an $n \times n$ self-adjoint matrix A is positive semidefinite if and only if n first expressions in (2) are nonnegative.

The above criterion leads into a new algorithm (using polynomial number of arithmetical operations) for verifying the positive semidefiniteness, even though the number of summands in (1) equals to the number of partitions of k, which is asymptotically $\frac{1}{4k\sqrt{3}}e^{\pi\sqrt{2k/3}}$ (see [1]), superpolynomial with respect to $k \leq n$. Despite the large number of summands, we show later how to compute (1) in polynomial time with respect to $k \leq n$.

The criterion of Theorem 1 establishes also a connection between the coefficients of the characteristic polynomial and the traces of the powers of the matrix, a connection which is interesting on its own.

2 Preliminaries and Terminology

Notation H_n stands for an n-dimensional Hilbert space, and $L(H_n)$ means the linear mappings $H_n \to H_n$. We will identify H_n with $\mathbb{C}^n$, and $L(H_n)$ with $n \times n$ matrices with complex entries.

We denote the complex conjugate of $z \in \mathbb{C}$ by z^*. The *adjoint matrix* A^* of A is given by taking the complex conjugate of the transpose matrix of A. That is, if the matrix entries of A are A_{ij}, then the matrix entries of A^* are given by $(A^*)_{ij} = (A_{ji})^*$. Matrix A is *self-adjoint*, if $A = A^*$.

Space $H_n = \mathbb{C}^n$ is equipped with so-called *Hermitian inner product*

$$\langle \boldsymbol{x} \mid \boldsymbol{y} \rangle = x_1^* y_1 + \ldots + x_n^* y_n.$$

We understand the vectors $\boldsymbol{x} \in H_n$ as column vectors, hence the image of $\boldsymbol{x}$ under mapping $A \in L(H_n)$ is $A\boldsymbol{x}$, understood as the product of $n \times n$-matrix and a column vector.

The *trace* of matrix A is denoted by $\mathrm{Tr}(A)$ and defined as the sum of the diagonal elements of A:

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} A_{ii}.$$

Matrix A is *positive* or *positive semidefinite*, if

$$\langle \boldsymbol{x} \mid A\boldsymbol{x} \rangle \geq 0$$

for each $\boldsymbol{x} \in H_n$. It is a well-known fact that if $\langle \boldsymbol{x} \mid A\boldsymbol{x} \rangle \in \mathbb{R}$ for each $\boldsymbol{x} \in H_n$, then A is self-adjoint. Hence it makes no sense to define the notion of positivity in this way for $n \times n$-matrices that are not self-adjoint.

A *submatrix* of A is a matrix that is obtained by deleting some rows and columns of A. A *principal submatrix* of an $n \times n$ matrix is a submatrix obtained by deleting some rows and the *corresponding* columns. A *minor* of A is a determinant of a principal submatrix of A.

If A is a $n \times n$-matrix, and $I \subseteq \{1, \ldots, n\}$, notation $A[I]$ stands for the principal submatrix of A obtained by deleting ith row and column for each $i \notin I$. Hence

$A[I]$ is an $|I| \times |I|$-matrix that has those rows and columns whose indices are in I left. Matrix $A[\emptyset]$ is defined to be 1 and $A[\{1,\ldots,n\}]$ is naturally understood as A.

The *characteristic polynomial* of a $n \times n$ matrix A is defined as

$$p_A(\lambda) = (-1)^n \det(A - \lambda I) = \det(\lambda I - A),$$

where I is the $n \times n$ identity matrix. As it is well-known, $p_A(\lambda)$ is a monic polynomial, and the roots of $p_A(\lambda)$ are the eigenvalues of A. It is obvious that if A is self-adjoint, so is also $A[I]$ for any $I \subseteq \{1,\ldots,n\}$, and the basic properties of the determinants imply easily that the characteristic polynomial of a self-adjoint matrix has real coefficients.

The *symmetric basic functions* $V_i = V_i(\lambda_1,\ldots,\lambda_n)$ on variables λ_1, ..., λ_n are defined by identity

$$(\lambda - \lambda_1) \cdot \ldots \cdot (\lambda - \lambda_n) = V_0\lambda^n - V_1\lambda^{n-1} + V_2\lambda^{n-2} + \ldots + (-1)^n V_n. \quad (3)$$

The explicit expressions for V_i can also be found straightforwardly:

$$\left\{ \begin{array}{lcl} V_0(\lambda_1,\ldots,\lambda_n) & = & 1, \\ V_1(\lambda_1,\ldots,\lambda_n) & = & \lambda_1 + \lambda_2 + \ldots + \lambda_n, \\ V_2(\lambda_1,\ldots,\lambda_n) & = & \lambda_1\lambda_2 + \lambda_1\lambda_2 + \ldots + \lambda_{n-1}\lambda_n, \\ \ldots & & \ldots \\ V_n(\lambda_1,\ldots,\lambda_n) & = & \lambda_1\lambda_2 \cdot \ldots \cdot \lambda_n. \end{array} \right. \quad (4)$$

Notation S_n stands for the symmetric group on n elements, that is, the permutations on set $\{1,\ldots,n\}$. If $\pi \in S_n$ is a permutation, then $\mathrm{sgn}(\pi)$ stands for the *sign* of π. Any permutation in S_n can be represented as a product of *cycles*:

$$\pi = (1, \pi(1), \ldots, \pi^{l_1-1}(1)) \ldots (i_r, \pi(i_r), \ldots, \pi^{l_r-1}(i_r)), \quad (5)$$

and representation (5) is unique when ignoring the order of the cycles and cyclic shifts in each cycle. We say that permutation (5) has *cycle structure* $(l_1,\ldots,l_r)$, where l_1, ..., l_r are the lengths of the cycles. If $l_1 \leq \ldots \leq l_r$ and $l_1 + \ldots + l_r = n$, the subset of permutations in S_n having cycle structure $(l_1,\ldots,l_r)$ is denoted by $C_n(l_1,\ldots,l_r)$. The cardinality of $C_n(l_1,\ldots,l_r)$ is denoted by $N_n(l_1,\ldots,l_r)$.

Example 1. In S_3 there are 6 permutations: The identity $(1)(2)(3)$, three transpositions $(1)(23)$, $(2)(13)$, $(3)(12)$, and two three-cycles (123) and (132). Hence $C_3(1,1,1) = \{(1)(2)(3)\}$, $C_3(1,2) = \{(1)(23),(2)(13),(3)(12)\}$, and $C_3(3) = \{(123),(132)\}$. Consequently, $N_3(1,1,1) = 1$, $N_3(1,2) = 3$, and $N_3(3) = 2$. Exploiting the structure of S_4 in the same way, one can see that $N_4(1,1,1,1) = 1$, $N_4(1,1,2) = 6$, $N_4(1,3) = 8$, $N_4(2,2) = 3$, and $N_4(4) = 6$.

If a is an element of a ring, and $n \in \mathbb{Z}$, then $\mathbb{Z}$-module multiplication of a by n is defined as $n \cdot a = a + \ldots + a$ (n times), if $n \geq 0$, and $n \cdot a = -(a + \ldots + a)$ ($-n$ times), if $n < 0$. Hence a $\mathbb{Z}$-multiplication can always be interpreted as a repeated ring addition, but for example, in $\mathbb{C}[x]$, a $\mathbb{Z}$-module multiplication has a more natural implementation by multiplying each coefficient by the given integer. By a $\mathbb{Z}$-module division we understand recovering b from $a = n \cdot b$, where $n \in \mathbb{Z}$.

3 Positivity Criterion

The following fact is well-known, and sometimes it is used as an alternative definition of positivity.

Theorem 2. *A self-adjoint matrix A is positive if and only if all its eigenvalues are nonnegative.*

Remark 1. All eigenvalues of a self-adjoint matrix are real, see [2], for example.

To see that Theorem 2 is equivalent to the previous definition, it suffices just to notice that since A is self-adjoint, there are eigenvectors $\boldsymbol{x}_1$, ..., $\boldsymbol{x}_n$ of A that form an orthonormal basis of H_n [2], and then any vector can be represented as $\boldsymbol{x} = c_1\boldsymbol{x}_1 + \ldots + c_n\boldsymbol{x}_n$. It is then clear that

$$\langle \boldsymbol{x} \mid A\boldsymbol{x} \rangle = \lambda_1 |c_1|^2 + \ldots + \lambda_n |c_n|^2 . \tag{6}$$

Hence if $\lambda_i \geq 0$, then certainly A is positive, but if any $\lambda_i < 0$, then vector $\boldsymbol{x}_i$ substituted into Equation 6 shows that A is not positive.

According to Theorem 2, to determine whether a given matrix A is positive, we should determine if its eigenvalues are all nonnegative. For this purpose, one should discover the characteristic polynomial of A. There are previously known polynomial time algorithms for this purpose (see [4] and [5] and the references therein), but here we present a another method, which is rather based on linear algebra than combinatorics.

We will next present a modest lemma giving a part of the desired criterion, but for that purpose, we must first introduce the following definition: We define *an alternating polynomial* as a monic polynomial over real numbers having coefficients with alternating signs. In symbols, an alternating polynomial $p(x)$ is a polynomial

$$p(x) = x^n - a_1 x^{n-1} + a_2 x^{n-2} - \ldots + (-1)^{n-1} a_{n-1} x + (-1)^n a_n, \tag{7}$$

where each $a_i \geq 0$.

Lemma 1. *The real roots of an alternating polynomial are nonnegative.*

Proof. Assume the contrary: Alternating polynomial (7) has a root $-\lambda$, where $\lambda > 0$. Then

$$\begin{aligned} 0 &= p(-\lambda) \\ &= (-\lambda)^n - a_1(-\lambda)^{n-1} + \ldots + (-1)^{n-1} a_{n-1}(-\lambda) + (-1)^n a_n \\ &= (-1)^n(\lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \ldots + a_{n-1}\lambda + a_n). \end{aligned}$$

Since each a_i is nonnegative, the last expression in parenthesis is at least $\lambda^n > 0$. In particular, the last line is nonzero, a contradiction. □

The above lemma provides easily a positivity criterion, stated in the following theorem.

Theorem 3. *A self-adjoint matrix A is positive if and only if its characteristic polynomial is alternating.*

Proof. As discussed above, the characteristic polynomial of a self-adjoint matrix has real coefficients and real roots. If the characteristic polynomial is alternating, the eigenvalues of A are nonnegative by the previous lemma, and hence A is positive.

Assume then that A is positive. Then each eigenvalue of A is nonnegative, and by (4), each symmetric basic function on $\lambda_1, \ldots, \lambda_n$ is nonnegative, too. But then the characteristic polynomial of A is alternating, as Equation 3 shows. □

To proceed towards Theorem 1, we will present the coefficients of the characteristic polynomial in an alternate way. For that purpose, we refer to [3] for the following well-known representation:

$$p_A(\lambda) = \sum_{k=0}^{n} \lambda^{n-k}(-1)^k \sum_{\substack{I\subseteq\{1,\ldots,n\}\\|I|=k}} \det(A[I]). \tag{8}$$

Equation (8) says that the kth coefficient (counting from the highest power of λ) of the characteristic polynomial of A equals to $(-1)^k$ times the sum of the minors of A of size $k \times k$. Since anyway we have

$$\begin{aligned} p_A(\lambda) &= (\lambda - \lambda_1) \cdot \ldots \cdot (\lambda - \lambda_n) \\ &= V_0\lambda^n - V_1\lambda^{n-1} + V_2\lambda^{n-2} + \ldots + (-1)^n V_n, \end{aligned}$$

equation (8) implies that

$$V_k(\lambda_1, \ldots, \lambda_n) = \sum_{\substack{I\subseteq\{1,\ldots,n\}\\|I|=k}} \det(A[I]), \tag{9}$$

and hence Theorem 3 can be rephrased as

Theorem 4. *A matrix A is positive if and only if*

$$\sum_{\substack{I\subseteq\{1,\ldots,n\}\\|I|=k}} \det(A[I]) \geq 0$$

for each $k \in \{0, 1 \ldots, n\}$.

Notice that Theorem 4 seems like a weaker version of the criterion mentioned in the introduction: to guarantee the positivity of matrix A it is not necessary to have each individual minor $\det(A[I])$ nonnegative, but it is sufficient to have the sum of order k minors nonnegative for each $k \in \{1, \ldots, n\}$.

4 Connecting Characteristic Polynomial to the Power Traces

In this section, we will prove Theorem 1 by discovering alternative forms for sums (9) in terms of the traces of the powers of matrix A (see Theorem 5 below). By

Equation (8), these sums are equal to the coefficients of the characteristic polynomial of A (ignoring the alternating sign), which means that evaluating sums (9) is equivalent to finding the characteristic polynomial of A.

To begin with, we introduce the following simple lemma.

Lemma 2. *Let $A^{(1)}$, ..., $A^{(k)}$ be $n \times n$ matrices and $\pi \in S_k$ be a k-cycle $(1, \ldots, k)$. Then*

$$\sum_{i_1=1}^{n} \ldots \sum_{i_k=1}^{n} \prod_{j=1}^{k} A^{(j)}_{i_j, i_{\pi(j)}} = \mathrm{Tr}(A^{(1)} \cdot \ldots \cdot A^{(k)}). \tag{10}$$

Proof. For $k = 1$ the left hand size of (10) becomes

$$\sum_{i_1=1}^{n} A^{(1)}_{i_1, i_1} = \mathrm{Tr}(A^{(1)}),$$

as claimed. Assume then that the claim is true for $k - 1$. Then

$$\begin{aligned}
& \sum_{i_1=1}^{n} \ldots \sum_{i_k=1}^{n} \prod_{j=1}^{k} A^{(j)}_{i_j, i_{\pi(j)}} \\
= & \sum_{i_1=1}^{n} \ldots \sum_{i_{k-1}=1}^{n} \prod_{j=1}^{k-2} A^{(j)}_{i_j, i_{\pi(j)}} \sum_{i_k=1}^{n} A^{(k-1)}_{i_{k-1}, i_k} A^{(k)}_{i_k, i_1} \\
= & \sum_{i_1=1}^{n} \ldots \sum_{i_{k-1}=1}^{n} \prod_{j=1}^{k-2} A^{(j)}_{i_j, i_{\pi(j)}} (A^{(k-1)} A^{(k)})_{i_{k-1}, i_1} \\
= & \mathrm{Tr}(A^{(1)} \cdot \ldots \cdot A^{(k-2)} \cdot (A^{(k-1)} A^{(k)})),
\end{aligned}$$

which is the claim. The last equality is a result of applying the induction hypothesis on matrices $A^{(1)}$, ..., $A^{(k-2)}$, and $A^{(k-1)} A^{(k)}$. □

Remark 2. We will actually use only a special case of the previous lemma with $A^{(1)} = \ldots = A^{(k)} = A$. Equation (10) becomes then

$$\sum_{i_1=1}^{n} \ldots \sum_{i_k=1}^{n} A_{i_1, i_2} \cdot \ldots \cdot A_{i_{k-1}, i_k} A_{i_k, i_1} = \mathrm{Tr}(A^k). \tag{11}$$

In this case, we can denote $C = (1, \ldots, k)$ and write (11) in a form

$$\sum_{\substack{i_j=1 \\ j \in C}}^{n} \prod_{j \in C} A_{i_j, i_{\pi(j)}} = \mathrm{Tr}(A^k) \tag{12}$$

to shorten the notations. We will use form (12) later.

In the light of Theorem 4, Theorem 1 follows immediately from the theorem below.

Theorem 5. *Let A be an $n \times n$-matrix and $k \leq n$. Then*

$$\begin{aligned} & \sum_{|I|=k} \det(A[I]) \\ = \quad & \frac{1}{k!} \sum_{r=1}^{k} (-1)^{k-r} \sum_{\substack{l_1+\ldots+l_r=k \\ 1 \leq l_1 \leq \ldots \leq l_r \leq k}} N_k(l_1, \ldots, l_r) \operatorname{Tr}(A^{l_1}) \cdot \ldots \cdot \operatorname{Tr}(A^{l_r}). \end{aligned}$$

Proof. We introduce a slight variation to the notation used before: If i_1, ..., $i_k \in \{1, 2, \ldots, n\}$, then

$$A[i_1, i_2, \ldots, i_k] = \begin{pmatrix} A_{i_1,i_1} & A_{i_1,i_2} & \ldots & A_{i_1,i_k} \\ A_{i_2,i_1} & A_{i_2,i_2} & \ldots & A_{i_2,i_k} \\ \ldots & \ldots & \ldots & \ldots \\ A_{i_k,i_1} & A_{i_k,i_2} & \ldots & A_{i_k,i_k} \end{pmatrix}, \tag{13}$$

where $A_{i,j}$ are the matrix elements of matrix A. The difference between the previous notation $A[\{i_1, \ldots, i_k\}]$ is that now the order of numbers i_1, ..., i_k is significant, and a number can occur many times. On the other hand, if $i_1 < i_2 < \ldots < i_k$, then clearly $A[i_1, \ldots, i_k] = A[\{i_1, \ldots, i_k\}]$.

It is important to notice that if two indices are equal in (13), then there are two identical rows (and also two identical columns) in $A[i_1, i_2, \ldots, i_k]$, and consequently $\det(A[i_1, i_2, \ldots, i_k]) = 0$. Moreover, any permutation of 1, ..., k does not affect the value of $\det(A[i_1, i_2, \ldots, i_k])$. To see this, consider arbitrary transposition, say swapping indices i_r and i_s. The effect of this transposition is simply that one of swapping the rth and sth rows and the corresponding columns. But since both swappings result in a sign change, the determinant remains unchanged. Since all permutations can be expressed as a product of transpositions, we conclude that $\det(A[i_1, i_2, \ldots, i_k])$ is invariant under permuting the subindices 1, ..., k.

To get the desired result, we will evaluate the sum

$$\sum_{i_1=1}^{n} \ldots \sum_{i_k=1}^{n} \det(A[i_1, \ldots, i_k]) \tag{14}$$

in two different ways.

As discussed before, the terms having $i_r = i_s$ for some $r \neq s$ are zero. To avoid equalities, i_1 can be selected in n ways, i_2 in $n-1$ ways, and so on. Thus there are $n(n-1) \cdot \ldots \cdot (n-k+1) = k! \binom{n}{k}$ nontrivial summands in (14). We can then classify all the nontrivial terms into classes where sequences $(i_1, \ldots, i_k)$ differ from each other only by a permutation of subindices 1, ..., k. In each class, there is a unique representative with property $i_1 < \ldots < i_k$, so we get (recall that $\det(A[i_1, i_2, \ldots, i_k])$ is invariant under the permutation of subindices)

$$
\begin{aligned}
&\sum_{i_1=1}^{n} \dots \sum_{i_k=1}^{n} \det(A[i_1, \dots, i_k]) \\
= &\sum_{1 \le i_1 < \dots < i_k \le n} \sum_{\pi \in S_k} \det(A[i_{\pi(1)}, \dots, i_{\pi(k)}]) \\
= \; & k! \sum_{1 \le i_1 < \dots < i_k \le n} \det(A[i_1, \dots, i_k]) \\
= \; & k! \sum_{|I|=k} \det(A[I]).
\end{aligned}
$$

We will next show that

$$
\begin{aligned}
&\sum_{i_1=1}^{n} \dots \sum_{i_k=1}^{n} \det(A[i_1, \dots, i_k]) \\
= \; & \sum_{r=1}^{k} (-1)^{k-r} \sum_{\substack{l_1+\dots+l_r=k \\ 1 \le l_1 \le \dots \le l_r \le k}} N_k(l_1, \dots, l_r) \operatorname{Tr}(A^{l_1}) \cdot \dots \cdot \operatorname{Tr}(A^{l_r}), \quad (15)
\end{aligned}
$$

and Theorem 5 will follow immediately. To prove (15), we apply the definition of the determinant to the left hand side to get

$$
\begin{aligned}
&\sum_{i_1=1}^{n} \dots \sum_{i_k=1}^{n} \det(A[i_1, \dots, i_k]) \\
= \; & \sum_{i_1=1}^{n} \dots \sum_{i_k=1}^{n} \sum_{\pi \in S_k} \operatorname{sgn}(\pi) \prod_{j=1}^{k} A_{i_j, i_{\pi(j)}}. \quad (16)
\end{aligned}
$$

Consider then the product

$$
\operatorname{sgn}(\pi) \prod_{j=1}^{k} A_{i_j, i_{\pi(j)}} \quad (17)
$$

when $\pi \in S_k$ is a fixed permutation in $C_k(l_1, \dots, l_r)$, that is, π has cycle structure $(l_1, \dots, l_r)$. Clearly $\operatorname{sgn}(\pi) = (-1)^{l_1 - 1 + \dots + l_r - 1} = (-1)^{k-r}$, and using notations C_1, …, C_r as in Remark 2 for the cycles of π, we can write (17) as

$$
(-1)^{k-r} \prod_{m=1}^{r} \prod_{j \in C_m} A_{i_j, i_{\pi(j)}} \quad (18)
$$

Now the inmost sum of (16) consists of terms of form (18). Computing the outer sums over each such term, we have

$$
\begin{aligned}
&(-1)^{k-r} \sum_{i_1=1}^{n} \dots \sum_{i_k=1}^{n} \prod_{m=1}^{r} \prod_{j \in C_m} A_{i_j, i_{\pi(j)}} \\
= \; & (-1)^{k-r} \sum_{\substack{i_j=1 \\ j \in C_1}}^{n} \prod_{j \in C_1} A_{i_j, i_{\pi(j)}} \dots \sum_{\substack{i_j=1 \\ j \in C_m}}^{n} \prod_{j \in C_m} A_{i_j, i_{\pi(j)}} \\
= \; & (-1)^{k-r} \operatorname{Tr}(A^{l_1}) \cdot \dots \cdot \operatorname{Tr}(A^{l_r}) \quad (19)
\end{aligned}
$$

by lemma 2. This means that in sum (16) each permutation $\pi \in C_k(l_1, \ldots, l_r)$ contributes term (19). Then

$$\begin{aligned} & \sum_{i_1=1}^{n} \ldots \sum_{i_k=1}^{n} \det(A[i_1, \ldots, i_k]) \\ = & \sum_{r=1}^{k} \sum_{\substack{l_1+\ldots+l_r=k \\ 1\le l_1\le\ldots\le l_r\le k \\ \pi\in C_k(l_1,\ldots,l_r)}} (-1)^{k-r} \operatorname{Tr}(A^{l_1}) \cdot \ldots \cdot \operatorname{Tr}(A^{l_r}) \qquad (20) \\ = & \sum_{r=1}^{k} (-1)^{k-r} \sum_{\substack{l_1+\ldots+l_r=k \\ 1\le l_1\le\ldots\le l_r\le k}} N_k(l_1, \ldots, l_r) \operatorname{Tr}(A^{l_1}) \cdot \ldots \cdot \operatorname{Tr}(A^{l_r}), \end{aligned}$$

and the claim follows. □

5 Computational Aspects

We will now explain how to compute the sum

$$\sum_{r=1}^{k} (-1)^{k-r} \sum_{\substack{l_1+\ldots+l_r=k \\ 1\le l_1\le\ldots\le l_r\le k}} N_k(l_1, \ldots, l_r) \operatorname{Tr}(A^{l_1}) \cdot \ldots \cdot \operatorname{Tr}(A^{l_r}) \qquad (21)$$

occurring in Theorem 1 efficiently. It is clearly enough to explain how sum

$$S(r,k) = \sum_{\substack{l_1+\ldots+l_r=k \\ 1\le l_1\le\ldots\le l_r\le k}} N_k(l_1, \ldots, l_r) \operatorname{Tr}(A^{l_1}) \cdot \ldots \cdot \operatorname{Tr}(A^{l_r}) \qquad (22)$$

can be computed efficiently, and for that purpose, we will first present the following simple lemma.

Lemma 3. *Let l_1, ..., l_r be positive integers such that $l_1 + \ldots + l_r = k$ and $l_1 \le \ldots \le l_r$. Then*

$$N_k(l_1, \ldots, l_r) = \begin{cases} \frac{k!}{l^r r!} & \text{if } l_1 = \ldots = l_r = l = k/r, \\ \binom{k}{m} N_m(l_1, \ldots, l_{r'}) N_{k-m}(l_{r'+1}, \ldots, l_r) & \text{if } l_1 + \ldots + l_{r'} = m \text{ and } l_{r'} < l_{r'+1}. \end{cases}$$

Proof. To get the first claim, notice that all $k!$ permutations of numbers $1, \ldots, k$ can be obtained exactly once by starting from all representations (5) of all permutations consisting of r l-cycles, then arranging the cycles in all $r!$ ways, and finally applying cyclic shifts to these cycles in all l^r ways. In symbols: $k! = N_k(l, \ldots, l) r! l^r$, and the first claim follows.

For the second claim, we notice that there are $\binom{k}{m}$ ways to choose a set I of m elements of $\{1, \ldots, k\}$, and then one has $N_m(l_1, \ldots, l_{r'})$ possibilities for

having r' cycles in I, and $N_{k-m}(l_{r'+1}, \ldots, l_r)$ for having the rest of the cycles in the complement of I. The claim follows immediately, because now $\{l_1, \ldots, l_{r'}\} \cap \{l_{r'+1}, \ldots, l_r\} = \emptyset$. □

Lemma 3 provides the recursion we can use for computing (22). Generalizing (22), we define

$$S(r,k,M) = \sum_{\substack{l_1+\ldots+l_r=k \\ 1\le l_1\le\ldots\le l_r<M}} N_k(l_1,\ldots,l_r)\,\mathrm{Tr}(A^{l_1})\cdot\ldots\cdot\mathrm{Tr}(A^{l_r}), \tag{23}$$

but it turns out that a further generalization is useful. For any $e \in \{1, \ldots, r\}$ we define also

$$S(r,k,M,e) = \sum_{\substack{l_1+\ldots+l_r=k \\ 1\le l_1\le\ldots\le l_r<M \\ l_{r-e+1}=\ldots=l_r}} N_k(l_1,\ldots,l_r)\,\mathrm{Tr}(A^{l_1})\cdot\ldots\cdot\mathrm{Tr}(A^{l_r}) \tag{24}$$

to be a restricted version of (23), with e last variables l_{r-e+1}, …, l_r having the same value. Then we have clearly $S(r,k,M) = S(r,k,M,1)$ and $S(r,k) = S(r,k,k+1,1)$.

Some boundary conditions are very easy to obtain. For example, by Lemma 3 it is clear that

$$S(1,k,M,e) = \begin{cases} N_k(k)\,\mathrm{Tr}(A^k) = (k-1)!\,\mathrm{Tr}(A^k), & \text{if } k < M, \\ 0 & \text{otherwise.} \end{cases} \tag{25}$$

Moreover, if all variables are required to be equal, we have again a very easy case. In fact,

$$\begin{aligned} S(r,k,M,r) &= \sum_{\substack{rl=k \\ 1\le l<M}} N(l,\ldots,l)\,\mathrm{Tr}(A^l)^r \\ &= \begin{cases} \frac{k!}{l^r r!}\,\mathrm{Tr}(A^l)^r, & \text{if } r \mid k \text{ and } l = k/r < M, \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \tag{26}$$

Since each variable l_1, …, l_r is at least 1, we have also evidently

$$S(r,k,M,e) = \begin{cases} N_k(1,\ldots,1)\,\mathrm{Tr}(A)^r = \mathrm{Tr}(A)^r, & \text{if } k = r, \\ 0 & \text{if } r > k. \end{cases}$$

We will now find a recursion for $S(r,k,M,e)$, and for that purpose, we separate the terms in the defining sum according to whether $l_{r-e} = l_{r-e+1}$ or not.

$$\begin{aligned} &S(r,k,M,e) \\ = &\sum_{\substack{l_1+\ldots+l_r=k \\ 1\le l_1\le\ldots\le l_r<M \\ l_{r-e}<l_{r-e+1}=\ldots=l_r}} N_k(l_1,\ldots,l_r)\,\mathrm{Tr}(A^{l_1})\cdot\ldots\cdot\mathrm{Tr}(A^{l_r}) + S(r,k,M,e+1) \end{aligned}$$

Denoting $l = l_{r-e+1} = \ldots = l_r$ and using the recursion of Lemma 3 we get

$$
\begin{aligned}
& S(r,k,M,e) \\
= & \sum_{l<M} \sum_{\substack{l_1+\ldots+l_{r-e}=k-le \\ 1\le l_1\le\ldots\le l_r<l}} \binom{k}{le} N_{k-le}(l_1,\ldots,l_{r-e}) \\
\times & \operatorname{Tr}(A^{l_1})\cdot\ldots\cdot\operatorname{Tr}(A^{l_{r-e}}) N_{le}(l,\ldots,l)\operatorname{Tr}(A^l)\cdot\ldots\cdot\operatorname{Tr}(A^l) \\
+ & S(r,k,M,e+1)
\end{aligned}
$$

If $k-le$ is less than $r-e$, the inner sum is empty, so we can write

$$
\begin{aligned}
& S(r,k,M,e) \\
= & \sum_{l=1}^{\min(\lfloor\frac{k-r}{e}\rfloor+1,M-1)} \sum_{\substack{l_1+\ldots+l_{r-e}=k-le \\ 1\le l_1\le\ldots\le l_r<l}} \binom{k}{le} N_{k-le}(l_1,\ldots,l_{r-e}) \\
\times & \operatorname{Tr}(A^{l_1})\cdot\ldots\cdot\operatorname{Tr}(A^{l_{r-e}})\frac{(le)!}{l^e e!}\operatorname{Tr}(A^l)^e \\
+ & S(r,k,M,e+1) \\
= & \sum_{l=1}^{\min(\lfloor\frac{k-r}{e}\rfloor+1,M-1)} \binom{k}{le}\frac{(le)!}{l^e e!}\operatorname{Tr}(A^l)^e S(r-e,k-le,l,1) \\
+ & S(r,k,M,e+1)
\end{aligned}
$$

The recursion

$$
\begin{aligned}
S(r,k,M,e) &= \sum_{l=1}^{\min\{\lfloor\frac{k-r}{e}\rfloor+1,M-1\}} \binom{k}{el}\frac{(le)!}{l^e e!}\operatorname{Tr}(A^l)^e S(r-e,k-le,l,1) \\
&+ S(r,k,M,e+1) \qquad (27)
\end{aligned}
$$

obtained gives rise to the following procedure:

Step 1. Compute and store values $\operatorname{Tr}(A)$, ..., $\operatorname{Tr}(A^n)$. A matrix multiplication of an $n\times n$ matrix can be done with $\mathcal{O}(n^3)$ ring multiplications and additions, and A^n can be computed with $\mathcal{O}(\log n)$ matrix multiplications by repeated squaring. Hence this first step can be done with $\mathcal{O}(n^3\log n)$ ring multiplications and additions.

Step 2. Compute and store $\operatorname{Tr}(A^l)^e$ for each $1\le e,l\le n$. Basing on the values stored in the first step, this step can be done by using $\mathcal{O}(n^2\log n)$ ring multiplications.

Step 3.0. For each $M\in[r+1,k+1]$ and $r\in[1,k]$ compute

$$
S(r,k,M,r)=\begin{cases}\frac{k!}{l^r r!}\operatorname{Tr}(A^l)^r, & \text{if } r\mid k \text{ and } l=k/r<M,\\ 0 & \text{otherwise.}\end{cases}
$$

For Step 3.j assume that values $S(r,k,M,r-i)$ are computed for each $M\in[r+1,k+1]$, $r\in[i+1,k]$, and $i\in[0,1,\ldots,j-1]$. After step 3.0 the assumption

holds for $i = 0$, and step 3.j implies that the assumption holds for $i = j$ after step 3.j.

For j=1 to k−1

Step 3.j. Compute $S(r, k, M, r-j)$ for each $M \in [r+1, k+1]$ and $r \in [j+1, k]$ by using recursion

$$\begin{aligned} & S(r,k,M,r-j) \\ = & \sum_{l=1}^{\min(\lfloor \frac{k-r}{r-j} \rfloor + 1, M-1)} C(r,k,l,j)\,\mathrm{Tr}(A^l)^{r-j} S(j, k-(r-j)l, l, 1) \\ + & \; S(r,k,M,r-j+1), \end{aligned} \tag{28}$$

where $C(r,k,l,j) = \binom{k}{(r-j)l} \frac{((r-j)l)!}{l^{r-j}(r-j)!}$ (cf. (27)). By the assumption, values $S(j, k-(r-j)l, l, 1)$ and $S(r,k,M,r-j+1)$ are computed in the previous steps.

Step 4. Compute

$$C(k) = \sum_{r=1}^{k} (-1)^{r-k} S(r,k,k+1,1). \tag{29}$$

As

$$S(r,k,k+1,1) = \sum_{\substack{l_1+\ldots+l_r=k \\ 1 \le l_1 \le \ldots \le l_r < k+1}} N(l_1,\ldots,l_r)\,\mathrm{Tr}(A^{l_1}) \cdot \ldots \cdot \mathrm{Tr}(A^{l_r}),$$

we see that (29) equals to sum (21), that is,

$$\begin{aligned} C(k) &= \sum_{i_1=1}^{n} \ldots \sum_{i_k=1}^{n} \det(A[i_1,\ldots,i_k]) \\ &= k! \sum_{\substack{I \subseteq \{1,\ldots,n\} \\ |I|=k}} \det(A[I]) = k!(-1)^k c_k, \end{aligned}$$

where c_k is the kth coefficient of the characteristic polynomial of A ($c_0 = 1$ being the leading coefficient).

Now we estimate how many algebraic operations are needed for all steps $3.j$, $0 \le j \le k-1$. Recursion (28) reveals that one needs at most $\frac{k-r}{r-j} + 1 = \frac{k-j}{r-j}$ ring multiplications, additions, and $\mathbb{Z}$-module multiplications for fixed r and M. Clearly this holds also for $j = 0$. Since there are $k-r+1$ possibilities for M, one needs at most

$$\sum_{r=j+1}^{k} (k-r+1)\frac{k-j}{r-j}$$

aforementioned operations for step $3.j$. Summing over each j, we see that the

number of required algebraic operations is at most

$$\sum_{j=0}^{k-1}\sum_{r=j+1}^{k}(k-r+1)\frac{k-j}{r-j} \leq \sum_{j=0}^{k-1}\sum_{r=j+1}^{k}k^2\frac{1}{r-j} = k^2\sum_{j=0}^{k-1}\sum_{r=1}^{k-j}\frac{1}{r}$$
$$\leq \quad k^2\sum_{j=0}^{k-1}(\log(k-j)+1) \leq k^2\sum_{j=0}^{k-1}(\log k+1) \leq 2k^3\log k,$$

when $k \geq 3$.

Step 4 can be done with $\mathcal{O}(k)$ algebraic operations, hence we need altogether $\mathcal{O}(k^3 \log k)$ algebraic operations for computing $C(k)$ and the coefficient $c_k = \frac{(-1)^k}{k!}C(k)$. Now that determinant is given by $(-1)^n c_n$, the claimed goal of the algorithm is reached.

We conclude by noticing that even though computing $c(k) = \frac{(-1)^k}{k!}C(k)$ requires division by $k!$, the procedure can be still implemented over some rings over finite characteristics, such as polynomial rings over prime fields $\mathbb{Z}_p$, for instance. In such cases, the polynomials have to be interpreted as polynomials over $\mathbb{Z}$ when performing steps 1–4. After that, the outcoming polynomial must be divided by $k!$, and its coefficients can then be projected to the prime field $\mathbb{Z}_p$ to get the final outcome.

References

[1] G. H. Hardy and S. Ramanujan: *Asymptotic Formulae in Combinatory Analysis*. Proc. London Math. Soc. 17, 75–115 (1918).

[2] M. Hirvensalo: *Quantum Computing*, 2nd edition, Springer (2004).

[3] P. Lancaster and M. Tismenetsky: *The theory of matrices*. Academic Press (1985).

[4] M. Mahajan and V. Vinay: *Determinant: Old Algorithms, New Insights*. SIAM Journal on Discrete Mathematics 12(4), 474–490 (1999).

[5] G. Rote: *Division-free algorithms for the determinant and the Pfaffian: algebraic and combinatorial approaches*. In: Helmut Alt (ed.) : *Computational Discrete Mathematics*. Springer LNCS 2122, pp. 119–135 (2001).

Turku
Centre *for*
Computer
Science

Lemminkäisenkatu 14 A, 20520 Turku, Finland | www.tucs.fi

University of Turku
- Department of Information Technology
- Department of Mathematics

Åbo Akademi University
- Department of Computer Science
- Institute for Advanced Management Systems Research

Turku School of Economics and Business Administration
- Institute of Information Systems Sciences

ISBN 952-12-1649-2
ISSN 1239-1891