On Computing the Determinant and Smith Form of an Integer Matrix

Wayne Eberly*
Department of Computer Science
University of Calgary
Calgary, AB, Canada, T2N 1N4
email: eberly@cpsc.ucalgary.ca

Mark Giesbrecht*
Department of Computer Science
University of Western Ontario
London, ON, Canada, N6A 5B7
email: mwg@csd.uwo.ca

Gilles Villard
CNRS
LMC-IMAG Grenoble, France
email: Gilles.Villard@imag.fr

Abstract

A probabilistic algorithm is presented to find the determinant of a nonsingular, integer matrix. For a matrix $A \in \mathbb{Z}^{n\times n}$ the algorithm requires $O(n^{3.5}(\log n)^{4.5})$ bit operations (assuming for now that entries in A have constant size) using standard matrix and integer arithmetic. Using asymptotically fast matrix arithmetic, a variant is described which requires $O(n^{2+\theta/2} \cdot \log^2 n \log\log n)$ bit operations, where two $n \times n$ matrices can be multiplied with $O(n^\theta)$ operations. The determinant is found by computing the Smith form of the integer matrix, an extremely useful canonical form in itself. Our algorithm is probabilistic of the Monte Carlo type. That is, it assumes a source of random bits and on any invocation of the algorithm there is a small probability of error.

1 Introduction

One of the most fundamental invariants of a square matrix is the determinant. Applications for computing the determinant of a matrix are numerous. For integer matrices alone they include computational number theory [4], computational group theory [9], and computational geometry [2, 3]. In this paper we present a new algorithm for the determinant which is faster than any previously known. For a matrix $A \in \mathbb{Z}^{n\times n}$ this algorithm requires

$$O(n^3(\log n + \log\|A\|)^2\sqrt{\log|\det A|} \cdot \log^2 n)$$

bit operations using standard matrix arithmetic, where $\|A\| = \max_{ij}|A_{ij}|$. Since, by Hadamard's bound, $|\det A| = O(n(\log n + \log\|A\|))$, this cost is at worst $O(n^{3.5}(\log n + \log\|A\|)^{2.5} \cdot \log^2 n)$ bit operations, though the sensitivity to the size of the determinant can be beneficial.

We will consider only the exact computation of the determinant of an integer matrix. Computing cost will be counted in bit operations (and hence will reflect both the number of integer operations and the size of integers involved). The fastest previously known method for computing the determinant of an integer matrix uses the Chinese remainder algorithm and matrix arithmetic modulo primes. For a matrix $A \in \mathbb{Z}^{n\times n}$, this requires $O(n^4(\log n + \log\|A\|)(\log n + \log\log\|A\|) + n^2\log^2\|A\|)$ bit operations, and is deterministic (see Abbott et al. [1]). The best known Monte Carlo algorithm requires $O(n^3\log|\det A| \cdot (\log n + \log\log\|A\|) + \log^2|\det A|)$ bit operations (see below).

It is well known that every nonsingular integer matrix is equivalent to a matrix in Smith canonical form. That is, there exist unimodular $X, Y \in \mathbb{Z}^{n\times n}$ (i.e., $\det X$, $\det Y = \pm 1$) such that

$$S = XAY = \begin{bmatrix} s_n & & 0 \\ & \ddots & \\ 0 & & s_1 \end{bmatrix} \in \mathbb{Z}^{n\times n} \qquad (1)$$

and $s_i \mid s_{i+1}$ for $1 \le i \le n$. S is called the *Smith normal form* of A and $s_1, \ldots, s_n \in \mathbb{Z}_{>0}$ the *invariant factors* of A. Once we have the Smith form, the determinant of A is $s_1 s_2 \cdots s_n$ (and this is how our algorithm for the determinant proceeds). The Smith normal form also has many applications in computational number theory and group theory [4] as well as computations in homology theory (e.g., Dumas & Saunders [7]). The best known deterministic algorithm to compute the Smith form of an integer matrix is by Storjo-

*Research was supported in part by Natural Sciences and Engineering Research Council of Canada

To Appear: 41st IEEE Annual Symposium on Foundations of Computer Science, FOCS'2000, November 2000.

hann [18] and requires $O(n^4 \log\|A\| + n^3 \log^2\|A\|)$ bit operations (ignoring poly-logarithmic factors). When the matrix is sparse the algorithms of Giesbrecht [12, 13] do substantially better, but are comparable when the matrix is dense.

In this paper we present an algorithm which requires $O(n^3(\log n + \log\|A\|)^2\sqrt{\log|\det A|}\cdot\log^2 n)$ bit operations for dense matrices, or $O(n^{3.5}(\log n+\log\|A\|)^{2.5}\cdot\log^2 n)$ bit operations in the worst case (i.e., independent of the magnitude of the determinant), using standard matrix arithmetic.

Using similar methods and asymptotically fast matrix arithmetic, we derive an asymptotically faster algorithm for computing the Smith form and determinant. Suppose that two $n\times n$ over an arbitrary ring R can be multiplied with $O(n^\theta)$ operations in R. Using standard matrix arithmetic gives $\theta = 3$, while the best known algorithm of Coppersmith & Winograd [5] allows $\theta = 2.376$. Our algorithm for the Smith form and determinant then requires

$$O\Big(n^{2+\theta/2}\cdot(\log n+\log\|A\|)^{3/2}(\log n)^{1/2} \cdot(\log\log n+\log\log\|A\|)\Big)$$

bit operations.

In Section 6 we examine the cost of our algorithm when computing the determinant and Smith form of a "random" integer matrix. In particular, we show that if the entries a matrix are chosen uniformly and randomly from an interval $\Lambda = \{a, a+1,\ldots,a+\lambda-1\}$ for any integer $a\in\mathbb{Z}$ and $\lambda\in\mathbb{Z}_{\geq 2}$, the expected number of non-trivial (i.e., not equal to ± 1) invariant factors is $O(\log_\lambda n)$. This is consistent with previous experimental evidence (and, perhaps, "folklore") that the number of invariant factors is small but is, to our knowledge, the first proof of this sort of bound. In this case, our algorithm for the Smith form and determinant will require $O(n^3(\log n+\log\|A\|)^2\cdot\log(n)\log_\lambda(n))$ bit operations.

1.1 An Overview of the Algorithm

The algorithm itself is relatively simple. It employs some analogous ideas for computation of the Frobenius form of a sparse matrix over an abstract field developed by Villard [20] and adapts them to Smith forms over the integers. The algorithm consists of three main ideas:

1. The largest invariant factor s_n of any matrix can be computed with an expected number of $O(n^3(\log n + \log\|A\|)^2)$ bit operations. This is done by solving the equation $Ax = b$ for random vectors $b\in\mathbb{Z}^{n\times 1}$. With appropriate random selection of b, it can be proven that the GCD of the denominators of the entries in the solution vector x is the largest invariant factor of A with high probability. We solve for x using p-adic lifting to obtain the desired cost. This is discussed in Section 2.
2. We show that we can capture the kth invariant factor of A by means of random perturbations of A. Let $B\in\mathbb{Z}^{n\times n}$ be an appropriately constructed random matrix with rank k. If the largest invariant factor of $A+B$ is σ_n, then with sufficiently high probability $\gcd(\sigma_n, s_n) = s_{n-k}$. This is discussed in Section 3.
3. The number of *distinct* invariant factors $s_1,\ldots,s_n$ of A is at most $\sqrt{\log|\det A|}$ or $O(\sqrt{n(\log n+\log\|A\|)})$. This will allow us to do what amounts to a binary search for the distinct invariant factors, which requires $O(\sqrt{\log|\det A|}\cdot\log n)$ computations of the kth invariant factor by means of (1) and (2) above. This leads to the total expected cost of our Smith form and determinant algorithms. The completed algorithm and its analysis is discussed in Section 4.

1.2 Previous Algorithms

The best known methods for computing the determinant are fraction free Gaussian elimination and homomorphic imaging. The latter (which is asymptotically faster) simply computes the determinant modulo a collection of small (typically word-sized) primes and reconstructs the integral determinant via the Chinese remainder theorem. By Hadamard's bound, the product of these primes must have $O(n(\log n+\log\|A\|))$ bits to ensure correctness. The algorithm obtained will require $O(n^4(\log n+\log\|A\|)(\log n + \log\log\|A\|)+n^2\log^2\|A\|)$ bit operations (see [1]).

By using asymptotically fast matrix multiplication we can obtain a better exponent, though practicality quickly vanishes. Using fast matrix arithmetic, the above homomorphic imaging scheme to compute the determinant requires $O(n^{\theta+1}\cdot(\log n+\log\|A\|)(\log n+\log\log\|A\|)+n^2(\log n + \log\|A\|)^2)$ bit operations.

Monte Carlo algorithms for the determinant have the advantage that their cost is sensitive to the size of the determinant. The idea is to compute the determinant modulo a collection of small random primes in sequence and build the integer determinant as the residues are obtained. Once the determinant remains stable modulo a small number of random primes, it is straightforward to bound the probability that the obtained determinant is the correct one. A Monte Carlo algorithm is easily obtained which requires $O(n^3\log|\det A|\cdot(\log n+\log\log\|A\|)+\log^2|\det A|)$ bit operations. On any input, on any invocation, the algorithm is correct with constant probability.

The algorithm of Kaltofen [14] computes the determinant of a $n\times n$ matrix over an arbitrary ring with $O(n^{3.5}\log n\log\log n)$ ring operations. A careful analysis of this algorithm for integer inputs reveals that this algorithm also requires $O(n^{3.5}(\log n + \log\|A\|)(\log n + \log\log\|A\|)(\log\log n+\log\log\log\|A\|))$ bit operations, using asymptotically fast integer arithmetic. Using asymptotically fast matrix arithmetic one should obtain further improvements [15].

2 Computing the Largest Invariant Factor of an Integer Matrix

Consider the following Monte Carlo type probabilistic algorithm to compute the largest invariant factor s_n of an integer matrix $A \in \mathbb{Z}^{n\times n}$. The method is essentially the same as that of Abbott et al. [1], which is derived from the algorithm of Pan [17] for polynomial matrices.

Algorithm: `LargestInvariantFactor`
Input: • $A \in \mathbb{Z}^{n\times n}$;
Output: • $s_n \in \mathbb{Z}$, the largest invariant factor of A;
(1) $M := 6 + 2n(\log_2 n + \log_2 \|A\|)$; $\mathcal{L} := \{0,\ldots,M-1\}$;
(2) $s_n^{(0)} := 1$;
(3) For k from 1 to 2 do
(4) Choose random $b^{(k)} \in \mathcal{L}^{n\times 1}$;
(5) Solve $Ax^{(k)} = b^{(k)}$ for $x^{(k)} \in \mathbb{Q}^{n\times 1}$;
(6) Let $t_n^{(k)} := \mathrm{lcm}(\mathrm{denom}(x_1^{(k)}),\ldots,\mathrm{denom}(x_n^{(k)}))$;
$s_n^{(k)} := \mathrm{lcm}(s_n^{(k-1)}, t_n^{(k)})$;
where $x_j^{(k)}$ is the jth entry in $x^{(j)}$ and $\mathrm{denom}(x_j^{(k)})$ is its denominator;
od;
(7) Return $s_n^{(2)}$.

THEOREM 2.1. *The algorithm* `LargestInvariantFactor` *always returns a factor of the largest invariant factor s_n of A. The algorithm returns s_n with probability at least $1/3$ on any input matrix A.*

PROOF. In [1] it is shown that in any iteration k, $t_n^{(k)}$ is a factor of s_n, whence $s_n^{(k)}$ is always a factor of s_n. It is also shown that for any prime $p \mid s_n^{(k)}$,

$$\mathrm{Prob}\{\mathrm{ord}_p t_n^{(k)} < \mathrm{ord}_p s_n\} \le \frac{1}{M}\cdot\left\lceil \frac{M}{p}\right\rceil.$$

By repeating this twice,

$$\begin{aligned}
\mathrm{Prob}\left\{s_n \ne s_n^{(2)}\right\} &\le \sum_{p\,|\,s_n} \mathrm{Prob}\left\{\mathrm{ord}_p(s_n) < \mathrm{ord}_p(s_n^{(2)})\right\}\\
&\le \sum_{\substack{p\,|\,s_n\\ p\ prime}} \mathrm{Prob}\left\{\begin{array}{c}\mathrm{ord}_p(s_n) < \mathrm{ord}_p(t_n^{(1)})\\ \wedge\,\mathrm{ord}_p(s_n) < \mathrm{ord}_p(t_n^{(2)})\end{array}\right\}\\
&= \sum_{\substack{p\,|\,s_n\\ p\ prime}} \left(\frac{1}{M}\left\lceil\frac{M}{p}\right\rceil\right)^2 < \sum_{\substack{p\,|\,s_n\\ p\ prime}} \left(\frac{1}{p}+\frac{1}{M}\right)^2\\
&< \sum_{\substack{p\,|\,s_n\\ p\ prime}} \frac{1}{p^2} + \frac{1}{M}\cdot\sum_{\substack{p\,|\,s_n\\ p\ prime}} \frac{1}{p} + \frac{1}{M^2}\cdot\sum_{\substack{p\,|\,s_n\\ p\ prime}} 1\\
&< \frac{1}{2} + \frac{1}{M}\left(\frac{1}{2}+\frac{1}{3}+\frac{1}{5}+\frac{\log_2(s_n)}{7}\right) + \frac{\log_2(s_n)}{M^2}\\
&< 2/3,
\end{aligned}$$

□

THEOREM 2.2. *The algorithm* `LargestInvariantFactor` *requires $O(n^3(\log n + \log\|A\|)^2)$ bit operations on input $A \in \mathbb{Z}^{n\times n}$.*

PROOF. The dominant cost in each iteration of the main loop is the solution of the linear system in step (5), which we do with p-adic lifting [10, 6, 19, 16]. Each iteration requires $O(n^3(\log n + \log\|A\|)^2)$ bit operations. Two iterations are sufficient. □

3 Computing the kth Invariant Factor with Random Rank k Perturbations

In this section we show that we can compute the kth invariant factor of a matrix $A \in \mathbb{Z}^{n\times n}$ by adding an appropriately generated random rank k matrix B to A. With sufficiently high probability the GCD of the largest invariant factor of $A+B$ with the largest invariant factor of A is the kth largest invariant factor of A. The proof of this is nontrivial and will require a number of lemmas. Throughout we assume that $A \in \mathbb{Z}^{n\times n}$, $s_1,\ldots,s_n \in \mathbb{Z}_{>0}$ are the invariant factors of A, and that $\sigma_1,\ldots,\sigma_n$ are the invariant factors of $A+B$.

LEMMA 3.1. *If B has rank at most k then s_i divides σ_{i+k} for $1 \le i \le n-k$.*

PROOF. The proof of an analogous result of [20], Lemma 2.1, for matrices whose entries are polynomials over a field carries over to the above result without change. □

In order to develop an efficient method to compute s_{n-k}, it will be useful to bound the probability that the highest power of a prime p dividing the largest invariant factor of $A+B$ is the same as that dividing s_{n-k}, when the above matrix B is computed as the product $B = UV$, where $U \in \mathbb{Z}^{n\times k}$ and $V \in \mathbb{Z}^{k\times n}$, whose entries are uniformly and independently chosen from the set of integers between 0 and $\beta-1$, and for a positive integer β to be determined later. A suitable bound will be derived in the next three subsections. It will be used, in the final subsection, to produce a reliable Monte Carlo algorithm to compute s_{n-k}.

3.1 A Sufficient Determinantal Condition

To begin, an additional assumption will be made — namely, that $A = S$ is in Smith normal form as in equation (1). Once again, let p be prime, let $h = \mathrm{ord}_p s_{n-k+1}$, so that

$h > 0$ and

$$A = \begin{bmatrix} p^h D_1 & 0 \\ 0 & D_2 \end{bmatrix}, \quad D_1 = \begin{bmatrix} p^{-h} s_n & & 0 \\ & \ddots & \\ 0 & & p^{-h} s_{n-k+1} \end{bmatrix},$$

$$D_2 = \begin{bmatrix} s_{n-k} & & 0 \\ & \ddots & \\ 0 & & s_1 \end{bmatrix}.$$

By the choice of h, both D_1 and D_2 are integer matrices. Suppose we partition U and V so that $U = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}$ and $V = \begin{bmatrix} V_1 & V_2 \end{bmatrix}$, where $U_1, V_1 \in \mathbb{Z}^{k \times k}$, and $U_2, V_2^t \in \mathbb{Z}^{(n-k) \times k}$.

Finally, let m be a positive integer such that the determinant of A is not divisible by p^m. Since none of the entries of the diagonal matrix D_2 is divisible by p^{h+1}, there exists a diagonal matrix E_2 with integer entries such that $D_2 E_2 = E_2 D_2 = p^h I + p^m F_2$ for yet another (diagonal) integer matrix F_2.

LEMMA 3.2. *Suppose A is in Smith form* (1), *and that both of the following conditions are satisfied:*

- $\det U_1 \not\equiv 0 \bmod p$;
- $\det[U_1(V_1 + V_2 E_2 U_2 \overline{U_1} D_1) + p^h D_1] \not\equiv 0 \bmod p$, *where $\overline{U_1}$ is an integer matrix such that $\overline{U_1} U_1 = I + p^m \widehat{U}$, for an integer matrix $\widehat{U}$.*

Then $\gcd(p, \sigma_n / s_{n-k}) = 1$.

PROOF. We will show that these conditions imply a stronger result, namely, that $\gcd(p, \sigma_{i+k}/s_i) = 1$ for all i such that $1 \leq i \leq n-k$. By Lemma 3.1, s_i divides σ_{i+k} for all such i, so that $\det D_2 = \prod_{i=1}^{n-k} s_i$ divides $\prod_{i=1}^{n-k} \sigma_{i+k}$, which in turn divides the determinant of $A + UV$. Let $\delta_{A,U,V} = \det(A+UV)/\det D_2 \in \mathbb{Z}$; clearly, the result now follows if

$$\operatorname{ord}_p \delta_{A,U,V} = 0. \tag{2}$$

The remainder of this proof will serve to establish this identity.

If A, U and V have the decompositions given above, then

$$A + UV = \begin{bmatrix} p^h D_1 + U_1 V_1 & U_1 V_2 \\ U_2 V_1 & D_2 + U_2 V_2 \end{bmatrix}.$$

Since $\det U_1 \not\equiv 0 \bmod p$, matrices $\overline{U_1}$ and $\widehat{U}$ with the properties mentioned in the final condition in the statement of the Lemma do exist,

$$\begin{aligned} &\begin{bmatrix} I & 0 \\ -U_2 \overline{U_1} & I \end{bmatrix} (A+UV) \\ &\quad = \begin{bmatrix} U_1 V_1 + p^h D_1 & U_1 V_2 \\ -p^h U_2 \overline{U_1} D_1 + p^m X_{2,1} & D_2 + p^m X_{2,2} \end{bmatrix} \end{aligned}$$

for integer matrices $X_{2,1}$ and $X_{2,2}$, and

$$\begin{aligned} &\begin{bmatrix} I & 0 \\ -U_2 \overline{U_1} & I \end{bmatrix} (A+UV) \begin{bmatrix} I & 0 \\ E_2 U_2 \overline{U_1} D_1 & I \end{bmatrix} \\ &\quad = \begin{bmatrix} U_1(V_1 + V_2 E_2 U_2 \overline{U_1} D_1) + p^h D_1 & U_1 V_2 \\ p^m Y_{2,1} & D_2 + p^m X_{2,2} \end{bmatrix} \\ &\quad = \begin{bmatrix} U_1(V_1 + V_2 E_2 U_2 \overline{U_1} D_1) + p^h D_1 & U_1 V_2 \\ 0 & D_2 \end{bmatrix} + p^m Z \end{aligned}$$

for another pair of integer matrices $Y_{2,1}$ and Z. Therefore

$$\begin{aligned} &\det(A+UV) \\ &\quad \equiv \det \begin{bmatrix} U_1(V_1 + V_2 E_2 U_2 \overline{U_1} D_1) + p^h D_1 & U_1 V_2 \\ 0 & D_2 \end{bmatrix} \bmod p^m. \end{aligned} \tag{3}$$

Let

$$\widehat{A} = \begin{bmatrix} U_1(V_1 + V_2 E_2 U_2 \overline{U_1} D_1) + p^h D_1 & U_1 V_2 \\ 0 & D_2 \end{bmatrix}.$$

Then, since $\det[U_1(V_1 + V_2 E_2 U_2 \overline{U_1} D_1) + p^h D_1] \not\equiv 0 \bmod p$,

$$\begin{aligned} \operatorname{ord}_p(\det \widehat{A}) &= \operatorname{ord}_p(\det[U_1(V_1 + V_2 E_2 U_2 \overline{U_1} D_1) + p^h D_1]) \\ &\quad + \operatorname{ord}_p(\det D_2) \\ &= \operatorname{ord}_p(\det D_2) \\ &\leq \operatorname{ord}_p(\det A) < m, \end{aligned}$$

and it follows by equation (3) that

$$\operatorname{ord}_p(\det(A+UV)) = \operatorname{ord}_p(\det \widehat{A}),$$

so that

$$\begin{aligned} \operatorname{ord}_p \delta_{A,U,V} &= \operatorname{ord}_p(\det(A+UV)) - \operatorname{ord}_p(\det D_2) \\ &= \operatorname{ord}_p(\det \widehat{A}) - \operatorname{ord}_p(\det D_2) \\ &= (\operatorname{ord}_p(\det[U_1(V_1 + V_2 E_2 U_2 \overline{U_1} D_1) + p^h D_1]) \\ &\quad + \operatorname{ord}_p(\det D_2)) - \operatorname{ord}_p(\det D_2) \\ &= 0 + \operatorname{ord}_p(\det D_2) - \operatorname{ord}_p(\det D_2) \\ &= 0, \end{aligned}$$

as required. □

3.2 Special Case: p Divides β

Suppose now that p divides β. To eliminate the assumption (used in Lemma 3.2) that A is in Smith normal form, note that there always exist unimodular integer matrices L and R such that $A = LSR$, where S is in Smith form as in (1). Now, the invariant factors of $A + UV$ are identical to those of $S + \widehat{U}\widehat{V}$ if $\widehat{U} = L^{-1}U$ and $\widehat{V} = VR^{-1}$. Assuming once again that p divides β, one can argue that the matrices $\widehat{U}$ and $\widehat{V}$ are chosen under the same probability distribution as U and V. The next result now follows.

LEMMA 3.3. *Suppose p divides β, and that the entries of $U \in \mathbb{Z}^{n\times k}$ and $V \in \mathbb{Z}^{k\times n}$ are selected uniformly and independently from the set of integers between 0 and $\beta - 1$. Let $B = UV$. Then $\gcd(p, \sigma_n/s_{n-k}) = 1$ with probability at least $\frac{2}{25}$ if $p = 2$ and with probability at least $(1-1/(p-1))^2$ if $p \geq 3$.*

PROOF. As noted above we may assume without loss of generality that A is in Smith normal form and that Lemma 3.2 is applicable. This lists two conditions that, together, imply that $\gcd(\sigma_n/s_{n-k}) = 1$.

The first of these conditions is that $\det U_1$ is not congruent to 0 modulo p. It is well known that if the entries of an $n\times n$ matrix are chosen uniformly from the finite field $\mathbb{Z}_p$ with p elements for any prime p, then the matrix is nonsingular with probability at least $1/4$ if $p = 2$ and with probability at least $1/(p-1)$ if $p \geq 3$. A slightly more careful approximation shows that the probability is at least $\frac{1}{5}\sqrt{2}$ if $p = 2$ — see, for example, [8] for details. Thus, the first condition is satisfied with probability at least $\frac{1}{5}\sqrt{2}$ if $p = 2$ and at least $1/(p-1)$ if $p \geq 3$, and the result will follow if the same bounds can be established for the conditional probability that the second condition is satisfied when the first is.

Suppose, therefore, that the first condition is satisfied, so that the matrix $\widetilde{U_1} = U_1 \bmod p$ with entries in $\mathbb{Z}_p$ is nonsingular. Since $\overline{U_1}U_1 \equiv I \bmod p$, it is sufficient to work in $\mathbb{Z}_p$ and to bound the probability that the matrix

$$V_1 + V_2E_2U_2\widetilde{U_1}^{-1}D_1 + p^h\widetilde{U_1}^{-1}D_1$$

is nonsingular: When the first condition is satisfied, the second condition is equivalent to the condition that the determinant of this matrix is nonzero in $\mathbb{Z}_p$. Now, it suffices to note that for any choice of U_2 and V_2, and for uniformly and randomly chosen V_1, the above matrix is uniformly and randomly chosen from $\mathbb{Z}_p^{k\times k}$. The result therefore follows from the bounds, that a randomly chosen square matrix with entries in $\mathbb{Z}_p$ is nonsingular, that have been given above. □

3.3 The General Case

In order to eliminate the assumption that p divides β, we will consider two independent types of trials of the process sketched above, involving two choices of β.

The second type of trial will use a value $\beta_2 \geq \lceil 2n^2(\log_2 n + \log_2\|A\|)\rceil$ and will be discussed later. The first type of trial will use an even value $\beta_1 \geq 21n^2\beta_2 \geq 21n^2\lceil 2n^2(\log_2 n + \log_2\|A\|)\rceil$.

Suppose now that p is a prime that is less than or equal to β_2; then $\frac{\beta_1}{p} \geq 21n^2$, so that

$$\left(1-\tfrac{p}{\beta_1}\right)^{2n^2} \geq \left(1-\tfrac{1}{21n^2}\right)^{2n^2} \geq \tfrac{9}{10},$$

and, similarly,

$$\left(1+\tfrac{p}{\beta_1}\right)^{2n^2} \leq \left(1+\tfrac{1}{21n^2}\right)^{2n^2} \leq \tfrac{11}{10}.$$

Suppose the entries of integer matrices U and V are chosen uniformly and independently from the set of integers between 0 and β_1, and consider any pair of matrices (of the same sizes) $\widehat{U}$ and $\widehat{V}$ with entries between 0 and $p-1$. Since p does not necessarily divide β_1, the probability that

$$U \equiv \widehat{U} \bmod p \qquad \text{and} \qquad V \equiv \widehat{V} \bmod p \tag{4}$$

may depend on the choice of $\widehat{U}$ and $\widehat{V}$. However, if an integer α is uniformly chosen between 0 and $\beta_1 - 1$, then for any given integer b such that $0 \leq b \leq p-1$, the probability that $\alpha \equiv b \bmod p$ is at least $\frac{\beta_1 - p}{p\beta_1} = \frac{1}{p}\left(1-\frac{p}{\beta_1}\right)$ and at most $\frac{\beta_1+p}{p\beta_1} = \frac{1}{p}\left(1+\frac{p}{\beta_1}\right)$. Consequently the ratio of the probability that equation (4) is satisfied, when $p \leq \beta_2$, to the probability that the equation is satisfied in the case that p divides β_1, is at least

$$\left(1-\tfrac{p}{\beta_1}\right)^{2nk} \geq \left(1-\tfrac{p}{\beta_1}\right)^{2n^2} \geq \frac{9}{10}$$

and at most

$$\left(1+\tfrac{p}{\beta_1}\right)^{2nk} \leq \left(1+\tfrac{p}{\beta_1}\right)^{2n^2} \leq \frac{11}{10}.$$

In combination with Lemma 3.3, this implies the following.

LEMMA 3.4. *Let β_1 be as defined above and suppose p is a prime such that $p \leq \beta_2$. If the entries of matrices $U \in \mathbb{Z}^{n\times k}$ and $V \in \mathbb{Z}^{k\times n}$ are selected uniformly and independently from a set of integers between 0 and $\beta_1 - 1$, and $B = UV$, then $\gcd(p, \sigma_n/s_{n-k}) = 1$ with probability at least*

$$\begin{cases} \frac{2}{25} & \text{if } p = 2, \\ \frac{1}{5} & \text{if } p = 3, \\ \frac{1}{2} & \text{if } p = 5, \\ 1 - \frac{11}{5(p-1)} & \text{if } p \geq 7. \end{cases}$$

In order to simplify the required mathematics, it will be useful to consider sets of five independent trials. The following is a trivial consequence of the above lemma and the fact that $(11/(5(p-1)))^5 \leq (2/(p-1))^4$ whenever $p \geq 7$.

COROLLARY 3.5. *Let β_1 be as defined above, suppose p is a prime such that $7 \leq p \leq \beta_2$, and suppose the entries of matrices $U_1, U_2, U_3, U_4, U_5 \in \mathbb{Z}_{n\times k}$ and $V_1, V_2, V_3, V_4, V_5 \in \mathbb{Z}^{k\times n}$ are selected uniformly and independently from a set of integers between 0 and $\beta_1 - 1$. Let $B_i = A + U_iV_i$ and suppose B_i has largest invariant factor $\sigma_{i,n}$ for $1 \leq i \leq 5$. Then*

$$\gcd\left(p, \tfrac{\sigma_{1,n}}{s_{n-k}}, \tfrac{\sigma_{2,n}}{s_{n-k}}, \tfrac{\sigma_{3,n}}{s_{n-k}}, \tfrac{\sigma_{4,n}}{s_{n-k}}, \tfrac{\sigma_{5,n}}{s_{n-k}}\right) = 1$$

with probability at least $1-\left(\frac{2}{p-1}\right)^4$.

Of course, these probabilities are quite low when p is small. However, a constant number of independent trials can be used to boost the probability of success.

LEMMA 3.6. *Let* β_1 *be as defined above, and suppose the entries of* $U_i \in \mathbb{Z}^{n\times k}$ *and* $V_i \in \mathbb{Z}^{k\times n}$ *are integers selected uniformly and independently between* 0 *and* $\beta_1 - 1$, *for* $1 \le i \le 15$. *Let* $\sigma_{i,1}, \sigma_{i,2}, \ldots, \sigma_{i,n}$ *be the invariant factors of* $A + U_iV_i$ *for* $1 \le i \le 15$. *Let*

$$\gamma = \gcd(\tfrac{\sigma_{1,n}}{s_{n-k}}, \tfrac{\sigma_{2,n}}{s_{n-k}}, \ldots, \tfrac{\sigma_{15,n}}{s_{n-k}}).$$

Then the probability that γ *has a prime factor less than or equal to* β_2 *is less than* $\frac{1}{3}$.

PROOF. Since β_1 is even, it follows by Lemma 3.3 that the probability that γ is even is at most $\left(\frac{23}{25}\right)^{15}$. It follows by Lemma 3.4 that γ is divisible by 3 (if $\beta_1 \ge 3$) with probability at most $\left(\frac{4}{5}\right)^{15}$, and that γ is divisible by 5 (if $\beta_1 \ge 5$) with probability at most $\left(\frac{1}{2}\right)^{15}$. Finally, it follows by Corollary 3.5 that if p is a prime between 7 and β_2, then the probability that p divides γ is at most $\left(\frac{2}{p-1}\right)^{12}$. Thus the probability that γ has a prime factor less than or equal to β_2 is at most

$$\begin{aligned}
&\left(\tfrac{23}{25}\right)^{15} + \left(\tfrac{4}{5}\right)^{15} + \left(\tfrac{1}{2}\right)^{15} + \sum_{\substack{7 \le p \le \beta_2 \\ p \text{ is prime}}} \left(\frac{2}{p-1}\right)^{12} \\
&\le \left(\tfrac{23}{25}\right)^{15} + \left(\tfrac{4}{5}\right)^{15} + \left(\tfrac{1}{2}\right)^{15} + \sum_{\substack{p \ge 7 \\ p \text{ is odd}}} \left(\frac{2}{p-1}\right)^{12} \\
&= \left(\tfrac{23}{25}\right)^{15} + \left(\tfrac{4}{5}\right)^{15} + \left(\tfrac{1}{2}\right)^{15} + \sum_{j \ge 3} \left(\frac{1}{j^{12}}\right) \\
&< \tfrac{1}{3},
\end{aligned}$$

as claimed. □

A single trial using β_2 is sufficient to filter out larger primes:

LEMMA 3.7. *Let* β_2 *be as defined above and suppose the entries of* $U_{16} \in \mathbb{Z}^{n\times k}$ *and* $V_{16} \in \mathbb{Z}^{k\times n}$ *are chosen uniformly and independently between* 0 *and* $\beta_2 - 1$. *Let* $\sigma_{16,1}, \sigma_{16,2}, \ldots, \sigma_{16,n}$ *be the invariant factors of* $A + U_{16}V_{16}$. *Then the probability that* $\gcd(s_n, \sigma_{16,n}/s_{n-k})$ *is divisible by a prime that is greater than* β_2 *is at most* $1/n$.

PROOF. Let p be a prime greater than β_2, and (abusing notation) consider $\Gamma = \{0, 1, \ldots, \beta_2 - 1\}$ as a subset of β_2 distinct values in the finite field $\mathbb{Z}_p$.

Suppose first that A is in Smith normal form. Then, by Lemma 3.2, $\gcd(p, \sigma_{16,n}/s_{n-k}) = 1$ as long as two determinantal conditions are satisfied.

The first of these conditions is that the determinant of the leading $k \times k$ submatrix of U_{16} is relatively prime to p and, since the entries of (the residue mod p of) U_{16} are chosen uniformly and independently from the above set Γ, it can be argued by the Schwartz-Zippel lemma that this first condition fails with probability at most $k/|\Gamma| \le n/|\Gamma| \le 1/(2n^2(\log_2 n + \log_2 \|A\|))$.

Suppose the first condition succeeds, and consider the determinant used in the second condition, as a function of the entries of the leading $k \times k$ submatrix of V_{16}: For every choice of the remaining entries of U_{16} and V_{16}, this determinant is a nonzero function with total degree at most k in the entries of the leading $k \times k$ submatrix of V_{16}, so it follows once again by the Schwartz-Zippel lemma that the probability that the first condition succeeds and the second fails is at most $1/(2n^2(\log_2 n + \log_2 \|A\|))$.

The condition that A is in Smith normal form can be eliminated, as usual, by noting that $A = LSR$ for unimodular matrices L and R and for a matrix S in Smith normal form, so that

$$A + U_{16}V_{16} = L(S + (L^{-1}U_{16})(V_{16}R^{-1}))R.$$

We must now consider determinantal conditions involving $L^{-1}U_{16}$ and $V_{16}R^{-1}$, where L and R are fixed unimodular matrices depending only on A. The first determinantal condition is that the leading $k \times k$ submatrix of $L^{-1}U_{16}$ has a determinant relatively prime to p. Once again, since the entries of this submatrix are (at most) linear in the entries of U_{16}, it can be argued that this condition fails with probability at most $1/(2n^2(\log_2 n + \log \|A\|))$. The second condition involves the determinant (mod p) of a $k \times k$ matrix as well. Once again, assuming that the first condition succeeds, it can be observed that this determinant is a nonzero function of the entries of V_{16}, for every possible choice of U_{16}. Since each entry of the submatrix (whose determinant one must check) is at most linear in the entries of V_{16}, the Schwartz-Zippel can be applied, once again, to establish that the probability that the second condition fails when the first succeeds is at most $1/(2n^2(\log_2 n + \log_2 \|A\|))$. Thus, the probability that $\gcd(p, \sigma_{16,n}/s_{n-k})$ is different from 1 is at most $1/n^2(\log_2 n + \log_2 \|A\|)$ for any prime $p > \beta_2$.

Since $\log_2 s_n \le \log_2(\det A) \le n(\log_2 n + \log_2 \|A\|)$, there are at most $n(\log_2 n + \log_2 \|A\|)$ distinct primes p that divide s_n. Thus the probability that there exists any prime $p > \beta_2$ dividing $\gcd(s_n, \sigma_{16,n}/s_{n-k})$ is at most $1/n$.

Obviously, no prime p that is relatively prime to s_n can divide $\gcd(s_n, \sigma_{16,n}/s_{n-k})$, so this implies the desired result. □

3.4 Computing an Invariant Factor

THEOREM 3.8. *Suppose* $n \ge 6$ *and let* β_1 *and* β_2 *be as defined in Subsection 3.3, above. Suppose that the entries of*

$U_i \in \mathbb{Z}^{n\times k}$ *and* $V_i \in \mathbb{Z}^{k\times n}$ *are chosen uniformly and independently from the set of integers between* 0 *and* $\beta_1 - 1$*, for* $1 \le i \le 15$*, and that the entries of* $U_{16} \in \mathbb{Z}^{n\times k}$ *and* $V_{16} \in \mathbb{Z}^{k\times n}$ *are chosen uniformly and independently from the set of integers between* 0 *and* $\beta_2 - 1$.

Let $\sigma_{i,1}, \sigma_{i,2}, \ldots, \sigma_{i,n}$ *be the invariant factors of the matrix* $A + U_iV_i$ *for* $1 \le i \le 16$*. Then*

$$\gcd(s_n, \sigma_{1,n}, \sigma_{2,n}, \ldots, \sigma_{16,n}) = s_{n-k}$$

with probability at least $1/2$.

PROOF. Since s_{n-k} divides s_n, and $\sigma_{i,n}$ is divisible by s_{n-k} for $1 \le i \le 16$, the above condition is satisfied if

$$\gcd(s_n, \tfrac{\sigma_{1,n}}{s_{n-k}}, \tfrac{\sigma_{2,n}}{s_{n-k}}, \ldots, \tfrac{\sigma_{16,n}}{s_{n-k}}) = \gcd(\gcd(s_n, \tfrac{\sigma_{16,n}}{s_{n-k}}), \gcd(\tfrac{\sigma_{1,n}}{s_{n-k}}, \ldots, \tfrac{\sigma_{15,n}}{s_{n-k}})) = 1$$

The desired probability now follows from Lemmas 3.6 and 3.7. □

In the sequel we will assume that a Monte Carlo algorithm `OneInvariantFactor`, on input A and k, computes s_k as suggested in the above theorem. Since β_0, β_1 are small, the number of bit operations used is $O(n^3(\log n + \log\|A\|)^2)$ – see Theorem 2.2 and note that a constant number of trials of the required algorithms suffice to reduce the failure probability to $1/2$.

4 Computing the Smith form and determinant

In this section we present the algorithm for the Smith form of an integer matrix $A \in \mathbb{Z}^{n\times n}$. This will also give an algorithm for computing the determinant once we have computed its sign.

As we noted earlier, the algorithm is essentially a binary search for the distinct invariant factors, using the algorithm `LargestInvariantFactor` to capture the largest invariant factor, and the perturbation theory described in Section 3, which isolates an arbitrary invariant factor as the greatest common divisor of the largest invariant factor of A and a constant number of perturbed matrices.

A key point is that A has a small number of distinct invariant factors:

THEOREM 4.1. *Any* $A \in \mathbb{Z}^{n\times n}$ *has at most* $\sqrt{3\log_2|\det A|} = O(n^{1/2}(\log n + \log\|A\|)^{1/2})$ *distinct invariant factors.*

PROOF. Suppose A has invariant factors $s_1, \ldots, s_n \in \mathbb{Z}_{>0}$ with $s_j \mid s_{j+1}$ for $1 \le j < n$. Let $s_{i_1}, \ldots, s_{i_\mu}$ be the *distinct* invariant factors of A in increasing order. We know

$$\prod_{1\le j\le \mu} s_{i_j} \le \prod_{1\le j\le n} s_j = |\det A|.$$

As well $2s_{i_j} \le s_{i_{j+1}}$ for $1 \le j < \mu$ so $s_{i_j} \ge 2^{j-1}$ for $1 \le j \le \mu$ and

$$\prod_{1\le j\le\mu} s_{i_j} \ge 2^{\sum_{1\le j\le\mu} j-1} = 2^{\mu(\mu-1)/2} \ge 2^{\mu^2/3}$$

for $\mu > 1$. Thus $\log_2|\det A| \ge \mu^2/3$ and $\mu \le \sqrt{3\log_2|\det A|}$. □

The algorithm `InvariantFactors` recursively computes the invariant factors of A. It maintains a list $\sigma = (\sigma_1, \ldots, \sigma_n)$ of the invariant factors known so far, with $\sigma_i = s_i$, the ith invariant factor (with high probability), or $\sigma_i = *$, if it has not yet been computed. The complete algorithm is invoked by

$$\texttt{InvariantFactors}(A, 1, n, (s_1, *, \ldots, *, s_n)); \qquad (5)$$

where $s_1 = \gcd_{1\le i,j\le n}(A_{ij})$ is the first invariant factor of A, and s_n is the last (largest) invariant factor of A, computed correctly with probability at least $1 - 1/(2n)$ using `LargestInvariantFactor(A)` (called $\lceil\log_2(2n)\rceil$ times to achieve the desired probability). We assume that $s_1 < s_n$ (otherwise we are done – all the invariant factors are the same).

Algorithm: `InvariantFactors`

Input:
- $A \in \mathbb{Z}^{n\times n}$;
- $i, j \in \mathbb{Z}$;
- $\sigma = (\sigma_1, \ldots, \sigma_n) \in (\mathbb{Z} \cup \{*\})^{n+1}$ with $\sigma_i, \sigma_j \in \mathbb{Z}$ the ith and jth invariant factors of A respectively and $\sigma_i < \sigma_j$;

Output:
- $\sigma \in (\mathbb{Z} \cup \{*\})^n$, such that $\sigma_k = s_k$ is the kth invariant factor of A for $i \le k \le j$;

(1) If $i + 1 = j$ return σ;
(2) $m := \lfloor (i+j)/2 \rfloor$;
(3) Let $\sigma_m :=$ `OneInvariantFactor`(A, m), the mth invariant factor of A (computed correctly with probability at least $1 - 1/(2n)$);
(4) If $(\sigma_m = \sigma_i)$
(5) Then For $\ell := i+1$ to $m-1$ do $\sigma_\ell := \sigma_m$;
(6) Else $\sigma :=$ `InvariantFactors`(A, i, m, σ);
(7) If $(\sigma_m = \sigma_j)$
(8) Then For $\ell := m+1$ to $j-1$ do $\sigma_\ell := \sigma_m$;
(9) Else $\sigma :=$ `InvariantFactors`(A, m, j, σ);
(10) Return σ;

THEOREM 4.2. *Given any* $A \in \mathbb{Z}^{n\times n}$*, the algorithm* `InvariantFactors` *computes all the invariant factors of* A *correctly with probability at least* $1/2$ *on any invocation as in (5). The algorithm requires* $O(n^3(\log n + \log\|A\|)^2\sqrt{\log|\det A|}\cdot\log^2 n)$ *bit operations. In the worst case (i.e., insensitive to the size of the determinant of* A*) it requires* $O(n^{3.5}(\log n + \log\|A\|)^{2.5}\cdot\log^2 n)$ *bit operations.*

PROOF. We first address the issue of correctness. Clearly, when it fills in any invariant factor, it does so correctly with probability at least $1-1/(2n)$. On output, every element of σ is an invariant factor. Thus, with probability $1/2$ the algorithm correctly computes the entire Smith form.

To see the complexity, we first note that we compute the mth invariant factor of A in Step (3) using the technique described in Theorem 3.8. It requires a constant number of calls to `LargestInvariantFactor` on very small perturbations of A, and hence each can be executed with $O(n^3(\log n+\log\|A\|)^2)$ bit operations.

The algorithm is essentially doing a binary search for each of the points at which the invariant factors change. There are $O(\sqrt{\log|\det A|})$ such points, so the total number of evaluations of step (3) is $O(\sqrt{\log|\det A|}\cdot\log n)$. Since `OneInvariantFactor` returns correctly with probability at most $1/2$, we must run it $\lceil 1+\log_2 n\rceil$ times and use the smallest value returned to achieve probability of correctness greater than $1-1/(2n)$. □

Once we have computed the invariant factors of A, $\det A$ is simply their product $d=s_1\cdots s_n$, up to the sign. To determine the sign of the determinant, find a prime p greater than 2 which does not divide $\det A$ – a random prime with $6+\log\log|\det A|=O(\log n+\log\log\|A\|)$ bits will satisfy this with probability at least $1/2$ (see, e.g., Giesbrecht [11]). Compute $\det A \bmod p$ and check whether $\det A\equiv d \bmod p$ or $\det A\equiv -d \bmod p$. This can be done with $O(n^3(\log n+\log\log|\det A|)^2)$ bit operations, which is dominated by the time required for `InvariantFactors`. Notice that this also gives a fast check that the proposed determinant from `InvariantFactors` is correct: if $\det A\not\equiv \pm d \bmod p$ then our computation for the determinant is incorrect and should be repeated.

THEOREM 4.3. *Given any $A\in\mathbb{Z}^{n\times n}$, we can compute $\det A$ correctly with probability at least $1/2$ on any invocation of the algorithm discussed above. The algorithm requires $O(n^3(\log n+\log\|A\|)^2\sqrt{\log|\det A|}\cdot\log^3 n)$ bit operations using standard matrix arithmetic. In the worst case (i.e., insensitive to the size of the determinant of A) it requires $O(n^{3.5}(\log n+\log\|A\|)^{2.5}\cdot\log^3 n)$ bit operations.*

The determinant can be checked for correctness with probability at least $1/2$ (on any invocation of the check) using $O(n^3(\log n+\log\log|\det A|)^2)$ bit operations.

5 An asymptotically faster algorithm

If asymptotically fast matrix arithmetic is available, we can exploit it through a tradeoff to the algorithm of Storjohann [18] for computing the Smith form over $\mathbb{Z}_d$ for some integer d. More specifically, we use the method described in Sections 2 and 3 to compute all the invariant factors larger than some pre-determined bound C. We then compute the remaining invariant factors by computing the Smith form of A modulo the first invariant factor which is smaller than C using Storjohann's algorithm.

Algorithm: `FastInvariantFactors`

Input: • $A\in\mathbb{Z}^{n\times n}$;

Output: • $\sigma=(\sigma_1,\ldots,\sigma_n)\in\mathbb{Z}^n$, where σ_i is the ith invariant factor of A;

(1) Let
$$C:=\left\lceil \exp\left(\frac{n^{2-\theta/2}(\log n+\log\|A\|)^{3/2}}{(\log n+\log\log\|A\|)^{1/2}}\right)\right\rceil;$$
(2) $i:=n+1$;
(3) Repeat
(4) $\quad i:=i-1$;
(5) $\quad \sigma_i$:= `OneInvariantFactor`(A,i), the ith invariant factor of A (computed correctly with probability at least $1-1/(2n)$);
(6) Until $\sigma_i<C$;
(7) Compute the Smith form of $A \bmod \sigma_i$ using the algorithm of Storjohann [18] and extract invariant factors $\sigma_1,\ldots,\sigma_{i-1}\in\mathbb{Z}$;
(8) Return $\sigma=(\sigma_1,\ldots,\sigma_n)$.

THEOREM 5.1. *Given any $A\in\mathbb{Z}^{n\times n}$, the algorithm `FastInvariantFactors` computes all the invariant factors of A correctly with probability at least $1/2$ on any invocation. The algorithm requires $O(n^{2+\theta/2}\cdot(\log n+\log\|A\|)^{3/2}\log(n)^{1/2}\cdot(\log\log n+\log\log\|A\|))$ bit operations.*

PROOF. To see that the algorithm works we note that by the same argument as in Theorem 4.2 that $\sigma_n,\ldots,\sigma_i$ are correctly computed with probability at least $1/2$ (where σ_i is the largest invariant factor smaller than C). The integer pre-images of the invariant factors of $A \bmod \sigma_i$ are exactly those of A up to a unit in $\mathbb{Z}_{\sigma_i}$. By normalizing with a GCD with σ_i (as done in [18]), we obtain the desired invariant factors $\sigma_1,\ldots,\sigma_{i-1}$ of A.

To determine the complexity of this algorithm, we note that $C^i\le|\det A|<n^n|A|^n$, so we need to perform $O(n(\log n+\log\|A\|)/\log C)$ iterations of the loop (3)-(6). By the analysis of the previous section, this requires
$$T_1=O\left(\frac{n(\log n+\log\|A\|)}{\log C}\cdot n^3(\log n+\log\|A\|)^2\log n\right)$$
bit operations, and the probability that all those invariant factors computed are correct is at least $1/2$.

In step (6), Storjohann's algorithm requires $O(n^\theta)$ operations with integers in $\mathbb{Z}_{\sigma_i}$, each of which has $O(\log C)$ bits. The cost of this step is thus
$$T_2=O\left(n^\theta(\log C)(\log\log C)(\log\log\log C)\right)$$

bit operations, using fast integer arithmetic.

With the selected C, the total cost of the algorithm is

$$T_1 + T_2 = O\left(n^{2+\theta/2}(\log n + \log\|A\|)^{3/2}\log(n)^{1/2} \cdot (\log\log n + \log\log\|A\|)\right).$$

as specified. $\square$

6 On the Expected Number of Invariant Factors

In this section we consider the expected number of invariant factors of an $n \times n$ matrix whose entries are chosen uniformly and independently from a small finite set. Suppose, in particular, that $a \in \mathbb{Z}$, $\lambda \in \mathbb{Z}_{\geq 2}$ and

$$\Lambda = \{a, a+1, a+2, \ldots, a+\lambda-1\} \tag{6}$$

is a set containing λ contiguous integers. Suppose the entries of an $n \times n$ matrix A are chosen uniformly and independently from Λ. We will show that the expected number of nontrivial invariant factors of A is then in $O(\log_\lambda n)$.

It will be useful to consider two kinds of events. For $1 \leq i \leq n$, let Dep_i denote the event that the first i columns of A are linearly dependent (over the rationals), and let MDep_i denote the event that there exists at least one prime p such that the submatrix including the first i columns of A mod p has rank at most $i-2$, in the field $\mathbb{Z}_p$ of integers mod p.

Since a set of vectors of size one is only linearly dependent if the vector in the set is the zero vector, the probability of event Dep_1 is at most λ^{-n} and the probability of event MDep_1 is zero, whence

$$P[\mathsf{Dep}_1 \vee \mathsf{MDep}_1] \leq \lambda^{-n}. \tag{7}$$

Suppose now that $2 \leq i \leq n$ and that neither the event Dep_{i-1} nor the event MDep_{i-1} is satisfied. Then the first $i-1$ columns of the matrix A are linearly independent, and there exists a set R_{i-1} of $i-1$ rows such that the submatrix $A_{R_{i-1}}$ including the entries of A in the first $i-1$ columns and the rows in R_{i-1} is nonsingular. Consider any choice of entries of A in the rows in R_{i-1} and in column i, and let $\widehat{A}_{R_{i-1}}$ be the $(i-1) \times i$ submatrix of A including entries in these rows and the first i columns. The entries of A in the remaining rows and in column i are selected independently of the entries of $\widehat{A}_{R_{i-1}}$, one another, and of all other entries of A. Thus if $1 \leq j \leq n$, $j \notin R_{i-1}$, and v_j is the vector of dimension i including entries in the first i columns of row j of A, then this vector is a linear combination of the rows of $\widehat{A}_{R_{i-1}}$ with probability at most λ^{-1}, because there is only choice of the final entry of v_j achieving this condition for any choice of the first $i-1$ entries. Consequently, since there are $n-i+1$ rows that do not belong to R_{i-1}, the probability that the first i columns of A are linearly dependent is at most λ^{i-n-1}. That is,

$$P[\mathsf{Dep}_i \mid \neg(\mathsf{Dep}_{i-1} \vee \mathsf{MDep}_{i-1})] \leq \left(\tfrac{1}{\lambda}\right)^{n-i+1}. \tag{8}$$

Next, note that if p is any prime that does not divide the determinant of $A_{R_{i-1}}$ then the first $i-1$ columns of the matrix A mod p are clearly linearly independent (over the finite field $\mathbb{Z}_p$), so the submatrix of A mod p including the first i columns clearly has rank at least $i-1$ over the field of integers mod p. Thus the event MDep_i can only occur if there exists a prime p dividing the determinant of $A_{R_{i-1}}$ such that the submatrix including the first i columns of A mod p has rank at most $i-2$ over $\mathbb{Z}_p$.

In order to bound the probability that this occurs, it will be useful to consider primes $p < \lambda$ and primes $p \geq \lambda$ separately.

Lemma 6.1. *Let $\lambda \in \mathbb{Z}_{\geq 2}$ and $i < n-1$, and suppose that the event MDep_{i-1} is false. The probability that there exists* any *prime $p < \lambda$ such that the submatrix of A* mod *p including the first i columns has rank at most $i-2$ is at most $(2/3)^{n-i+2} + (2/3)\cdot(1/2)^{n-i}$.*

Proof. The required analysis depends on the size λ; if $\lambda = 2$ then no such prime p exists and the result is trivial, so the case that $\lambda = 3$ is the first nontrivial one.

In this case, the only prime p to consider is $p = 2$. Since the event MDep_{i-1} did not occur, the submatrix including the first $i-1$ columns of A mod 2 has rank at least $i-2$. If it has maximal rank $i-1$ then the submatrix including the first i columns of A mod 2 has rank at least $i-1$ as well and the event MDep_i could not arise. On the other hand, if it has rank $i-2$ then there exists a set $\overline{R}_{i-2}$ of $i-2$ rows of A such that the $(i-2) \times i$ submatrix of A mod 2 including entries in rows from $\overline{R}_{i-2}$ and the first i columns has full rank. For each row $j \notin \overline{R}_{i-2}$, if v_j is the vector of dimension i including the first i entries of row j, then the probability that v_j is a linear combination of the rows of the above submatrix is at least $\frac{2}{3}$, since there is only one choice (mod 2) for the final entry of v_j achieving this condition, for each choice of the initial entries, and since this choice is made with probability at most $2/3$ when $p = 2$ and $\lambda = 3$. Since the entries of rows are selected independently, it follows that the submatrix including the first i columns of A mod 2 has rank less than or equal to $i-2$ with probability at most $\left(\frac{2}{3}\right)^{n-i+2}$.

If $\lambda = 4$ then one must consider the primes $p = 2$ and $p = 3$. For each prime, one can argue as above to obtain a bound of "failure" of $\left(\frac{2}{4}\right)^{n-i+2}$, so the probability that there exists any prime $p < \lambda = 4$ such that the rank of the submatrix with the first i columns of A mod p is too small is at most $2\left(\frac{1}{2}\right)^{n-i+2}$, and this is less than or equal to $\left(\frac{2}{3}\right)^{n-i+2}$ whenever $i \leq n-1$.

Similar analyses yield the bounds $\left(\frac{3}{5}\right)^{n-i+2}+\left(\frac{2}{5}\right)^{n-i+2}$ when $\lambda=5$ and $\left(\frac{1}{2}\right)^{n-i+2}+\left(\frac{1}{3}\right)^{n-i+2}+\left(\frac{2}{5}\right)^{n-i+2}$ when $\lambda=6$. Each of these is less than or equal to $\left(\frac{2}{3}\right)^{n-i+2}$ whenever $i\leq n-1$.

Finally, if $\lambda\geq 7$ and $i\leq n-1$ then it follows by a similar analysis that the relevant probability is at most

$$\begin{aligned}\left(\tfrac{4}{7}\right)^{n-i+2}+\left(\tfrac{3}{7}\right)^{n-i+2}+\sum_{\substack{5\leq p\\ p\text{ is odd}}}\left(\frac{2}{p-1}\right)^{n-i+2}&\\ \leq\left(\tfrac{2}{3}\right)^{n-i+2}+\left(\tfrac{1}{2}\right)^{n-i}(\zeta(2)-1)&\\ <\left(\tfrac{2}{3}\right)^{n-i+2}+\tfrac{2}{3}\left(\tfrac{1}{2}\right)^{n-i},&\end{aligned}$$

as needed to complete the proof. □

Now, if p is a prime that is greater than or equal to λ that divides the determinant of $A_{R_{i-1}}$, then the probability that the submatrix including the first i columns of A mod p has rank at most $i-2$ is at most $\left(\frac{1}{\lambda}\right)^{n-i+2}$, since the likelihood that a given entry of this matrix assumes any fixed value, mod p, is either 0 or $\frac{1}{\lambda}$. The determinant of $A_{R_{i-1}}$ is a nonzero integer with absolute value at most

$$(i-1)!\lambda^{i-1}<((i-1)\lambda)^{i-1},$$

and the number of primes $p\geq\lambda$ that divide this determinant is at most

$$\log_\lambda((i-1)\lambda)^{i-1}<(i-1)(1+\log_\lambda n).$$

Therefore, when neither of the events Dep_{i-1} or MDep_i arise, the probability that there exists a prime $p\geq\lambda$ such that the submatrix including the first i columns of A mod p has rank at most $i-2$ over $\mathbb{Z}_p$ is at most

$$(i-1)(1+\log_\lambda n)\left(\tfrac{1}{\lambda}\right)^{n-i+2}<n(1+\log_\lambda n)\left(\tfrac{1}{\lambda}\right)^{n-i+2}.$$

It now follows by the definition of the event MDep_i that

$$\begin{aligned}&P\big[\mathsf{MDep}_i\mid\neg(\mathsf{Dep}_{i-1}\vee\mathsf{MDep}_{i-1})\big]\\ &\quad\leq\left(\tfrac{2}{3}\right)^{n-i+2}+\tfrac{2}{3}\left(\tfrac{1}{2}\right)^{n-i}+n(1+\log_\lambda n)\left(\tfrac{1}{\lambda}\right)^{n-i+2}.\end{aligned}\tag{9}$$

Now, equations (8) and (9) imply that if $1\leq i\leq n-1$ then

$$\begin{aligned}&P[\mathsf{Dep}_i\vee\mathsf{MDep}_i\mid\neg(\mathsf{Dep}_{i-1}\vee\mathsf{MDep}_{i-1})]\\ &\leq\left(\tfrac{2}{3}\right)^{n-i+2}+\tfrac{2}{3}\left(\tfrac{1}{2}\right)^{n-i}+\left(\tfrac{1}{\lambda}\right)^{n-i+1}+n(1+\log_\lambda n)\left(\tfrac{1}{\lambda}\right)^{n-i+2}.\end{aligned}\tag{10}$$

This clearly implies (by the definition of conditional probability) that

$$\begin{aligned}&P[\mathsf{Dep}_i\vee\mathsf{MDep}_i\wedge\neg(\mathsf{Dep}_{i-1}\vee\mathsf{MDep}_{i-1})]\\ &\leq\left(\tfrac{2}{3}\right)^{n-i+2}+\tfrac{2}{3}\left(\tfrac{1}{2}\right)^{n-i}+\left(\tfrac{1}{\lambda}\right)^{n-i+1}+n(1+\log_\lambda n)\left(\tfrac{1}{\lambda}\right)^{n-i+2}\end{aligned}\tag{11}$$

as well. Thus if $n\geq 2$ then

$$\begin{aligned}P[\mathsf{MDep}_i]&\leq P\left[\bigvee_{j=1}^{i}(\mathsf{Dep}_j\vee\mathsf{MDep}_j)\right]\\ &\leq\lambda^{-n}+\sum_{j=2}^{i}\left(3\left(\tfrac{2}{3}\right)^{n-j+1}+n^3\left(\tfrac{1}{\lambda}\right)^{n-j+2}\right)\\ &=\lambda^{-n}+3\left(\tfrac{2}{3}\right)^{n-i+1}\sum_{h=0}^{i-2}\left(\tfrac{2}{3}\right)^h+n^3\left(\tfrac{1}{\lambda}\right)^{n-i+2}\sum_{h=0}^{i-2}\left(\tfrac{1}{\lambda}\right)^h\\ &\leq\lambda^{-n}+9\left(\tfrac{2}{3}\right)^{n-i+1}+n^3\left(\tfrac{1}{\lambda}\right)^{n-i+1}.\end{aligned}$$

THEOREM 6.2. *If the entries of an $n\times n$ matrix A are chosen uniformly and independently from the set Λ shown in equation* (6), *then the probability that A has at least j nontrivial invariant factors is at most*

$$\lambda^{-n}+9\left(\tfrac{2}{3}\right)^{j-1}+n^3\left(\tfrac{1}{\lambda}\right)^{j-1}.$$

PROOF. The claim is trivial if $j\leq 3$, since the given probability bound exceeds one in this case. Suppose, therefore, that $j\geq 4$.

If A has at least j nontrivial invariant factors then there is some prime p dividing the largest j invariant factors, so that A mod p is equivalent to a diagonal matrix with at least j zeroes on its diagonal. Therefore A mod p has rank at most $n-j$ over $\mathbb{Z}_p$. Clearly, then, the submatrix including the first $n-j+2$ columns of A mod p has rank at most $n-j$ as well, implying that the condition MDep_{n-j+2} is satisfied.

The claim now follows by the above probability bound for MDep_i, with $i=n-j+2$. □

COROLLARY 6.3. *If the entries of an $n\times n$ matrix A are chosen uniformly and independently from the set Λ shown in equation* (6), *then the expected number of nontrivial invariant factors of A is in $O(\log_\lambda n)$.*

PROOF. It follows trivially by the previous claim that the probability that A has at least j nontrivial invariant factors is at most the minimum of 1 and

$$\lambda^{-n}+9\left(\tfrac{2}{3}\right)^{j-1}+n^3\left(\tfrac{1}{\lambda}\right)^{j-1}.$$

The expected number of nontrivial invariant factors is therefore at most

$$\sum_{j\le\lceil 3\log_\lambda n\rceil} 1 + \sum_{j=\lceil 3\log_\lambda n\rceil+1}^{n} \left(\lambda^{-n} + 9\left(\tfrac{2}{3}\right)^{j-1} + n^3\left(\tfrac{1}{\lambda}\right)^{j-1}\right)$$
$$\le \sum_{j\le\lceil 3\log_\lambda n\rceil} 1 \;+ \sum_{j=\lceil 3\log_\lambda n\rceil+1}^{n} \lambda^{-\log_\lambda n}$$
$$+9 \sum_{j=\lceil 3\log_\lambda n\rceil+1}^{n} \left(\tfrac{2}{3}\right)^{j-1} + n^3 \sum_{j=\lceil 3\log_\lambda n\rceil+1}^{n} \left(\tfrac{1}{\lambda}\right)^{j-1}$$
$$\le 3\lceil\log_\lambda n\rceil + \sum_{j=1}^{n}\tfrac{1}{n} + 9\sum_{h\ge 0}\left(\tfrac{2}{3}\right)^h$$
$$+ n^3\left(\tfrac{1}{\lambda}\right)^{3\lceil\log_\lambda n\rceil}\sum_{h\ge 0}\left(\tfrac{1}{\lambda}\right)^h$$
$$\le 3\lceil\log_\lambda n\rceil + 1 + 27 + \tfrac{n^3}{n^3}\left(\tfrac{\lambda}{\lambda-1}\right)$$
$$\le 3\lceil\log_\lambda n\rceil + 29 \le 3\log_\lambda n + 32$$
$$= O(\log_\lambda n),$$

as required. □

The algorithm `InvariantFactors` from Section 4 will quickly identify the non-trivial invariant factors. The cost, as we have noted earlier, is dependent upon the *number* of distinct non-trivial invariant factors.

THEOREM 6.4. *Let $\Lambda = \{a, a+1, \ldots, a+\lambda-1\}$ for $a \in \mathbb{Z}$ and $\lambda \in \mathbb{Z}_{\ge 2}$, and suppose that the $n\times n$ matrix $A \in \mathbb{Z}$ is chosen uniformly and randomly from $\Lambda^{n\times n}$. The expected cost of the algorithm `InvariantFactors` to find the Smith form and determinant of A is $O(n^3(\log n + \log\|A\|)^2 \cdot \log(n)\log_\lambda(n))$.*

Acknowledgement

The authors would like to thank Erich Kaltofen for pointing out the bit-complexity analysis of [14].

References

[1] J. Abbott, M. Bronstein, and T. Mulders. Fast deterministic computation of determinants of dense matrices. In *Proc. of ACM International Symposium on Symbolic and Algebraic Computation (ISSAC'1999)*, pages 197–204, Vancouver, 1999. ACM Press.

[2] F. Avnaim, J.-D. Boissonnat, O. Devillers, F. Preparata, and M. Yvinec. Evaluation of a new method to compute signs of determinants. In *Communication at the 11th Annu. ACM Sympos. Comput. Geom.*, pages C16–C17, 1995.

[3] F. Avnaim, J.-D. Boissonnat, O. Devillers, F. Preparata, and M. Yvinec. Evaluating signs of determinants using single-precision arithmetic. *Algorithmica*, 17:111–132, 1997.

[4] H. Cohen. *A Course in Computational Number Theory*. Springer, 1996.

[5] D. Coppersmith and S. Winograd. Matrix multiplication via arithmetic progressions. In *Proc. 19th Annual ACM Symposium on Theory of Computing*, pages 1–6, New York, NY, 1987.

[6] J. Dixon. Exact solution of linear equations using p-adic expansions. *Numer. Math.*, 40(1):137–141, 1982.

[7] J. Dumas, B. Saunders, and G. Villard. Integer Smith form via the valence: experience with large sparse matrices from homology. In *Proc. of ACM International Symposium on Symbolic and Algebraic Computation (ISSAC'2000)*, 2000. To appear.

[8] W. Eberly. Processor-efficient parallel matrix inversion over abstract fields: Two extensions. In M. Hitz and E. Kaltofen, editors, *Proceedings, Second International Symposium on Parlallel Symbolic Computation, PASCO '97*, pages 38–45, Maui, Hawaii, July 1997.

[9] The GAP Group, Aachen, St Andrews. *GAP – Groups, Algorithms, and Programming, Version 4.2*, 1999. (`http://www-gap.dcs.st-and.ac.uk/~gap`).

[10] J. von zur Gathen and M. Sieveking. Weitere zum Erfüllungsproblem polynomial äquivalente kombinatorial sche Aufgaben. In *Komplexität von Entscheidungsproblemen: Ein Seminar*, pages 49–70, 1976.

[11] M. Giesbrecht. *Nearly Optimal Algorithms for Canonical Matrix Forms*. PhD thesis, University of Toronto, 1993. 196 pp.

[12] M. Giesbrecht. Probabilistic computation of the Smith normal form of a sparse integer matrix. In H. Cohen, editor, *Algorithmic Number Theory: Second International Symposium*, pages 175–188, 1996.

[13] M. Giesbrecht. Fast computation of the Smith form of a sparse integer matrix. *Computational Complexity*. Submitted.

[14] E. Kaltofen. On computing determinants of matrices without division. In *Proc. ACM International Symposium on Symbolic and Algebraic Computation (ISSAC'92)*, pages 342–349, Berkeley, USA, 1992.

[15] E. Kaltofen, 2000. Private Communication.

[16] T. Mulders and A. Storjohann. Diophantine linear system solving. In *Proc. of ACM International Symposium on Symbolic and Algebraic Computation (ISSAC'1999)*, pages 181–188, Vancouver, 1999. ACM Press.

[17] V. Pan. Computing the determinant and the characteristic polynomial of a matrix via solving linear systems of equations. *Information Processing Letters*, 28(2):71–75, 1988.

[18] A. Storjohann. Near optimal algorithms for computing Smith normal forms of integer matrices. In *Proceedings of ISSAC'96*, pages 267–274, Zurich, Switzerland, 1996.

[19] G. Villard. *Calcul formel et parallélisme: Resolution de Systemes Linear*. PhD thesis, L'institut National Polytechnique de Grenoble, 1988.

[20] G. Villard. Computing the Frobenius normal form of a sparse matrix. In *Proc. Third International Workshop on Computer Algebra in Scientific Computing (CASC-2000)*, 2000. To appear.