Black Box Linear Algebra: Extending Wiedemann’s Analysis of a Sparse Matrix Preconditioner for Computations over Small Fields

Wayne Eberly
Department of Computer Science
University of Calgary
Calgary, Alberta, Canada, T2N 1N4
eberly@ucalgary.ca

Abstract

Wiedemann’s paper, introducing his algorithm for sparse and structured matrix computations over arbitrary fields, also presented a pair of matrix preconditioners for computations over small fields. The analysis of the second of these is extended here in order to provide more explicit statements of the expected number of nonzero entries in the matrices obtained as well as bounds on the probability that the matrices being considered have maximal rank. It is hoped that this will make Wiedemann’s second preconditioner of more practical use.

This is part of ongoing work to establish that this matrix preconditioner can be used to bound the number of nontrivial nilpotent blocks in the Jordan normal form of a preconditioned matrix, in such a way that one can also sample uniformly from the null space of the originally given matrix. If successful this will result in a black box algorithm for the type of matrix computation required when using the number field sieve for integer factorization that is provably reliable (unlike some heuristics, presently in use) and — by a small factor — asymptotically more efficient than alternative provably reliable techniques that make use of other matrix preconditioners or require computations over field extensions.

1 Introduction

Suppose that $F = \mathbb{F}_q$ is a finite field with size $q$. Let $m$ and $n$ be integers such that $0 \leq m \leq n$. The paper that introduced Wiedemann’s algorithm [3] also includes a proof of the following claim — which concerns an $n \times n$ matrix obtained by appending an additional set of row vectors to a matrix $A$, with entries in $F$, $m$ rows, and $n$ columns, and with maximal rank:

\textbf{Theorem 1′ [Wiedemann]}: Numbers $\epsilon > 0$ and $c_1$ exist, both independent of $q$, with the following property: For any integers $n > m \geq 0$ a random procedure exists for generating $n - m$ row vectors with length $n$ such that if $A$ is an $m \times n$ matrix of rank $m$, then with probability at least $\epsilon$, the resulting $n \times n$ matrix is nonsingular and the total Hamming weight of the generated rows is at most $1 + c_1 n \log n$.

Unfortunately the unknown constants $\epsilon$ and $c_1$ are neither supplied nor estimated. Furthermore, it seems that if the proof in [3] is applied without change in order to determine these values then either $c_1$ must be so large or $\epsilon$ so tiny that the result is of limited practical interest.

This can be somewhat rectified: Wiedemann’s analysis can be modified to relax bounds on various parameters and to simplify the estimation of the unknown parameters $\epsilon$ and $c_1$.

1Wiedemann attributes much of the proof of this claim to an anonymous referee who is thanked for allowing this work to be included in Wiedemann’s report.
Theorem 1. Let $\mathbb{F} = \mathbb{F}_q$ be the finite field with size $q$. Let $m$ and $n$ be integers such that $n \geq m \geq 0$. Let $\ell$ be a nonnegative integer and let $\sigma$, $\tau$ and $v$ be positive constants (depending on $q$, but independent of $n$ and $m$) as given in Figure 1.

A random procedure exists such that if $A$ is an $m \times n$ matrix of rank $m$, then an additional $n - m + \ell$ rows (each with length $n$) are produced, and the expected number of nonzero entries in these rows is $\sigma n \ln n + \tau n$ if $n - m \geq v$, and at most $\frac{q-1}{q} (n - m + \ell)n \leq \frac{q-1}{q} (v + \ell)n$, otherwise.

If $q \leq n^2$ then the matrix $B \in (n + \ell) \times n$ obtained from $A$ by adding these rows has maximal rank $n$ with probability at least $\frac{q}{10}$. If $q > n^2$ then this matrix has maximal rank with probability at least $\frac{q}{10} - \frac{q}{10n}$.

This result was obtained by modifying Wiedemann’s description of the matrix $B$, by adding a constant number, $\ell$, of additional rows to this matrix in order to increase the probability that $B$ has maximal rank $n$ when $q$ is small. Treating extremely small finite fields as special cases allowed various restrictions on constants to be relaxed. Perhaps the most significant change made to Wiedemann’s analysis concerned the identification of the constant $\sigma$, for various choices of $q$; one can set $\sigma = \frac{q-1}{q} c$, for a positive constant $c$ such that $c > 6$ and

$$\frac{1}{\beta^2 (1 - \beta)^{1 - \beta}} f(q) \leq \beta^\beta \quad \text{when} \quad 0 < \beta \leq \frac{6}{c}$$

where

$$f(x) = (x-1)^\beta \left( x^{-1} + (1-x)^{-1} \left( \frac{\beta}{x} \right)^{\beta/3} \right).$$

A second result establishes that the constant $\sigma$, mentioned above, can be made arbitrarily close to $6 \left(1 - \frac{1}{q}\right)$, provided that the minimum field size $q$ and the constant $v$ are both increased — at the cost of increasing the constants $\ell$, $\tau$ and $v$ that are listed.

Theorem 2. Let $N$ be an integer such that $N \geq 18$. Let $\mathbb{F}_q$ be a finite field with size $q \geq 16N + 9$. Let $m$ and $n$ be integers such that $n \geq m \geq 0$. Let $\sigma = \left(1 - \frac{1}{q}\right) \cdot \left(6 + \frac{3}{N}\right)$, $\ell = \left\lfloor \frac{53}{10 \ln q} \right\rfloor$ if $q \leq 201$ and let $\ell = 0$ otherwise, and let $v = \left\lceil (2N + 1) \ln(2N + 1) + \frac{167}{N} \ln(2N + 1) \right\rceil + \ell$. Finally, let $\tau = \left\lceil \frac{6}{\ln q} \right\rceil + \ell$.

A random procedure exists such that if $A$ is an $m \times n$ matrix of rank $m$, then an additional $n - m + \ell$ rows (each with length $n$) are produced, and the expected number of nonzero entries in this row is $\sigma n \ln n + \tau n$ if $n - m \geq v$, and at most $\frac{q-1}{q} (n - m)n \leq \frac{q-1}{q} v n$, otherwise.

If $q \leq n^2$ then the matrix $B \in \mathbb{F}_q^{(n+\ell) \times n}$ obtained from $A$ by adding these rows has maximal rank $n$ with probability at least $\frac{q}{10}$. If $q > n^2$ then this matrix is nonsingular with probability at least $\frac{q}{10} - \frac{q}{10n}$.
In both cases, the probability bounds listed here are quite arbitrary. Probability bounds that are closer to one can be obtained by applications of the same techniques, at the cost of increasing the values of the constants $\ell$, $\tau$ and $\nu$.

A system providing plotting and high-precision numerical computations proved to be invaluable here: A plotting function was extremely useful to estimate a constant $c$ that could be selected for each small field size $q$. While other techniques were used to verify these constants, these also depended on high precision computations. In particular (while other systems could certainly have been used here too) the computer algebra system Maple was used and certainly aided this research.

A draft of a report [2], including a full proof of Theorems 1 and 2, is now available.

In a previous study of matrix preconditioners [1], it was noted that this type of preconditioner can be applied to a matrix $A \in F^{n\times n}$ to reduce the number of nontrivial nilpotent blocks in the Jordan normal form of a conditioned matrix, providing that the rank $r$ is known. Future goals for this research include a removal of the requirement that the rank of the input matrix is known, in advance, before it can be used for system solving. This would allow one to sample uniformly from the null space of the originally given matrix, as needed to support sieving-based applications in computational number theory.

References

[1] L. Chen, W. Eberly, E. Kaltoven, B. D. Saunders, W. J. Turner, and G. Villard. Efficient matrix preconditioners for black box linear algebra. *Linear Algebra and Its Applications*, 343–344:119–146, 1993.

[2] W. Eberly. Black box linear algebra: Extending Wiedemann’s analysis of a sparse matrix preconditioner for computations over small fields. Available at [http://www.cpsc.ucalgary.ca/~eberly/Research/sparse_conditioner.pdf](http://www.cpsc.ucalgary.ca/~eberly/Research/sparse_conditioner.pdf), 2016.

[3] D. H. Wiedemann. Solving sparse linear equations over finite fields. *IEEE Transactions on Information Theory*, 32:54–62, 1986.