Probabilistic Analysis of Wiedemann's Algorithm for Minimal Polynomial Computation

Gavin Harrison, Jeremy Johnson, B. David Saunders
Department of Computer Science
Drexel University, University of Delaware
gmh33@drexel.edu,jjohnson@cs.drexel.edu,saunders@udel.edu

Blackbox algorithms for linear algebra problems start with one sided (Lanczos) or two sided (Wiedemann) projection of the sequence of powers of a matrix to a sequence of scalars or a sequence of smaller matrices. Such algorithms usually require that the minimal polynomial of the resulting sequence should be that of the given matrix. Exact formulas are given for the probability that this occurs based on the Jordan structure of a matrix, and from these formulas sharp bounds follow. The bounds are valid for all finite field sizes and show that a small blocking factor can give high probability of success for all cardinalities and matrix dimensions.

Let K be a finite field with cardinality q. Given $A \in K^{n\times n}$, and $\bar{A}$ be the linearly generated sequence $\{I, A, A^2, \ldots\}$. Given $U, V \in K^{n\times b}$ whose elements are selected uniformly randomly from K, $U^T\bar{A}V$ is a linearly generated sequence of smaller matrices, and with high probability, the minimal generating polynomial of $U^T\bar{A}V = \{U^TV, U^TAV, U^TA^2V, \ldots\}$ is the minimal polynomial of the matrix A. All square matrices are similar to a generalized Jordan form matrix, $A = PJP^{-1}$, where $J, P \in K^{n\times n}$. If U and V are selected uniformly randomly, then $X = P^TU$ and $Y = P^{-1}V$ are also uniformly random. $U\bar{A}V = X^T\bar{J}Y = \{X^TY, X^TJY, X^TJ^2Y, \ldots\}$, and $X\bar{J}Y$ has the same probability as $U\bar{A}V$ of having its minimal generating polynomial match the minimal polynomial of A and J. We call this probability $Prob_{q,b}(A)$.

Let $C_f \in K^{d\times d}$ represent the companion matrix for the polynomial $f(x) = f_0 + f_1x + \ldots + f_{d-1}x^{d-1} + x^d$ with coefficients in K. Let J_{f^e} be the generalized Jordan block of an irreducible f occurring with multiplicity e. Since $Prob_{q,b}(J_f \oplus J_g) = Prob_{q,b}(J_f)Prob_{q,b}(J_g)$ when $\gcd(f, g) = 1$, unique irreducibles can be treated separately, and for each irreducible only its highest multiplicity affects the probability that the projection preserves the minimal polynomial. Furthermore we show $Prob_{q,b}(J_f) = Prob_{q,b}(J_{f^e})$ for any e. Therefore, letting $T = \{(f_1, e_1, t_1), (f_2, e_2, t_2), \ldots\}$, where the polynomials f_i are the irreducibles occuring in the invariant factors of A, e_i is the highest multiplicity of f_i, and t_i is the number of occurrences of $f_i^{e_i}$, it follows that

$$Prob_{q,b}(J) = \prod_{k=1}^{|T|} Prob_{q,b}\left(\bigoplus_{t_k} J_{f_k}\right).$$

For an irreducible polynomial f of degree d, the probability that $U^T\bar{C}_fV$ has minimal polynomial f is easy to determine. The minimal polynomial of the projection is always a factor of f, which for irreducible f is 1 or f. It is 1 only if the sequence is a sequence of zero matrices, which is to say that one of U, V is zero. Thus

$$Prob_{q,b}(C_f) = (1 - q^{-db})^2.$$

If $J = \bigoplus_t C_f$, then for $U, V \in K^{dt\times b}$, with blocking conformal to the diagonal blocks of J, we have $U^T\bar{J}V = \sum_{k=1}^{t} U_k^T\bar{C}_fV_k$. We show $Prob_{q,b}(C_f) \leq Prob_{q,b}\left(\bigoplus_t J\right)$.

	field cardinality			
block size	2	3	10007	$2^{31}-1$
1	0.000467	0.00112	0.0499	0.911
2	0.25	0.444	$1-2\times10^{-4}$	$1-4.3\times10^{-11}$
4	0.766	0.927	$1-2\times10^{-12}$	$1-9.4\times10^{-30}$
8	0.984	$1-9.1\times10^{-4}$	$1-2\times10^{-28}$	$1-4.4\times10^{-67}$
16	$1-6\times10^{-5}$	$1-1.4\times10^{-7}$	$1-2\times10^{-60}$	$1-9.8\times10^{-142}$
32	$1-9.3\times10^{-10}$	$1-3.2\times10^{-15}$	$1-2\times10^{-124}$	$1-4.8\times10^{-291}$

Table 1: Bounds for worst case probability of success to preserve minimum polynomial, matrix size $10^8 \times 10^8$

It is evident that the probability of success increases with d as well as with b. The worst case is a matrix whose minimal polynomial is a distinct product of the smallest possible irreducibles. This yields an exact lower bound formula for the probability that a projection $U^T\bar{A}V$ of A has the same minimal polynomial. Let $L_q(d,n)$ be the number of degree d irreducible factors over the finite field of cardinality q that fit in a matrix of dimension n after all smaller degree irreducibles have been inserted. Then, for an $n \times n$ matrix A,

$$Prob_{q,b}(A) \geq \prod_{d=1}^{\infty} \left(1 - \frac{2q^{db}-1}{q^{2db}}\right)^{L_q(d,n)}$$

We compare this bound to previously given lower bounds in the case when field cardinality and matrix dimension are of similar size. For small primes, Wiedemann (proposition 3) treats the case $b = 1$ and he fixes the projection on one side because he is interested in linear system solving and thus in the sequence $\bar{A}b$ [2] . For small q, his formula, $1/(6\log_q(N))$, computed with some approximation, is nonetheless quite close to our exact formula. However as q approaches N the discrepancy with our exact formula increases. At the large/small crossover, $q = N$, Kaltofen/Pan's lower bound is 0, Wiedemann's is 1/6, and ours is $1/e$. The Kaltofen/Pan probability bound improves as q grows larger from N [1] . The Wiedemann bound becomes more accurate as q goes down from N. But the area $q \approx N$ is of some practical importance. In integer matrix algorithms where the finite field used is a choice of the algorithm, sometimes practical considerations of efficient field arithmetic encourages the use of primes in the vicinity of N. For instance, exact arithmetic in double precision and using BLAS works well with $q \in 10^6..10^7$. Sparse matrices of order N in that range are tractable. Our bound may help justify the use of such primes.

But the primary value we see in our analysis here is the understanding it gives of the value of blocking, $b > 1$. Table 1 shows the bounds for the worst case probability that a random projection will preserve the minimal polynomial of a matrix $A \in K^{10^8 \times 10^8}$ for various fields and projection block sizes. It shows that the probability of finding the minimal polynomial correctly under projection converges rapidly to 1 as the projected block size increases. Even over $GF(2)$, with block size $b = 16$ the probability is very good.

References

[1] Erich Kaltofen and B. David Saunders. On wiedemann's method of solving sparse linear systems. In *Proceedings of the 9th International Symposium, on Applied Algebra, Algebraic Algorithms and Error-Correcting Codes*, AAECC-9, pages 29–38, London, UK, UK, 1991. Springer-Verlag.

[2] D. Wiedemann. Solving sparse linear equations over finite fields. *Information Theory, IEEE Transactions on*, 32(1):54–62, 1986.