10-1-1988

Cycle Lengths In $A^k b$

Charles M. Grinstead
Swarthmore College, cgrinst1@swarthmore.edu

Recommended Citation
Charles M. Grinstead. (1988). "Cycle Lengths In $A^k b$". *SIAM Journal On Matrix Analysis And Applications*. Volume 9, Issue 4. 537-542.
http://works.swarthmore.edu/fac-math-stat/59

SIAM J. MATRIX ANAL. APPL.
Vol. 9, No. 4, October 1988

007

CYCLE LENGTHS IN $A^k b$*

CHARLES M. GRINSTEAD†

Abstract. Let A be a nonnegative, $n \times n$ matrix, and let b be a nonnegative, $n \times n$ vector. Let S be the sequence $\{A^k b\}$, $k = 0, 1, 2, \cdots$. Define $m(A, b)$ to be the length of the cycle of zero-nonzero patterns into which S eventually falls. Define $m(A)$ to be the maximum, over all nonnegative b of $m(A, b)$. Finally, define $m(n)$ to be the maximum, over all nonnegative, $n \times n$ matrices A of $m(A)$. This paper shows given A and b, that $m(A, b)$ is a divisor of a certain number, which is determined by the structure of A and b. It is also shown that $\log m(n) \sim (n \log n)^{1/2}$.

Key words. positive matrices, symmetric group, Prime Number Theorem

AMS(MOS) subject classifications. 15A48, 10H25

Let A be a nonnegative, $n \times n$ matrix, and let b be a nonnegative, $n \times 1$ vector. In this paper, we are concerned with the zero-nonzero patterns in the sequence $S = \{A^k b\}$, $k = 0, 1, 2, \cdots$. Since there are 2^n possible patterns for an $n \times 1$ vector, the sequence S must eventually fall into a cycle, and the length of the cycle is at most 2^n. We define $m(A, b)$ to be the length of this cycle. We also define $m(A)$ to be the maximum, over all b of $m(A, b)$. Finally, we define $m(n)$ to be the maximum, over all nonnegative, $n \times n$ matrices A of $m(A)$. Given A and b, we will show that $m(A, b)$ is a divisor of a certain number, which is determined by the structure of A and b. We will also show that

$$\log (m(n)) \sim \sqrt{(n \log n)}.$$

Each nonnegative, $n \times n$ matrix A corresponds to a directed graph $G(A)$ with vertices $1, 2, \cdots, n$, and with an edge from j to i if $a_{ij} > 0$. (If $a_{ii} > 0$, then there is a loop at vertex i.) This is *not* the usual definition. However, the present definition aids in the exposition. We note that our graph could be obtained by applying the usual definition to the transpose of A. Each nonnegative vector b corresponds to a subset $P(b)$ of vertices, defined by $i \in P(b)$ if and only if $b_i > 0$. Given a vector b, the set corresponding to Ab is the set of all vertices of distance one from $P(b)$, i.e., all vertices j such that there is a vertex $i \in P(b)$ and an edge (i, j) in $G(A)$. If we define

$$\Gamma^{(k)}(v) = \{w: \text{there is a path of length } k \text{ from } v \text{ to } w\},$$
$$\Gamma^{(k)}(T) = \bigcup_{v \in T} \Gamma^{(k)}(v),$$

then we have

$$P(A^k b) = \Gamma^{(k)}(P(b)).$$

A subgraph H of a graph G is said to be *strongly connected* if, for any two (not necessarily distinct) vertices in H, there is a path in H from each vertex to the other. A subgraph is a *strongly connected component* (scc) if it is a maximal strongly connected subgraph of G. It is easy to show that the scc's are pairwise disjoint. They need not, however, partition the graph.

In terms of matrices, given A, we let P be a permutation matrix such that PAP^T is in block lower triangular form, with square matrices on the diagonal. If no such P exists,

* Received by the editors June 9, 1986; accepted for publication (in revised form) March 29, 1988. This research was partly supported by National Science Foundation grant DMS-8406451.

† Swarthmore College, Swarthmore, Pennsylvania 19081.

other than $P = I$, then the matrix A is said to be irreducible. If we find P such that each diagonal block is irreducible, then the diagonal blocks of size greater than 1×1, together with the 1×1 nonzero diagonal blocks, correspond to the scc's of $G(A)$.

LEMMA 1. *Let $S_1, S_2, \cdots$ be a sequence of subsets of a finite set S, and let d be a positive integer. Suppose that for all j and for all sufficiently large h, we have $S_j \subset S_{j+hd}$. Then the sequence has an eventual cycle whose length divides d.*

Proof. Let j be any nonnegative integer less than d. Let S^j be the set of all $v \in S$ such that $v \in S_{j+hd}$ for some $h \geqq 0$. For each $v \in S^j$, there exists an integer h_v such that $v \in S_{j+hd}$ for all $h \geqq h_v$. Let h_0 be the maximum of the h_v, taken over all $v \in S^j$. Then $S_{j+hd} = S^j$ for all $h \geqq h_0$. Thus, the sequence has an eventual cycle consisting of S^0, $S^1, \cdots, S^{(d-1)}$. This implies that the sequence has an eventual cycle the length of which divides d. □

We define the *index* of a graph to be the greatest common divisor of the lengths of the circuits in the graph. The index of a nonnegative matrix A is then defined to be the index of $G(A)$. (We note that this is not the usual definition of the index of a matrix.) Let b be a nonnegative vector, and for $j \geqq 0$ denote $P(A^j b)$ by P_j. We first examine $m(A, b)$ in the case that A is irreducible.

LEMMA 2. *Let A be a nonnegative, irreducible $n \times n$ matrix with index d. Then, for all b, $m(A, b)$ divides d.*

Proof. It is well known (see [5, Thm. 2.9, p. 49]) that the greatest common divisor of the lengths of the circuits through any vertex v is d, independent of the vertex v. It is also well known that if we call these circuit lengths $c_1, c_2, \cdots, c_j$, then there is a multiple of d, say $N_v d$, such that every multiple of d greater than or equal to $N_v d$ can be written as a nonnegative linear combination of the $\{c_i\}$. If we let $N = \max_{v \in G} N_v$, then every vertex in G is on a circuit of length hd, for all $h \geqq N$. Thus, if $v \in P_j$, then $v \in P_{j+hd}$ for all sufficiently large h. Then Lemma 1 applies. This completes the proof. □

LEMMA 3. *Let A be a nonnegative, $n \times n$ matrix, and let C be an* scc *in $G(A)$. Let b be a nonnegative vector, and, as before, let $P_j = P(A^j b)$ and $P_0 = P(b)$. Let the index of C be d. Then the sequence $\{P_j \cap C\}$ eventually repeats with a cycle length that divides d.*

Proof. Let j be a positive integer, and let v be any vertex in $P_j \cap C$. Since C is an scc with index d, v is on a circuit in C of length hd, for all sufficiently large h. This means that v is in $P_{j+hd} \cap C$ for all sufficiently large h. Thus, from Lemma 1, we see that the sequence $\{P_j \cap C\}$ eventually repeats with a cycle length that divides d. □

THEOREM 1. *Let A be a nonnegative, $n \times n$ matrix, and let $G(A)$ have* scc*'s C_1, $C_2, \cdots, C_k$. Let b be any nonnegative vector. Suppose that the sequence $\{P_j \cap C_i\}$ has an eventual cycle of length r_i. Then $m(A, b)$ equals the least common multiple of the r_i.*

Proof. Let r equal the least common multiple (lcm) of the r_i. First we show that $m(A, b)$ divides r. To show this, we shall show that for all vertices v in $G(A)$, the sequence $\{P_j \cap \{v\}\}$ has an eventual cycle whose length divides r.

Let $v \in C_i$ for some i. If $v \in P_k$ for some $k \geqq 0$, then for some positive integer j and for all sufficiently large h, we have $v \in P_{j+hr}$, since r is a multiple of r_i. So Lemma 1 implies that the sequence $\{P_j \cap \{v\}\}$ has an eventual cycle whose length divides r. If $v \neg\in P_k$ for all $k \geqq 0$, then the sequence $\{P_j \cap \{v\}\}$ has a cycle of length 1.

Now suppose that v is not an element of any C_i. Let $C_{i1}, C_{i2}, \cdots, C_{ik}$ be the scc's containing vertices that have paths to v. If $k = 0$, then v is in no P_i with $i \geqq n$. Finally, if $k > 0$, then for $j \geqq n$, $v \in P_j$ if and only if there is a v^* in $C_{is} \cap P_t$, for some $s \leqq k$, and a path of length $(j - t)$ from v^* to v. (We need to assume that $j \geqq n$ because, for smaller values of j, the fact that $v \in P_j$ could be due to v being at the end of a path of length j from a vertex in P_0 not in any C_i.) Since $v^* \in C_{is}$, the eventual cycle in the sequence $\{P_j \cap \{v^*\}\}$ has a length that divides r_{is}. This means that the contribution of v^* to the eventual cycle in the sequence $\{P_j \cap \{v\}\}$ has a length that divides r_{is}. Since the contributions of all other vertices in $C_{i1}, C_{i2}, \cdots, C_{ik}$ to the sequence $\{P_j \cap \{v\}\}$ are similar, we see that the sequence $\{P_j \cap \{v\}\}$ has an eventual cycle whose length certainly divides r. So we have shown that for all vertices in $G(A)$, either they appear in none of the P_j for sufficiently large j, or they appear with a pattern having a length that divides r. This means that $m(A, b)$ divides r.

Since the eventual cycle of the sequence $\{P_j \cap C_i\}$ is of length r_i, it is clear that r_i must divide $m(A, b)$. $\square$

COROLLARY 1. *Let A be a nonnegative, $n \times n$ matrix, and let $G(A)$ have* scc*'s C_1, $C_2, \cdots, C_l$ with indices $d_1, d_2, \cdots, d_l$, respectively. Let b be any nonnegative vector. Let T be the set of* scc*'s that intersect at least one element of the sequence $\{P_k\}$. Let d_T equal the least common multiple of the set of d_i's corresponding to the C_i's in T. Then $m(A, b)$ divides d_T.*

Proof. If $C_i \in T$, and we let r_i equal the length of the eventual cycle of the sequence $\{P_j \cap C_i\}$, then using Lemma 3, we know that r_i divides d_i. If $C_i \neg\in T$, then the length of the eventual cycle of the sequence $\{P_j \cap C_i\}$ is 1. Thus, since $m(A, b)$ equals the lcm of the r_i's corresponding to the C_i's in T, we see that $m(A, b)$ divides d_T. $\square$

We now turn our attention to the maximization of $m(A)$ over all nonnegative $n \times n$ matrices A. Given positive integers $d_1, d_2, \cdots, d_k$, with sum n, it is possible to construct a nonnegative, $n \times n$ matrix A and a nonnegative vector b such that $m(A, b) = \text{lcm}\,(d_1, d_2, \cdots, d_k)$. We accomplish this by letting $G(A)$ be the disjoint union of circuits of lengths $d_1, d_2, \cdots, d_k$, and letting P_0 be a set containing exactly one vertex from each circuit.

Next, we note that $d_i \leqq |C_i|$ for each i, and that

$$\sum_{i=1}^{k} |C_i| \leqq n,$$

so we know that

$$\sum_{i=1}^{k} d_i \leqq n.$$

Thus, we wish to maximize the lcm of $d_1, d_2, \cdots, d_k$ over all sets of positive integers with sum not exceeding n. The maximum value will be $m(n)$. Let us call a set $\{d_i\}$ whose lcm is this maximum value an n-extremal set. We note that this problem can be stated as follows. Among all elements of the symmetric group S_n, which elements have the largest order, and what is their order? In terms of the original problem, the group elements of largest order are those that, when written as a product of disjoint cycles, have cycle lengths forming an n-extremal set. The order of these elements is $m(n)$. This problem was studied by Landau (see [3, Vol. 1, pp. 222–229]), who proved Theorem 2 below (see also [4]). We will give a shorter proof of this result.

LEMMA 4. *For every $n \geqq 1$, there exists an n-extremal set X such that each element of X is either a prime power or the number* 1.

Proof. Assume that X is an n-extremal set containing an integer r, which is neither 1 nor a prime power. Then r is divisible by at least two different primes, p and q. Suppose that the powers of p and q appearing in the prime factorization of r are p^y and q^z. In X, if we replace r by the integers p^y, q^z, and (r/p^yq^z), then the lcm remains unchanged, and the sum of the elements of X decreases by the quantity

$$\Delta = r - \left(p^y + q^z + \left(\frac{r}{p^y q^z}\right)\right).$$

It remains to show that $\Delta \geqq 0$. We have

$$\Delta = r\left(1 - \frac{1}{p^y q_z}\right) - p^y - q^z$$
$$\geqq r\frac{5}{6} - p^y - q^z.$$

Since p^y and q^z divide r,

$$\Delta \geqq r\frac{5}{6} - \frac{r}{p^y} - \frac{r}{q^z}$$
$$= r\left(\frac{5}{6} - \frac{1}{p^y} - \frac{1}{q^z}\right)$$
$$\geqq r\left(\frac{5}{6} - \frac{1}{2} - \frac{1}{3}\right)$$
$$= 0. \qquad \square$$

Let p_i denote the ith prime. At first glance, it might seem that an n-extremal set should consist of $p_1, p_2, \cdots, p_k$, where k is the largest prime such that the sum of the first k primes does not exceed n. However, it is easy to show that this is not, in general, the best way to proceed. As an example, suppose that n is the sum of all of the primes not exceeding the prime 1231. The numbers 2, 3, 5, and 1231 have the same sum as the numbers 2^9 and 3^6, but the lcm of the second set is larger than the lcm of the first set. So, by replacing the first set with the second set, we obtain a set with a larger lcm.

It is nevertheless the case that by taking the first k primes, we obtain a set whose lcm has, asymptotically, the same logarithm as $m(n)$.

THEOREM 2. *Given a positive integer n, let k be the largest integer such that*

$$\sum_{i=1}^{k} p_i \leqq n.$$

Then, $\log(m(n)) \sim \sum_{i=1}^{k} \log(p_i)$. *Furthermore*, $k \sim 2\sqrt{(n)}/\sqrt{(\log(n))}$, *so*

$$\log(m(n)) \sim (\sqrt{n})(\sqrt{(\log(n))}).$$

Before proving this theorem, we need the following lemma.

LEMMA 5. *Let T and T' be two sets of real numbers with the following properties:*

(i) *Each element of T is less than each element of T'.*

(ii) *Every element of both sets is at least as large as e.*

(iii) *The sum of the elements in T is at least as large as the sum of the elements of T'.*

Then the product of the elements of T is at least as large as the product of the elements of T'.

Proof. Let B be a real number at least as large as each element of T and less than each element of T'. We note that B can be chosen to be at least e. Let S and S' be the sums of the elements in the sets T and T', respectively. Let P and P' be the products of the elements in the sets T and T', respectively.

First, fix $|T'| = k$. If two elements of T' are unequal, we can make P' larger without affecting S', by replacing each of the two elements by their average. Thus, we may assume that all of the elements in T' are equal. Then their common value is (S'/k), and $P' = (S'/k)^k$. Since each element in T' is greater than B, it is easy to show that $P' \leqq B^{(S'/B)}$.

If q is any real number such that $e \leqq q \leqq B$, then it is easy to check that $q \geqq B^{(q/B)}$. Thus, if the elements of T are $q_1, q_2, \cdots, q_k$, then the product of the elements of T is at least $(B^{(q_1/B)})(B^{(q_2/B)}) \cdots (B^{(q_k/B)})$, which equals $B^{(S/B)}$. Since $S \geqq S'$, we have $P \geqq P'$, which completes the proof. □

Proof of Theorem 2. The Prime Number Theorem implies that $p_i \sim i \log(i)$ (see [2, p. 10]). Using this, it is easy to show that

$$\sum_{i=1}^{k} p_i \sim \left(\frac{1}{2}\right) k^2 \log(k).$$

Thus, we have $n \sim (\frac{1}{2})k^2 \log(k)$, which implies that $\log(k) \sim (\frac{1}{2}) \log(n)$. Hence,

$$k \sim 2(\sqrt{(n)})/(\sqrt{(\log(n))}).$$

We note that this implies that

$$p_k \sim (\sqrt{(n)})(\sqrt{(\log(n))}).$$

Now assume for the moment that n is the sum of the first k primes. Let S be the set of primes not exceeding p_k, and let S' be an n-extremal set. The sum of the elements in S' is then less than or equal to the sum of the elements in S, which equals n. Let T' be the set of all elements of S' which are powers of primes p_j such that $j > k$. Let T be the set of all primes p_i in S such that no power of p_i appears in S'. Since each prime in $S - T$ appears to the first power in $S - T$, and appears to at least the first power in $S' - T'$, and since the sum of the elements in S is at least as great as the sum of the elements in S', we must have that

$$\sum_{q_j \in T'} q_j \leqq \sum_{p_i \in T} p_i,$$

where each q_j in the left-hand sum represents a prime power. We further note that each q_j in T' is greater than p_k, and that each p_i in T is less than or equal to p_k. We now note that Lemma 5 applies, except that one of the p_i in T might be the prime 2. If we temporarily change it to a 3, then, using Lemma 5, we see that the product of the elements of T is at least as great as the product of the elements in T'. Changing the 3 back to a 2 certainly does not affect the dominant term in the estimation for the logarithm of the product of the elements in S'. Thus in S', if the elements in T' are replaced by the elements in T, the product of the elements of the new S' is at least two-thirds as large as the product of the elements in the old S', hence we may assume that S' contains no powers of any prime greater than p_k. At this point we emphasize that S' may have a sum that exceeds n. Nevertheless, each element in S' is less than or equal to n. Using this fact, we shall show that the logarithm of the product of the elements in S' does not exceed $\sqrt{(n)(\log n)}$. We estimate:

$$\begin{aligned}
\log\left(\prod_{q_i \in S'} q_i\right) &= \log\left(\prod_{p_i \leqq \sqrt{n}} p_i^{\alpha_i}\right) + \log\left(\prod_{\substack{p_i > \sqrt{n} \\ i \leqq k}} p_i\right) \\
&\leqq \log(n^{\pi(\sqrt{n})}) + \sum_{\substack{p_i > \sqrt{n} \\ i \leqq k}} \log(p_i) \\
&\sim \frac{\sqrt{n}(\log n)}{\log \sqrt{n}} + \sqrt{n}\sqrt{\log n} - \sqrt{n} \\
&\sim \sqrt{n}\sqrt{\log n}.
\end{aligned}$$

We now show that this bound is achieved by the elements in S. We have

$$\sum_{i=1}^{k} \log(p_i) \sim p_k$$

(see [2, Thms. 420, 434]). Also, since $n = \sum_{i=1}^{k} p_i$ and $p_i \sim i(\log i)$, it is easy to show that

$$\begin{aligned} p_k &\sim (2\sqrt{(n)}/\sqrt{(\log n)})\log(2\sqrt{(n)}/\sqrt{(\log n)}) \\ &\sim \sqrt{n}\sqrt{\log n}. \end{aligned}$$

Thus, if n is the sum of the first k prime numbers, then

$$\log(m(n)) \sim \sqrt{n}\sqrt{\log n}.$$

Finally, we relax the assumption on n. Assume instead that

$$n_1 = \sum_{i=1}^{k} p_i < n \leqq \sum_{i=1}^{k+1} p_i = n_2.$$

Since $m(n)$ is clearly monotonic in n, we have $m(n_1) \leqq m(n) \leqq m(n_2)$, but it is easy to check that $m(n_1) \sim m(n_2)$, which completes the proof. $\square$

A comment on the conclusion of the theorem is in order. While it would be nicer to obtain a function to which $m(n)$ is asymptotic, it seems unlikely that this can be done. The task would require an estimation similar to the estimation of the product of the first k primes. Let P be this product. We would have to obtain an estimate of the following form:

$$\sum_{i=1}^{k} \log(p_i) = f(k) + o(1),$$

for then we would be able to say that $P \sim e^{f(k)}$. Although it is known that the above sum is asymptotic to p_k, at this time we cannot even say that the error term is $O(p_k^\delta)$ for even one value of $\delta < 1$.

Acknowledgement. The author thanks the referees for their numerous helpful observations.

REFERENCES

[1] P. G. Coxson and L. Larson, *Monomial patterns in the sequence $A^k b$*, preprint.

[2] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Clarendon Press, Oxford, 1960.

[3] E. Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*, *Teubner*, Leipzig, Stuttgart, 1909.

[4] W. Miller, *The maximum order of an element in a finite symmetric group*, Amer. Math. Monthly, 94 (1987), pp. 497–506.

[5] R. S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1962.