

THE RANK AND SIZE OF GRAPHS

Andrew Kotlov and László Lovász

DEPT. OF MATHEMATICS, YALE UNIVERSITY, NEW HAVEN, CT 06520

Abstract

We show that the number of points with pairwise different sets of neighbors in a graph is $O(2^{r/2})$ where r is the rank of the adjacency matrix. We also give an example that achieves this bound.

1 Introduction

It was conjectured in 1976 by C. van Nuffelen [vN] that the chromatic number of any graph with at least one edge does not exceed the rank of its adjacency matrix. Interest in this conjecture was increased substantially when it was re-discovered, independently, by a computer (Fajtlowicz 1988 [F]), and, perhaps even more significantly, when its connection to a fundamental problem in the theory of communication complexity was observed (see, e.g., [L]).

The first counterexample to van Nuffelen's conjecture was obtained in 1989 by Alon and Seymour [A-S]. They constructed a graph with chromatic number 32 and with an adjacency matrix of rank 29. In 1992 Razborov showed that the gap between the chromatic number and the rank of the adjacency matrix was superlinear [R]. In 1993 Raz and Spieker [R-S] proved that the chromatic number was not bounded by any polynomial in the rank. The best result known so far is due to Nisan and Wigderson [N-W]. It gives (if translated to these terms) an infinite family of graphs with rank r and with the chromatic number $\chi = 2^{\Omega(\log r)^\alpha}$ where $\alpha = \log_3 6 > 1$.

In this paper we investigate the upper bound on the number of points of a graph in terms of the rank, r , of its adjacency matrix. Such a bound exists as soon as we make the assumption that the graph has no *twin points*, i.e., non-adjacent points with the same neighborhood. In this case, it is trivial to prove that the number of points is at most 2^r . We improve this trivial upper bound to $O(2^{r/2})$, and show that this new bound is best possible.

Clearly, this gives an upper bound on the chromatic number as well. Using Ramsey's Theorem this bound can be improved by a factor of $\frac{1}{r}$ (we are grateful to the referee for this remark). Recently, an upper bound of $O((1.36\dots)^r)$ has been proved [K]. We conjecture,

however, that the chromatic number is bounded by a subexponential function of the rank.

2 The Upper Bound

Let G be a graph on n vertices, and M , its adjacency matrix. Denote $r = \text{rk}(M)$. Two vertices u and v of G are called *twins* if their sets of neighbors coincide:

$$\Gamma(u) = \Gamma(v).$$

(In particular, twins are non-adjacent.) In M twins correspond to identical rows (and columns). Hence deleting or adding a twin point does not change either the rank or the chromatic number of a graph. Therefore assume throughout this section that G has no twins. We first state and prove the following easy

Lemma Let H be a largest subgraph of G of a smaller rank, $\text{rk}(H) < r$. Denote $t = |H|$. Then

- (a) no two columns of M coincide on more than t positions;
- (b) H does not have triplets;
- (c) if H has twins, its rank is exactly $r - 2$.

Proof To see (a) suppose that some two columns, i and j , of M coincide on (at least) $t + 1$ positions, say positions 1 through $t + 1$. Then the first $t + 1$ rows of M form a submatrix whose right null-space is strictly larger than that of M , for in particular it contains the vector

$$(0, \dots, 0, \overset{i}{1}, 0, \dots, 0, \overset{j}{-1}, 0, \dots, 0).$$

Consequently, the subgraph of G spanned by $\{v_1, \dots, v_{t+1}\}$ has rank smaller than r , a contradiction with the maximality of t . To see (b) and (c), consider the graph spanned by H and u , where u is some vertex of $G - H$, and the corresponding adjacency matrix, N . Observe that N is of rank r (by the maximality of H) and (hence) it is twin-free. Since deleting a row from a 0-1 matrix without twin columns cannot produce three or more identical columns, (b) follows. Since deleting from N the row that corresponds to u produces twin columns, this row is not in the linear span of the rest of the rows of N , and thus deleting such a row lowers the rank (by 1). Similarly, deleting the corresponding column from the submatrix just obtained (of a symmetric N) lowers the rank by another 1, and (c) follows. \square

Theorem Let G be a twin-free graph on n vertices, r the rank of its adjacency matrix. Then

$$n = O(2^{r/2}).$$

Proof Let H denote a largest subgraph of G of a smaller rank, $rk(H) < r$. Let $t = |H|$.

Case 1 $t \leq 3n/4$.

Along with M , consider the matrix M' obtained from M by replacing all 0's by -1 's. In other words, $M' = 2M - J$ where J denotes the all-1 matrix of the appropriate size. Clearly, $|rk(M') - rk(M)| \leq 1$. One can look at the columns of M' as vectors of euclidean length \sqrt{n} in the (at most) $(r + 1)$ -dimensional space. Every two of them, say u and v , agree on at most t positions (part (a) of the lemma), and disagree on at least $n - t$ positions. Hence $u \cdot v \leq t - (n - t) = 2t - n$. Consequently, the cosine of the angle between u and v is at most $(2t - n)/n \leq 1/2$ (since $t \leq 3n/4$), and so the angle between them is at least 60° . But now we can reduce our problem to the problem of packing unit vectors in \mathbb{R}^{r+1} on the r -dimensional unit sphere S^r in such a way that the angle between any two of them is at least $\varphi = 60^\circ$. The upper bound on the number of vectors packed in this fashion is given by

$$(1 - \cos \varphi)^{-r/2} 2^{-r(0.099+o(1))} \quad \text{as } r \rightarrow \infty,$$

for all $\varphi \leq \varphi^* = 62.9974\dots$ (this result is due to Kabatjanskiĭ and Levenšteĭn [K-L], see also [G-W], p.825). Taking $\varphi = 60^\circ$ we obtain an upper bound of $O(2^{r/2})$, as required. This settles Case 1.

Let C denote a constant bounding from above $2^{-r(0.099+o(1))}$ in the above formula for $\varphi = 60^\circ$. In what follows we prove the inequality $n \leq C \cdot 2^{r/2} - 16$ by induction on r applied to twin-free graphs. We may assume that it is satisfied for $r = 2$ (i.e. that $C \geq 9$).

Case 2 $t > 3n/4$ and H does not have twins.

Then by the induction assumption applied to H , $t \leq C \cdot 2^{(r-1)/2} - 16$ and hence $n < 4t/3 < C \cdot 2^{r/2} - 16$, as required (note that we only need here that $t > n/\sqrt{2}$). This settles Case 2.

Thus for the rest of the proof assume that H has twins (cases 3 and 4 below). Let us denote the number of twin pairs in H by p (recall that by the lemma, H does not have triplets.)

Case 3 H has twins and $2(t - p) \geq n - 16$.

One can consider a twin-free subgraph, H' , of H on $t - p$ points. By part (c) of the lemma, $rk(H') = rk(H) = r - 2$. Hence by induction applied to H' , $n \leq 2(t - p) + 16 \leq 2(C \cdot 2^{(r-2)/2} - 16) + 16 = C \cdot 2^{r/2} - 16$ as required. This settles Case 3.

Case 4 $t > 3n/4$, H has p twin pairs, and $2(t - p) < n - 16$. In particular,

$$p > t - \frac{n}{2} + 8 > \frac{3n}{4} - \frac{n}{2} + 8 > \frac{n}{4} \quad \text{and} \quad n - t \geq n - 2(t - p) > 16. \quad (*)$$

In this final case we will apply the induction argument to the induced subgraph of G spanned by the twins of H . Let the vertices of G be arranged in the following order:

$$\{v_1, \dots, v_p, v'_1, \dots, v'_p, v_{2p+1}, \dots, v_t, \dots, v_n\}$$

where v'_i is the twin of v_i , and the first t points correspond to H . For each $1 \leq i \leq p$, $t+1 \leq j \leq n$ there is exactly one edge between v_j and $\{v_i, v'_i\}$ (part (a) of the lemma.) Let $\{v_i, v'_i\}$ be ordered so that (v_i, v_{t+1}) is an edge for all $1 \leq i \leq p$.

Now note that by the maximality of H , the first $t+1$ rows of M span its row-space. This in particular means that for a given j , $t+2 \leq j \leq n$, either (v_i, v_j) is an edge for all $1 \leq i \leq p$, or (v'_i, v_j) is an edge for all $1 \leq i \leq p$ (depending on the sign of the coefficient of row $t+1$ in the representation of row j as a linear combination of rows $1, \dots, t+1$). Let us reorder vertices v_{t+1}, \dots, v_n so that exactly the first k of them are adjacent to v_i and not to v'_i , $1 \leq i \leq p$ (and hence the other $n-t-k$ are adjacent to v'_i and not to v_i , $1 \leq i \leq p$).

Let A denote the subgraph spanned by $\{v_1, \dots, v_p\}$. We claim that A has no twins. Indeed, if A had twins, then the corresponding columns of M would coincide on the first $2p$ positions, as well as on positions $t+1$ through n . By part (a) of the lemma, we would have $t \geq 2p+(n-t)$ or $2(t-p) \geq n/2$ which contradicts our assumption above. Thus A is twin-free.

We next claim that $\text{rk}(A) \leq r-4$. Assume otherwise. Let D_1 and D_2 denote the submatrices of M spanned by the rows $\{v_{t+1}, \dots, v_{t+k}\}$ and $\{v_{t+k+1}, \dots, v_n\}$ respectively, and the columns $\{v_{2p+1}, \dots, v_n\}$. Observe that $\text{rk}(A) + \text{rk}(D_i) \leq r$, ($i=1, 2$), for the ij -entries of M are zeros if either $t+1 \leq i \leq t+k$ and $1 \leq j \leq p$ or $t+k+1 \leq i \leq n$ and $p+1 \leq j \leq 2p$ (due to the special vertex arrangement just discussed.) Thus $\text{rk}(D_1), \text{rk}(D_2) \leq 3$. Since neither D_1 nor D_2 can have twin rows (otherwise so would M), the total number of rows in D_1 and D_2 is at most $2 \cdot 2^3 = 16$. But this number is exactly $n-t > 16$, cf. (*). A contradiction.

We conclude that $\text{rk}(A) \leq r-4$. Since $|A| = p > n/4$, by induction applied to A

$$n \leq 4(C \cdot 2^{r-4/2} - 16) < C \cdot 2^{r/2} - 16,$$

as required. This concludes the proof of the theorem. \square

3 The Example

We conclude with an example of a graph on $\Omega(\sqrt{2^r})$ vertices. This construction is similar to (but not the same as) the well-known Mycielski graph.

Consider a twin-free graph on n vertices, of rank r . Duplicate each of its vertices. Add another vertex and connect it by an edge to exactly one vertex of each twin pair. The graph obtained is again twin-free, has $2n + 1$ vertices, and is of rank $r + 2$.

Using a single point graph as the base, one concludes easily by induction on (even) r that $n = 2 \cdot \sqrt{2^r} - 1$.

References

[A-S] N. Alon, P. Seymour: A Counterexample to the Rank-Coloring Conjecture. *J. Graph Theory* **13** pp. 523–525, 1989.

[F] S. Fajtlowicz. On Conjectures of Graffiti. *Discrete Math.* 72 (1988), pp. 113–118, North-Holland.

[G-W] P.M. Gruber, J.M. Wills (editors): *Handbook of Convex Geometry, Volume B*. North-Holland, 1993.

[K] A. Kotlov: Rank and Chromatic Number of a Graph, preprint

[K-L] G.A. Kabatjanskiĭ and V.I. Levenšteĭn: Bounds for packings on a sphere and in space (in Russian). *Problemy Peredachi Informatsii* 14, pp. 3–25, 1978.

[L] L. Lovász. Communication complexity: a survey, in: *Paths, flows, and VLSI-Layout*, (ed. B. Korte, L. Lovász, H. J. Prömel, A. Schrijver), Springer, 235–265. **YEAR**

[N-W] N. Nisan and A. Wigderson. On Rank vs. Communication Complexity. *Proc. of the 35th FOCS*, pp. 831–836, 1994.

[R] A. Razborov. The gap between the chromatic number of a graph and the rank of its adjacency matrix is superlinear. *Discrete Math.* 108, pp. 393–396, 1992.

[R-S] R. Raz and B. Spieker. On the Log-Rank conjecture in communication complexity. *Proc. of the 34th FOCS*, pp. 168–176, 1993.

[vN] C.van Nuffelen. A bound for the chromatic number of a graph. *Am. Math. Monthly* 83 (1976) 265–266.