On a Conjecture Concerning Forbidden Submatrices

R.P. Anstee*
Department of Mathematics, University of British Columbia,
#121-1984 Mathematics Road, Vancouver, B.C., Canada, V6T 1Z2.
and
Department of Mathematics and Statistics, University of Otago,
P.O. Box 56, Dunedin, New Zealand.

Abstract

Some extremal set problems can be phrased as follows. Given an $m \times n$ (0,1)-matrix A with no repeated columns and with no submatrix of a certain type, what is a bound on n in terms of m? We examine a conjecture of Frankl, Füredi, and Pach and the author that when we forbid a $k \times l$ submatrix F then n is $O(m^k)$. Two proof techniques are presented, one is amortized complexity and the other uses a result of Alon to show that n is $O(m^{2k-1-\epsilon})$ for $\epsilon = (k-1)/(13 \log_2 l)$, improving on the previous bound of $O(m^{2k-1})$.

Key words: (0,1)-matrices , forbidden configurations, forbidden submatrices.

1 Introduction

This paper considers a conjecture on forbidden submatrices. Define a matrix to be *simple* if it is a (0,1)-matrix and has no repeated columns. We remind the reader that a submatrix has the row and column order specified.

*support provided by NSERC. Work done while on sabbatical at Otago.

Conjecture 1.1 (Anstee, Frankl, Füredi, and Pach [2],[5]) *Let A be an $m \times n$ simple matrix and let F be a $k \times l$ (0,1)-matrix . Assume A has no submatrix F. Then there exists a constant c_F so that*

$$n \leq c_F m^k \qquad \blacksquare \qquad (1)$$

In [3] it was shown that n is bounded by a polynomial in m which was improved as follows, using a pigeonhole argument.

Theorem 1.2 (Frankl, Füredi, and Pach [5]) *Let A be an $m \times n$ simple matrix and let F be a $k \times l$ (0,1)-matrix . Assume A has no submatrix F. Then there exists a constant c_F so that*

$$n \leq c_F m^{2k-1} \qquad \blacksquare \qquad (2)$$

Note that (2) verifies the conjecture for $k = 1$. Section 2 considers an alternate proof technique (amortized complexity) that also handles the case $k = 1$.

The ordering of rows and columns implicit in using submatrices is what distinguishes the conjecture from the following.

Theorem 1.3 *Let A be an $m \times n$ simple matrix and let F be a $k \times l$ (0,1)-matrix . Assume A has no submatrix which is a column permutation of F. Then there exists a constant c_F so that (1) holds.*

Proof: Simply use the $O(m^k)$ bound (Füredi [5],[4]) obtained if you forbid the $k \times (l \cdot 2^k)$ submatrix of l copies of each column on k rows (an alternate proof is given at the end of Section 2). $\blacksquare$

Additional evidence supporting the conjecture can be found but until now (2) was the best bound. Section 3 uses a result of Alon [1] to obtain the bound

$$n \leq c_F m^{2k-1-\epsilon}, \qquad (3)$$

with $\epsilon = (k-1)/(13 \log_2 l)$ an improvement on (2) although still far from (1). If we somehow knew the bound was $c_F m^p$ for some integer p then we would have the conjecture for $k = 2$.

The conjecture is a useful benchmark on which to judge some of the more detailed conclusions possible when forbidding a submatrix.

2 Amortized Complexity

This section provides an elementary application of the amortized complexity idea popularized by Tarjan [9].

Theorem 2.1 *Let A be an $m \times n$ (0,1)-matrix with no consecutive pair of identical columns. Assume A has no $1 \times 2l$ submatrix $[1010 \cdots 10]$. Then*

$$n \leq (2l - 1)m + 1. \tag{4}$$

Proof: Consider submatrices [10] taken from a pair of consecutive columns. If A has $(l-1)m+1$ such [10]'s then by the pigeonhole principle there are l [10]'s in some row, one after the other, contradicting the forbidden $1 \times 2l$ submatrix.

To see how many [10]'s A must have, consider the potential function Φ : columns $\longrightarrow$ R defined as follows. Let c_i denote the ith column of A and

$$\Phi(c_i) = \text{ number of 1's in } c_i. \tag{5}$$

Now define

$$\text{RISE} = \{(i,i+1) \mid \Phi(c_i) < \Phi(c_{i+1})\}, \text{LEVEL} = \{(i,i+1) \mid \Phi(c_i) = \Phi(c_{i+1})\},$$

$$\text{FALL} = \{(i,i+1) \mid \Phi(c_i) > \Phi(c_{i+1})\}, \tag{6}$$

where $r = |\text{RISE}|, e = |\text{LEVEL}|, f = |\text{FALL}|$. Note that a level forces one [10] since consecutive columns are not identical and a fall $(i, i+1)$ forces $(\Phi(c_i) - \Phi(c_{i+1}))$ [10]'s. Now $\Phi(c_1) \geq 0$ and $\Phi(c_n) \leq m$, thus

$$\sum_{(i,i+1)\in\text{FALL}} (\Phi(c_i) - \Phi(c_{i+1})) \geq \sum_{(j,j+1)\in\text{RISE}} (\Phi(c_j) - \Phi(c_{j+1})) - m \geq r - m. \tag{7}$$

Substituting $r + e + f = n - 1$, we get

$$\sum_{(i,i+1)\in\text{FALL}} (\Phi(c_i) - \Phi(c_{i+1})) + e + f \geq n - m - 1. \tag{8}$$

Now the left hand side is at most twice the number of [10]'s (in the case $e = 0$ and $\Phi(c_i) - \Phi(c_{i+1}) = 1$ for $(i, i+1) \in$ FALL) and so A has at least $(n - m - 1)/2$ [10]'s. Thus if $n \geq (2l - 1)m + 2$, then A has $(l - 1)m + 1$ [10]'s, a contradiction. Thus (4) holds. ∎

Remarkably, there exist simple matrices A achieving the bound (4) for $m \geq \frac{1}{2}(l+3)$ as shown in [2] where the bound (4) is shown for simple matrices. The result extends to arbitrary $1 \times l$ forbidden submatrices. The task remains to generalize the proof idea to obtain Conjecture 1.1.

Here is one consequence which will appeal to Proposition 2.3 below in its proof.

Theorem 2.2 *Let β be a $(k-1) \times 1$ (0,1)-column and let α be a $1 \times l$ (0,1)-row. Define the $k \times l$ matrix F as*

$$F = \begin{bmatrix} \beta\beta\cdots\beta \\ \alpha \end{bmatrix}. \tag{9}$$

Then if A is an $m \times n$ simple matrix with no submatrix F, then there is a constant c_F so that

$$n \leq c_F m^k. \tag{10}$$

Proof: Note that forbidding the submatrix α results in a bound of $(2l-1)m+1$ (or less) by Theorem 2.1 since the $1 \times 2l$ row $[1010\cdots10]$ has α as a submatrix. Now apply Proposition 2.3 repeatedly to get (9) where if $\beta = (b_k, b_{k-1}, \ldots, b_1)$ then at the ith step apply Proposition 2.3 with 0 replaced by b_i. ■

Proposition 2.3 (Prop. 5.4 [2]) *Let F be a $k \times l$ matrix. Assume that there are constants c_F, r so that if A is an $p \times q$ simple matrix with no submatrix F, then*

$$q \leq c_F p^r. \tag{11}$$

Then there is a constant c_H so that if B is a $m \times n$ simple matrix with no $(k+1) \times l$ submatrix

$$H = \begin{bmatrix} 00\cdots0 \\ F \end{bmatrix}, \tag{12}$$

then

$$n \leq c_H m^{r+1}. \quad ■ \tag{13}$$

Our proof of Theorem 1.3 follows the proof idea of Theorem 2.1 and shows that, apart from a few columns, each column makes a 'contribution' to producing F. We state the following classic result. Let K_k denote any particular $k \times 2^k$ (0,1)-matrix consisting of all possible columns on k rows.

Theorem 2.4 (Sauer[7],Perles, Shelah[8], Vapnik, Chervonenkis[10]) *Let A be an $m \times n$ simple matrix with no submatrix which is a row and column permutation of K_k. Then*

$$n \leq \binom{m}{k-1} + \binom{m}{k-2} + \cdots + \binom{m}{0} \tag{14}$$

and (14) is best possible. ■

Proof of Theorem 1.3: We don't care about column order and so we process the columns in an appropriate order. Form $\binom{m}{k}$ 'buckets', one for each k subset of the rows. Each bucket S (where $S \subseteq \{1, 2, \ldots, m\}, |S| = k$) will keep track of contributions to F (where we view the columns of F as a multiset) in rows S from columns of A. When a bucket S has l contributions it will be considered full since then A will have a column permutation of F in rows S.

We show how to order the columns of A so that each of the first t columns contributes to a bucket for

$$t < n - \left(\binom{m}{k-1} + \binom{m}{k-1} + \cdots + \binom{m}{0} + 1\right). \tag{15}$$

Thus before $t > (l-1)\binom{m}{k} + 1$, some bucket will be full by the pigeonhole principle and so A will have a column permutation of F as a submatrix, a contradiction. The bound now follows using (15).

Given that we have ordered the columns so that the first t columns each in turn contribute to a bucket, then consider the next

$$\binom{m}{k-1} + \binom{m}{k-1} + \cdots + \binom{m}{0} + 1 \tag{16}$$

columns in A (which exist by (15)). But then we get a K_k in rows S for some k-set S (using Theorem 2.4) and so we get a contribution to bucket S from some column α. Reorder the remaining columns so that α is the $(t+1)$st column. By induction the result is true. ■

3 An application of a result of Alon

Alon [1] provided a far reaching generalization of Theorem 2.4. Let $S \subseteq \{1, 2, \ldots, m\}$ and let $A|_S$ denote the submatrix of A consisting of the rows of A indexed by S. Let $\mathcal{S}$ denote a family of subsets of $\{1, 2, \ldots, m\}$ and let $f(m, \mathcal{S})$ denote the number of (0,1)-columns α on m rows which have the property that, for each $S \in \mathcal{S}$, the column α, when restricted to the rows indexed by S, is not all 1's.

Theorem 3.1 **(Alon[1])** *Let A be an $m \times n$ simple matrix so that for each $S \in \mathcal{S}$, $A|_S$ does not have a column permutation of $K_{|S|}$ as a submatrix. Then*

$$n \leq f(m, \mathcal{S}). \quad \blacksquare \tag{17}$$

This is the (0,1)-version of Alon's result. We also need the following result.

Lemma 3.2 **([3])** *Let F be a $k \times l$ (0,1)-matrix . Then for any t satisfying*

$$t \geq 13k \log_2 l, \tag{18}$$

we have that any column permutation of K_t has F as a submatrix.

Note that the proof in [3] does not justify the claim we could take $t \geq k \cdot l$.

Theorem 3.3 *Let F be a $k \times l$ matrix ($l \geq 2$) and let A be an $m \times n$ simple matrix. Assume A has no submatrix F. Then*

$$n \text{ is } O(m^{2k-1-\epsilon}) \tag{19}$$

where $\epsilon = (k-1)/(13 \log_2 l)$.

Proof: In A we say two copies of K_k are *disjoint* if either they occupy different k-subsets of rows or, when they occupy the same set of rows, the rightmost column of one K_k is to the left of the leftmost column of the other K_k. There is a function $f(t)$ so that if we look at

$$\binom{m}{k-1} + \binom{m}{k-2} + \cdots + \binom{m}{0} + f(t) + 1 \tag{20}$$

columns of A, then we either can find a submatrix F or t copies of K_k all on different k-subsets of rows. If

$$t - 1 = \binom{s}{k} \tag{21}$$

for some $s \in \mathbf{R}$, not necessarily integral, then we can show

$$f(t) \leq \binom{s}{k} + \binom{s}{k+1} + \cdots + \binom{s}{(13k \log_2 l) - 1}. \tag{22}$$

We do this by finding the t copies of K_k one at a time using Theorem 3.1. Assume $t-1$ copies of K_k have already been found on k-subsets of rows $S_1, S_2, \ldots, S_{t-1}$. Let

$$\mathcal{S} = \{\text{all subsets of } \{1,2,\ldots,m\} \text{ of size } k \text{ or } \lceil 13k \log_2 l \rceil\} - \{S_1, S_2, \ldots, S_{t-1}\}. \tag{23}$$

Now by Theorem 3.1, $f(m,\mathcal{S})+1$ columns of A will either have a K_k on a k-subset of rows different from $S_1, S_2, \ldots, S_{t-1}$ or there will be a $K_{13k \log_2 l}$ and so will have a submatrix F. So we need an estimate for $f(m,\mathcal{S})$. Note that for $p \geq k$ a column of p 1's contributing to $f(m,\mathcal{S})$ must have all $\binom{p}{k}$ k-subsets of rows with 1's in $\{S_1, S_2, \ldots, S_{t-1}\}$. Now by the Lovász version of the Kruskal-Katona Theorem, there are at most $\binom{s}{k+1}$ $(k+1)$-subsets whose k-subsets are all in $\{S_1, S_2, \ldots, S_{t-1}\}$. Similarly there are at most $\binom{s}{k+2}$ $(k+2)$-subsets whose k-subsets are all in $\{S_1, S_2, \ldots, S_{t-1}\}$. Repeating we obtain

$$f(m,\mathcal{S}) \leq \binom{s}{k} + \binom{s}{k+1} + \cdots + \binom{s}{(13k \log_2 l) - 1} + \binom{m}{k-1} + \binom{m}{k-2} + \cdots + \binom{m}{0} \tag{24}$$

and hence we obtain (22). Now if we take $s = m^{(k-1)/(13k \log_2 l)}$ then we obtain that $f(t) \leq m^{k-1}$. So we may take

$$t = \frac{1}{k!} m^{(k-1)/(13 \log_2 l)} \tag{25}$$

and obtain (21) with '$\leq$'. Hence by (20), after

$$\binom{m}{k-1} + \binom{m}{k-2} + \cdots + \binom{m}{0} + m^{k-1} + 1 = O(m^{k-1}) \tag{26}$$

columns, we either get the submatrix F or t disjoint copies of K_k. After

$$((l-1)\binom{m}{k} + 1)/t \tag{27}$$

sets of $O(m^{k-1})$ columns, one after the other, we will either have F as a submatrix or, by the pigeonhole principle, get l mutually disjoint copies of

K_k in the same set of rows. But then we get F as a submatrix, finding the ith column of F in the ith copy of K_k. Combining (25),(26),(27) yields (19). ∎

Note that for k large, ϵ can be large yielding significant drops in the bound of Theorem 1.2. We cannot hope to directly improve the result of Theorem 3.3 since the bound (24) is real but note that we have neglected to obtain pairs of disjoint K_k in the same k-subset of rows. Lemma 3.2 can be improved. For example with $k = 2, l = 3$ we have any row and column permutation of K_8 has every 2×3 (0,1)-matrix as a submatrix. Then the bound when forbidding a specific 2×3 submatrix F drops to $O(m^{2.75})$.

References.

[1] N. Alon, On the Density of Sets of Vectors, *Discrete Math.* **46** (1983), 199-202.

[2] R.P. Anstee and Z. Füredi, Forbidden Submatrices, *Discrete Math.* **62** (1986), 225-243.

[3] R.P. Anstee, General Forbidden Configuration Theorems, *J. of Combin. Th. A* **40** (1985), 108-124.

[4] R.P. Anstee, A Forbidden Configuration Theorem of Alon, *J. of Combin. Th. A* **47** (1988), 16-27.

[5] P. Frankl, Z. Füredi and J. Pach, Bounding One-way Differences, *Eur. J. Combin.* **3** (1987), 341-347.

[6] Z. Füredi, private communication.

[7] N. Sauer, On the Density of Families of Sets, *J. Combin. Th. A* **13** (1972), 145-147.

[8] S. Shelah, A Combinatorial Problem: Stability and Order for Models and Theories in Infinitary Language, *Pac. J. Math.* **41** (1972), 247-261.

[9] R.E. Tarjan, Amortized Computational Complexity, *SIAM J. Alg. Discr. Math.* **6** (1985), 306-318.

[10] V.N. Vapnik and A.Ya. Chervonenkis, On the Uniform Convergence of Relative Frequencies of Events to their Probabilities, *Theory Prob. Applics.* **16** (1971), 264-280.