Forbidden Configurations and Indicator Polynomials

Richard Anstee
Balin Fleming
UBC, Vancouver

CanaDAM, May 25, 2009

Introduction

The use of indicator polynomials was explored in a joint paper with Fleming, Füredi and Sali. This talk focuses on joint work with Balin Fleming that led to a new bound for Forbidden Configurations. Füredi and Sali continue to explore applications to critical hypergraphs.

Forbidden Configuration Survey at www.math.ubc.ca/∼anstee

$[m] = \{1, 2, \ldots, m\}$

Let S be a finite set.

$2^S = \{T : T \subseteq S\}$

$\binom{S}{k} = \{T \in 2^S : |T| = k\}$

Definition We say that a matrix A is *simple* if it is a (0,1)-matrix with no repeated columns.

$[m] = \{1, 2, \ldots, m\}$

Let S be a finite set.

$2^S = \{T : T \subseteq S\}$

$\binom{S}{k} = \{T \in 2^S : |T| = k\}$

Definition We say that a matrix A is *simple* if it is a (0,1)-matrix with no repeated columns.

i.e. if A is an m-rowed simple matrix then A is the incidence matrix of some $\mathcal{F} \subseteq 2^{[m]}$.

Definition Given a matrix F, we say that A has F as a *configuration* if there is a submatrix of A which is a row and column permutation of F.

$$F = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad \in \quad A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

Definition Given a matrix F, we say that A has F as a *configuration* if there is a submatrix of A which is a row and column permutation of F.

$$F = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad \in \quad A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

We consider the property of forbidding a configuration F in A for which we say F is a *forbidden configuration* in A.

Definition Let $\text{forb}(m, F)$ be the largest function of m and F so that there exist a $m \times \text{forb}(m, F)$ simple matrix with *no* configuration F. Thus if A is any $m \times (\text{forb}(m, F) + 1)$ simple matrix then A contains F as a configuration.

Definition Let K_k denote the $k \times 2^k$ simple matrix of all possible columns on k rows (i.e. incidence matrix of $2^{[k]}$).

Theorem (Sauer 72, Perles and Shelah 72, Vapnik and Chervonenkis 71)

$$forb(m, K_k) = \binom{m}{k-1} + \binom{m}{k-2} + \cdots \binom{m}{0} = \Theta(m^{k-1})$$

Definition Let K_k denote the $k \times 2^k$ simple matrix of all possible columns on k rows (i.e. incidence matrix of $2^{[k]}$).

Theorem (Sauer 72, Perles and Shelah 72, Vapnik and Chervonenkis 71)

$$forb(m, K_k) = \binom{m}{k-1} + \binom{m}{k-2} + \cdots \binom{m}{0} = \Theta(m^{k-1})$$

Theorem (Füredi 83). *Let F be a $k \times l$ matrix. Then $forb(m, F) = O(m^k)$*

Definition Let K_k denote the $k \times 2^k$ simple matrix of all possible columns on k rows (i.e. incidence matrix of $2^{[k]}$).

Theorem (Sauer 72, Perles and Shelah 72, Vapnik and Chervonenkis 71)

$$forb(m, K_k) = \binom{m}{k-1} + \binom{m}{k-2} + \cdots \binom{m}{0} = \Theta(m^{k-1})$$

Theorem (Füredi 83). *Let F be a* $k \times l$ *matrix. Then* $forb(m, F) = O(m^k)$

Which F have forb(m, F) being $O(m^{k-1})$ and which F have forb(m, F) being $\Theta(m^k)$?

Let A be an m-rowed simple matrix which has no configuration K_k. For any k-set of rows $S \in \binom{[m]}{k}$ let $A|_S$ denote the submatrix of A given by the rows of S.

Since A has no K_k, then for every k-set $S \in \binom{[m]}{k}$ of rows we have that $A|_S$ has an absent $k \times 1$ (0,1)-column.

Remark If A is an m-rowed simple matrix with the property that for every k-set of rows $S \in \binom{[m]}{k}$ the submatrix $A|_S$ has an absent column, then A has no K_k and so has at most $O(m^{k-1})$ columns.

Let B be a $k \times (k+1)$ simple matrix with one column of each column sum. For $k=3$ a possible B is

$$B = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

For a matrix C, let $t \cdot C$ denote the matrix $[CCC \cdots C]$ from concatenating t copies of C. Let $F_B(t) = [K_k \ t \cdot [K_k \backslash B]]$ so for our choice of B,

$$F_B(t) = \left[\begin{matrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{matrix} \quad t \cdot \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}\right]$$

Let t be given. Let A be any m-rowed simple matrix which has no configuration $F_B(t)$. Then for any 3-set of rows $S \in \binom{[m]}{3}$, either

Let B be a $k \times (k+1)$ simple matrix with one column of each column sum. For $k = 3$ a possible B is

$$B = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

For a matrix C, let $t \cdot C$ denote the matrix $[CCC \cdots C]$ from concatenating t copies of C. Let $F_B(t) = [K_k \;\; t \cdot [K_k \backslash B]]$ so for our choice of B,

$$F_B(t) = \left[\begin{matrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{matrix} \quad t \cdot \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}\right]$$

Let t be given. Let A be any m-rowed simple matrix which has no configuration $F_B(t)$. Then for any 3-set of rows $S \in \binom{[m]}{3}$, either $A|_S$ has an absent column

Let B be a $k \times (k+1)$ simple matrix with one column of each column sum. For $k = 3$ a possible B is

$$B = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

For a matrix C, let $t \cdot C$ denote the matrix $[CCC \cdots C]$ from concatenating t copies of C. Let $F_B(t) = [K_k \; t \cdot [K_k \backslash B]]$ so for our choice of B,

$$F_B(t) = \left[\begin{matrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{matrix} \quad t \cdot \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \right]$$

Let t be given. Let A be any m-rowed simple matrix which has no configuration $F_B(t)$. Then for any 3-set of rows $S \in \binom{[m]}{3}$, either $A|_S$ has an absent column
or $A|_S$ has two columns which appear at most t times each.

Assume A is an m-rowed simple matrix with no $F_B(t)$.

Let S be a 3-set of rows in $[m]$ and let α be a 3×1 column.
We say that an $m \times 1$ column γ violates S (for the chosen α) if

$$\gamma|_S = \alpha.$$

We say 3×1 column α is
in short supply in A if it is violated by at most t columns of A.

Assume A is an m-rowed simple matrix with no $F_B(t)$.

Let S be a 3-set of rows in $[m]$ and let α be a 3×1 column.
We say that an $m\times 1$ column γ violates S (for the chosen α) if

$$\gamma|_S = \alpha.$$

We say 3×1 column α is
in short supply in A if it is violated by at most t columns of A.

Let $\mathcal{T}$ be the set of 3-sets S for which there is no absent column and hence (at least) two 3×1 columns α,β in short supply. Then by eliminating $\le t|\mathcal{T}|$ columns with violations on $S\in\mathcal{T}$ from A, we obtain a matrix which has an absent column on each 3-set of rows and so has $O(m^2)$ columns. Unfortunately $|\mathcal{T}|$ can be as big as $\Theta(m^3)$.

Multilinear Indicator Polynomials

Let $S = \{i, j, k\} \subseteq [m]$
Let $x_1, x_2, \ldots, x_m$ be variables. Let $\alpha = (\alpha_1, \alpha_2, \alpha_3)^T$ be a 3×1 (0,1)-column.

$$f_{S,\alpha}(\mathbf{x}) = (x_i - \bar{\alpha_1})(x_j - \bar{\alpha_2})(x_k - \bar{\alpha_3})$$

Multilinear Indicator Polynomials

Let $S = \{i, j, k\} \subseteq [m]$
Let $x_1, x_2, \ldots, x_m$ be variables. Let $\alpha = (\alpha_1, \alpha_2, \alpha_3)^T$ be a 3×1 (0,1)-column.

$$f_{S,\alpha}(\mathbf{x}) = (x_i - \bar{\alpha_1})(x_j - \bar{\alpha_2})(x_k - \bar{\alpha_3})$$

For a $m \times 1$ (0,1)-column γ

$$f_{S,\alpha}(\gamma) \begin{cases} \neq 0 & \text{if } \gamma|_S = \alpha \\ = 0 & \text{otherwise} \end{cases}$$

Multilinear Indicator Polynomials

$$f_{S,\alpha}(\mathbf{x}) = (x_i - \bar{\alpha_1})(x_j - \bar{\alpha_2})(x_k - \bar{\alpha_3})$$

We check that degree of $f_{S,\alpha}(\mathbf{x})$ is 3 with leading term

$$x_i x_j x_k$$

Assume $S \in \mathcal{T}$ and there are two 3×1 columns α, β in short supply (no column absent) and the two indicator polynomials are $f_{S,\alpha}$, $f_{S,\beta}$. We set

$$f_S(\mathbf{x}) = a_1 f_{S,\alpha}(\mathbf{x}) + a_2 f_{S,\beta}(\mathbf{x})$$

$$a_1 = +1, \qquad a_2 = -1$$

We have that for a $m \times 1$ (0,1)-column γ

$$f_S(\gamma) \begin{cases} \neq 0 & \text{if } \gamma|_S = \alpha \text{ or } \gamma|_S = \beta \\ = 0 & \text{otherwise} \end{cases}$$

and degree of $f_S(\mathbf{x})$ is (at most) 2 since the leading terms of degree 3 of $f_{S,\alpha}(\mathbf{x})$ and $f_{S,\beta}(\mathbf{x})$ will cancel.

Let $\mathcal{T}$ be the set of all 3-tuples S for which two columns, say α, β are in short supply.

A greedy approach would yield what we call a maximal independent set $\mathcal{I} = (S_i)$ which is an ordered list $S_1, S_2, \ldots \in \mathcal{T}$ and $m \times 1$ (0,1)-columns $\gamma_1, \gamma_2, \ldots$ of A so that

Let $\mathcal{T}$ be the set of all 3-tuples S for which two columns, say α, β are in short supply.

A greedy approach would yield what we call a maximal independent set $\mathcal{I} = (S_i)$ which is an ordered list $S_1, S_2, \ldots \in \mathcal{T}$ and $m \times 1$ (0,1)-columns $\gamma_1, \gamma_2, \ldots$ of A so that

γ_i violates S_i for one of the two chosen columns

but violates no S_j with $j < i$.

Let $\mathcal{T}$ be the set of all 3-tuples S for which two columns, say α, β are in short supply.

A greedy approach would yield what we call a maximal independent set $\mathcal{I} = (S_i)$ which is an ordered list $S_1, S_2, \ldots \in \mathcal{T}$ and $m \times 1$ (0,1)-columns $\gamma_1, \gamma_2, \ldots$ of A so that
γ_i violates S_i for one of the two chosen columns
but violates no S_j with $j < i$.
We could then delete $\leq 2t|\mathcal{I}|$ columns from A to obtain a matrix with at least one column absent on each triple of rows.

Let $\mathcal{T}$ be the set of all 3-tuples S for which two columns, say α, β are in short supply.

A greedy approach would yield what we call a maximal independent set $\mathcal{I} = (S_i)$ which is an ordered list $S_1, S_2, \ldots \in \mathcal{T}$ and $m \times 1$ (0,1)-columns $\gamma_1, \gamma_2, \ldots$ of A so that
γ_i violates S_i for one of the two chosen columns
but violates no S_j with $j < i$.
We could then delete $\leq 2t|\mathcal{I}|$ columns from A to obtain a matrix with at least one column absent on each triple of rows.

Theorem *If $\mathcal{I} = (S_i)$ is an independent set, then the indicator polynomials f_S are linearly independent.*

Theorem *If* $\mathcal{I} = (S_i)$ *is an independent set, and the indicator polynomials* f_S *are degree at most* 2 *then*

$$|\mathcal{I}| \leq \binom{m}{2} + \binom{m}{1} + \binom{m}{0}$$

$$= \Theta(m^2).$$

Theorem forb$(m, F_B(t))$ is $\Theta(m^2)$.

Proof: Assume A is a matrix with no $F_B(t)$. Let $\mathcal{I} = (S_i)$ be a maximal independent set with indicator polynomials f_{S_i}. For a given set $S \in \mathcal{I}$ there are at most $2t$ columns with violations of the two chosen columns in short supply on S. By our linear algebra, $|\mathcal{I}|$ is $O(m^2)$. Thus we may remove $2t|\mathcal{I}|$ or $O(m^2)$ columns and remove all violations on the two chosen 3×1 columns for each $S \in \mathcal{I}$ and so on each $S \in \mathcal{T}$ there will be an absent 3×1 column. The resulting matrix has at most $O(m^2)$ columns and so A has at most $O(m^2)$ columns.

We have one more 3-rowed configuration F with $\text{forb}(m, F)$ being $O(m^2)$. Let D be the 3×5 simple matrix with all columns of sum at least 1 that do not simultaneously have 1's in rows 1 and 2. We take $F_D(t) = [\mathbf{0}_3 \, (t+1) \cdot D]$ which becomes

$$F_D(t) = \left[\begin{matrix} 0 \\ 0 \\ 0 \end{matrix} \; (t+1) \cdot \left[\begin{matrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \end{matrix} \right] \right]$$

If a matrix A has no $F_D(t)$ then each 3-set of rows $\{i, j, k\}$ in some ordering has one of the following occur:

We have one more 3-rowed configuration F with $\text{forb}(m, F)$ being $O(m^2)$. Let D be the 3×5 simple matrix with all columns of sum at least 1 that do not simultaneously have 1's in rows 1 and 2. We take $F_D(t) = [\mathbf{0}_3 \, (t+1) \cdot D]$ which becomes

$$F_D(t) = \left[\begin{array}{cc} 0 & \\ 0 & (t+1) \cdot \left[\begin{array}{ccccc} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \end{array} \right] \\ 0 & \end{array} \right]$$

If a matrix A has no $F_D(t)$ then each 3-set of rows $\{i, j, k\}$ in some ordering has one of the following occur:

	no
i	0
j	0
k	0

We have one more 3-rowed configuration F with $\text{forb}(m, F)$ being $O(m^2)$. Let D be the 3×5 simple matrix with all columns of sum at least 1 that do not simultaneously have 1's in rows 1 and 2. We take $F_D(t) = [\mathbf{0}_3 \, (t+1) \cdot D]$ which becomes

$$F_D(t) = \left[\begin{array}{cc} 0 & \\ 0 & (t+1) \cdot \\ 0 & \end{array} \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \end{bmatrix} \right]$$

If a matrix A has no $F_D(t)$ then each 3-set of rows $\{i, j, k\}$ in some ordering has one of the following occur:

$$\begin{array}{cc} & \text{no} \\ i & 0 \\ j & 0 \\ k & 0 \end{array}$$ or at least two columns are in short supply

We have one more 3-rowed configuration F with $\text{forb}(m, F)$ being $O(m^2)$. Let D be the 3×5 simple matrix with all columns of sum at least 1 that do not simultaneously have 1's in rows 1 and 2. We take $F_D(t) = [\mathbf{0}_3 \, (t+1) \cdot D]$ which becomes

$$F_D(t) = \left[\begin{matrix} 0 \\ 0 \\ 0 \end{matrix} \; (t+1) \cdot \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \end{bmatrix} \right]$$

If a matrix A has no $F_D(t)$ then each 3-set of rows $\{i, j, k\}$ in some ordering has one of the following occur:

$$\begin{matrix} & \text{no} \\ i & 0 \\ j & 0 \\ k & 0 \end{matrix} \quad \text{or at least two columns are in short supply or} \quad \begin{matrix} \leq t \\ 0 \\ 0 \\ 1 \end{matrix} .$$

The case of one column in short supply (but not absent) makes the proof much more difficult. We can eliminate $O(m^2)$ columns and make a column absent on each 3-set of rows but the argument is more complex. The new proof idea, using indicator polynomials, was to consider what is in short supply on some 4-sets of rows and miraculously we are able to find indicator polynomials of degree 2 (we are able to cancel the terms of degree 4 and 3).

A typical situation if we avoid $F_D(t)$ could be:

	$\leq t$	$\leq t$	$\leq t$	$\leq t$	no
i	0	0	1	0	
j	0	0			0
k	1		0	1	0
l		1	1	1	0

Note that

	$\leq t$
i	1
j	
k	0
l	1

$\Rightarrow$

	$\leq t$	$\leq t$
i	1	1
j	1	0
k	0	0
l	1	1

We only need a few of these 4×1 columns to get our reduction in degree.

	$+1$	-1	$+1$	-1
i	0	1	1	0
j	0	0	0	0
k	0	0	0	0
l	1	1	0	0
	α	β	γ	δ

We form a new indicator polynomial for the 4 columns $\alpha, \beta, \gamma, \delta$ above as

$$f_S(\mathbf{x}) = +1 f_{S,\alpha}(\mathbf{x}) - 1 f_{S,\beta}(\mathbf{x}) + 1 f_{S,\gamma}(\mathbf{x}) - 1 f_{S,\delta}(\mathbf{x})$$

and we find that f_S is a degree 2 indicator polynomial for the 4 columns above.

Theorem *Let k, t be given positive integers with $k \geq 2$, $t \geq 1$. Let B as a $k \times (k+1)$ simple matrix with one column of each column sum And let $F_B(t) = [K_k \; t \cdot [K_k \backslash B]]$. Then*

$$forb(m, F_B(t)) \text{ is } \Theta(m^{k-1})$$

Let D be the $k \times (2^k - 2^{k-2} - 1)$ simple matrix with all columns of sum at least 1 that do not simultaneously have 1's in rows 1 and 2. We take $F_D(t) = [\mathbf{0}_k \; (t+1) \cdot D]$ Then

$$forb(m, F_D(t)) \text{ is } \Theta(m^{k-1})$$

Theorem *Let k, t be given positive integers with $k \geq 2$, $t \geq 1$. Let B as a $k \times (k+1)$ simple matrix with one column of each column sum And let $F_B(t) = [K_k\; t \cdot [K_k \backslash B]]$. Then*

$$forb(m, F_B(t)) \text{ is } \Theta(m^{k-1})$$

Let D be the $k \times (2^k - 2^{k-2} - 1)$ simple matrix with all columns of sum at least 1 that do not simultaneously have 1's in rows 1 and 2. We take $F_D(t) = [\mathbf{0}_k\; (t+1) \cdot D]$ Then

$$forb(m, F_D(t)) \text{ is } \Theta(m^{k-1})$$

Theorem Let F is a k-rowed configuration which is not a configuration in $F_B(t)$ (for any choice of B as a $k \times (k+1)$ simple matrix with one column of each column sum and for any t) and not in $F_D(t)$ (for any t). Then $\text{forb}(m, F)$ is $\Theta(m^k)$.

Merci/Thanks to the organizers for arranging this conference!